Functional Inequalities in Open and Closed Quantum Systems: Continuity, Correlations, and Applications

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Paul Elias Gondolf

aus Pforzheim

Tübingen

2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	07.10.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatterin:	Asst. Prof. Dr. Ángela Capel Cuevas
2. Berichterstatter:	Prof. Dr. Marius Lemm
3. Berichterstatter:	Prof. Dr. Michael J. Kastoryano

Für meine Familie
Friedrich, Joscha, Ulrike und Klaus

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Ángela Capel and Cambyse Rouzé, for their guidance, support, and encouragement throughout the course of this work. Their insights and mentorship have been invaluable, and their unwavering support for my academic career and opportunities in the field has been truly amazing. I could not have wished for two better supervisors.
This thesis has also greatly benefited from the many collaborators I had the pleasure of working with during this time. Their ideas and input helped shape the results in meaningful ways. In particular, special thanks go to Tim Möbus, who not only contributed as a collaborator but also offered thoughtful advice and careful proofreading. It has been a pleasure—and more importantly, a lot of fun—to work with him since the very beginning.
My sincere appreciation extends to the researchers and groups who hosted me during research visits. The M5 group at TUM Munich and the Quriosity group at Télécom Paris provided warm hospitality and a stimulating scientific environment, for which I am truly grateful.
Warm thanks are due to the Department of Mathematical Physics at the University of Tübingen—my home institution—and especially to the members of the quantum-cooking-group, whose nourishing meals sustained both body and mind throughout this journey. I would also like to thank Elena Kabagema-Bilan, the good soul of the department, whose support and kindness helped me (and many others) navigate the often rough waters of university administration.
Heartfelt thanks also go to my office mates, Karla Schön and David Seez, for their camaraderie and good humour, which made even the most challenging days more enjoyable. Their company and motivation—even on the rainiest of days—made all the difference.

Schliesslich gilt mein tief empfundener Dank meiner Familie: meinen Brüdern Joscha und Friedrich sowie meinen Eltern Ulrike und Klaus für ihre beständige Unterstützung, ihre Geduld und ihren unerschütterlichen Rückhalt während all dieser Zeit.
Ebenso danke ich meinen Freundinnen und Freunden aus Tübingen, sowohl aus meiner WG und darüber hinaus, sowie aus der Heimat – einige dort geblieben, die meisten längst in alle Himmelsrichtungen verstreut – und ganz besonders den PTFL ($\subset$ FRTFL) für die gemeinsamen Momente, die Unterstützung und die über Jahre gewachsene Verbundenheit.
Zuletzt gilt mein innigster Dank meiner Freundin Farah, die in den letzten drei Jahren an meiner Seite und mir stets eine Stütze war und mich durch alle Höhen und Tiefen begleitet hat.

Zusammenfassung

Diese Arbeit fasst die Resultate aus sechs Projekten zusammen, die sich thematisch zwischen Quanteninformationstheorie und Quanten-Vielteilchensystemen bewegen. Die ersten beiden Projekte beschäftigen sich mit der Herleitung von Stetigkeitsabschätzungen für Entropiefunktionale.

Konkret wird im ersten Projekt eine obere Schranke für die Differenz der von-Neumann-Entropie zweier Quantenzustände ρ und ρ' hergeleitet. Diese Schranke basiert auf der Entropiedifferenz der Jordan-Hahn-Zerlegung von $\rho - \rho'$, ergänzt um einen additiven Term, der die binäre Entropie ihrer Spurdistanz enthält. Aus dieser Ungleichung lässt sich nicht nur die bekannte Audenaert-Fannes-Ungleichung herleiten, sondern erstere verbessert letztere sogar. Als weitere Konsequenzen ergeben sich Stetigkeitsabschätzungen für die bedingte Entropie (CE) zweier Zustände mit übereinstimmenden Randverteilungen auf dem bedingten System sowie für die Umegaki-Relativentropie.

Das zweite Projekt behandelt Stetigkeitsabschätzungen von Funktionalen der Form $\rho \mapsto \inf_{\sigma\in\mathcal{C}} \widetilde{D}_\alpha(\rho\|\sigma)$, wobei $\widetilde{D}_\alpha$ die „sandwiched" Rényi-Divergenz bezeichnet und $\mathcal{C}$ eine konvexe, kompakte Menge von Zuständen ist, die mindestens ein positiv definites Element enthält. Aufbauend auf Arbeiten von Marwah and Dupuis sowie unabhängig davon von Beigi and Goodarzi zur „sandwiched" Rényi-bedingten Entropie, verbessern wir deren Methoden und verallgemeinern den Kontext im obigen Sinne. Ein Ansatz nutzt die Super- und Subadditivität sowie die Konvexität und Konkavität der Exponentialfunktion der Divergenz; der andere formuliert das Exponential der Divergenz als Norm in einem Interpolationsraum. Da sich die Ergebnisse der Interpolationstheorie jedoch nicht unmittelbar auf allgemeine konvexe und kompakte Mengen übertragen lassen, greifen wir zur Beweisführung der benötigten Normeigenschaften nicht auf die abstrakte Theorie zurück, sondern verwenden lediglich den Konkavitätssatz von Lieb sowie das Minimax-Prinzip von Sion. Als Konsequenz der Stetigkeitsschranken für die oben genannten Abbildungen ergibt sich unter anderem eine allgemeine Schranke für die Stetigkeit der „sandwiched" Rényi-Divergenz selbst.

Im dritten Projekt zeigen wir einen superexponentiellen Abfall der Belavkin-Staszewski'schen bedingten wechselseitigen Information (CMI) für Gibbs-Zustände eindimensionaler, lokaler, translationsinvarianter Wechselwirkungen bei beliebig positiver Temperatur. Mit der Eigenschaft dieser Größe als Fehlerabschätzung eines Rekonstruktionsschritts können wir beliebige Randverteilungen solcher Gibbs-Zustände mittels Matrixproduktoperatoren (MPO) approximieren. Diese Approximation hat eine subpolynomielle Bond-Dimension N/ε, wobei N die Größe der Randverteilung und ε den Fehler in der Spurnorm bezeichnet. Durch Kombination von lokaler Tomografie mit dieser MPO-Rekonstruktion ergibt sich abschließend ein effizienter Algorithmus zur Rekonstruktion solcher Randverteilungen aus Messdaten mit Laufzeit und Stichprobenkomplexität, die polynomiell in N/ε sind.

Das vierte Projekt untersucht die Mischungszeit der Davies-Halbgruppe für lokale, kommutierende Wechselwirkungen bei positiver Temperatur. Es wird gezeigt, dass sich durch die Kombination einer einheitlichen unteren Schranke an die lokalen Spektrallücken mit dem Abfall einer matrixwertigen bedingten wechselseitigen Information (MCMI) im entsprechenden Gibbs-Zustand eine exponentiell bessere Schranke in der Systemgröße für die Mischungszeit im Vergleich zur naiven Verwendung der Spektrallückenannahme ergibt. Darüber hinaus führt diese verbesserte Mischungszeit zu einer exponentiell verbesserten globalen Spektrallücke sowie zu einem Schrumpfungskoeffizienten für die Entropie der Halbgruppe für große und diskrete Zeiten.

Im fünften Projekt führen wir die quantenmechanischen Sobolev-Räume sowie soboleverhaltende Halbgruppen auf einem bosonischen Einmodensystem ein. Unter Verwendung dieser Theorie zeigen wir, dass ein unbeschränkter Operator, der einerseits (i) auf der Menge der endlich-rangigen Operatoren in der Fock-Basis eine Gorini-Kossakowski-Sudarshan-Lindblad (GKSL)-Form besitzt – wobei seine Komponenten Polynome in Erzeugungs- und Vernichtungsoperatoren sind – und andererseits (ii) die Momente des Zahloperators auf dieser Menge kontrolliert erhöht, ein Kern eines Generators einer Quanten-Markov-Halbgruppe ist, die zudem die quantenmechanischen Sobolev-Räume erhält. Je nach Stärke der Annahme (ii) zeigen wir darüber hinaus zeitunabhängige Schranken für die Sobolev-Normen der Halbgruppen und sogar eine

Regularisierungseigenschaft, wonach auch Zustände mit anfangs unendlicher Sobolev-Norm für positive Zeitentwicklung endlich normiert werden. Diese Theorie wird auf konkrete Beispiele angewendet, darunter eindimensionale gaußsche Halbgruppen und Generatoren, die Katzen-Code-Gatter modellieren; mittels der Theorie der quantenmechanischen Sobolev-Räume und soboleverhaltender Halbgruppen untersuchen wir zusätzlich deren Störungen.
Das sechste und letzte Projekt beweist das verallgemeinerte Quanten-Stein'sche Lemma für Subalgebra-Ressourcen. Das bedeuted die Alternativhypothese im n-ten Test ist durch die Fixpunkte eines n-fachen Tensorprodukts des Hilbert-Schmidt (HS)-Duals einer bedingten Erwartung auf eine von-Neumann-Unteralgebra gegeben. Aufbauend auf der Arbeit von Gao and Rahaman, die das Lemma für HS-symmetrische bedingte Erwartungen bewiesen haben, verallgemeinern wir deren Strategie und zeigen die Aussage für beliebige bedingte Erwartungen, vorausgesetzt, dass deren HS-Dual einen positiv definiten Fixpunkt besitzt.

Abstract

This thesis summarises and in some instances extends results arising from six projects situated at the intersection of quantum information theory and many-body systems. The first two projects concern the derivation of continuity bounds for entropy functionals.

More specifically, the first project establishes a tight upper bound for the difference in von Neumann entropy between two quantum states ρ and ρ'. This bound is expressed in terms of the entropy difference of the Jordan-Hahn decomposition of $(\rho - \rho')$, corrected by a binary entropy term involving their trace distance (TD). The resulting inequality not only recovers but strengthens the well-known Audenaert-Fannes inequality. As further consequences, the project yields continuity bounds for the conditional entropy (CE) when the marginal on the conditioning system agrees, as well as for the Umegaki relative entropy.

The second project concerns continuity bounds for functionals of the form $\rho \mapsto \inf_{\sigma \in \mathcal{C}} \widetilde{D}_\alpha(\rho \| \sigma)$, where $\widetilde{D}_\alpha$ is the sandwiched Rényi (SR) divergence and $\mathcal{C}$ a convex, compact set of quantum states, that contains at least one positive-definite element. Such quantities naturally arise in quantum resource theories as resource monotones. Building upon earlier work by Marwah and Dupuis and independently Beigi and Goodarzi on the SR-CE, we improve and generalise their strategies to this broader context. One approach uses the super/subadditivity and convexity/concavity of the exponential of the divergence; the other rewrites the divergence as a norm in an interpolation space. However, the existing scope of interpolation theory is insufficient for establishing the necessary norm properties. Thus, instead of relying on abstract theory, we base our analysis on Lieb's concavity theorem and Sion's minimax principle. As a consequence of these continuity bounds, we also obtain a general continuity estimate for the SR divergence itself.

The third project establishes a superexponential decay of the Belavkin-Staszewski (BS)-conditional mutual information (CMI) for Gibbs states of one-dimensional, local, translation-invariant interactions at arbitrary positive temperature. Using that this quantity forms an upper bound to the recovery error of a reconstruction map, we construct a matrix product operator (MPO) approximation of marginals of such Gibbs states. This approximation achieves a bond dimension subpolynomial in N/ε, where N is the size of the marginal and ε the TD reconstruction error. By combining local tomography with this MPO reconstruction, the project culminates in an efficient algorithm for recovering such marginals from measurement data, with runtime and sample complexity polynomial in N/ε.

The fourth project investigates the mixing time of the Davies semigroup for local, commuting interactions at positive temperature. We show that combining a uniform lower bound on local gaps with the decay of what we term the matrix-valued conditional mutual information (MCMI) in the corresponding Gibbs state yields an exponentially improved mixing time estimate compared to relying on the uniform bound on local gaps alone. Furthermore, this improved mixing time leads to an exponentially enhanced global spectral gap and establishes a large-time entropy contraction coefficient for the semigroup.

The fifth project develops the framework of quantum Sobolev space (QSS) and Sobolev-preserving semigroups on a single-mode bosonic system. Specifically, we show that if an unbounded operator (i) takes Gorini-Kossakowski-Sudarshan-Lindblad (GKSL)-form on the finite-rank operators in the Fock basis, with constituents polynomial in creation and annihilation operators, and (ii) increases the moments of the number operator in a controlled manner on this domain, then it is the core of a generator to a quantum Markov semigroup (QMS) that additionally preserves QSS. Depending on the strength of condition (ii), we further establish time-independent Sobolev norm bounds and even regularisation properties, whereby initially irregular states acquire finite Sobolev norms after any positive evolution time. Applying these results to concrete examples—including single-mode Gaussian semigroups and generators modelling cat code gates—we demonstrate the utility of the framework in generation and further perturbation theory.

The sixth and final project proves a generalised quantum Stein's lemma for subalgebra resources. This means the alternative hypothesis in the nth test is given by the fixed points of an n-fold tensor product of the Hilbert-Schmidt (HS)-adjoint of a conditional expectation onto a von Neumann subalgebra. Building on the work of Gao and Rahaman, who proved the lemma for HS-symmetric conditional expectations, we

generalise their approach and prove the result for arbitrary conditional expectations whose HS-dual has a positive-definite fixed point.

List of Publications

(a) Published articles

1. **Continuity bounds for quantum entropies arising from a fundamental entropic inequality**
   Koenraad Audenaert, Bjarne Bergh, Nilanjana Datta, Michael G. Jabbour, Ángela Capel, Paul Gondolf
   Published in *IEEE Transactions on Information Theory* (2025).
   Cited in the following as [Aud+25] and included in section A.1.
2. **Unified framework for continuity of sandwiched Rényi divergences**
   Andreas Bluhm, Ángela Capel, Paul Gondolf, Tim Möbus
   Published in *Annales Henri Poincaré* (2024).
   Cited in the following as [Blu+24] and included in section A.2.
3. **Energy preserving evolutions over Bosonic systems**
   Paul Gondolf, Tim Möbus, Cambyse Rouzé
   Published in *Quantum* (2024).
   Cited in the following as [GMR24] and included in section A.3.

(b) Submitted articles (available as preprints)

1. **Conditional independence of 1D Gibbs states with applications to efficient learning**
   Álvaro M. Alhambra, Ángela Capel, Paul Gondolf, Alberto Ruiz-de-Alarcón, Samuel O. Scalet
   Preprint, arXiv:2402.18500.
   Cited in the following as [Alh+24] and included in section B.1.
2. **Quasi-optimal sampling from Gibbs states via non-commutative optimal transport metrics**
   Ángela Capel, Paul Gondolf, Jan Kochanowski, Cambyse Rouzé
   Preprint, arXiv:2412.01732.
   Cited in the following as [Cap+24] and included in section B.2.

(c) Not submitted (rewritten and included)

1. **The generalised quantum Stein's lemma for subalgebra entropies**
   Paul Gondolf, Cambyse Rouzé
   Referenced in the following as [GR24b] and fully included in the main body of the thesis.

Contents

Acknowledgements **v**

Zusammenfassung **vii**

Abstract **ix**

List of Publications **xi**

1 Introduction **1**
1.1 Overview 1
1.2 Basic notation 3
1.3 General preliminaries 5
1.3.1 Entropy functionals 5
1.3.2 Von Neumann subalgebras of matrices and their conditional expectations 7
1.3.3 A short introduction to semigroup theory for quantum systems 11
1.3.4 Hamiltonians on lattice systems 20
1.4 Project specific preliminaries 23
1.4.1 Continuity bounds for some entropy functionals ([Aud+25] and [Blu+24]) 23
1.4.2 Reconstruction of Gibbs states on spin chains ([Alh+24]) 24
1.4.3 Convergence analysis for the Davies semigroup ([Cap+24]) 29
1.4.4 A short introduction to open bosonic systems ([GMR24]) 36
1.4.5 The generalised quantum Stein's lemma for subalgebra resources ([GR24b]) 39

2 Objectives **43**
2.1 Objectives of [Aud+25] and [Blu+24] 43
2.2 Objectives of [Alh+24] 44
2.3 Objectives of [Cap+24] 44
2.4 Objectives of [GMR24] 46
2.5 Objectives of [GR24b] 46

3 Results and Discussion **47**
3.1 Results and discussion of [Aud+25] and [Blu+24] 47
3.1.1 A novel bound for the von Neumann entropy and its consequences 47
3.1.2 Continuity bounds for entropy functionals derived from the sandwiched-Rényi divergence 49
3.1.3 Open questions and future work 52
3.2 Results and discussion of [Alh+24] 52
3.2.1 Upper and lower bounds on the BS-DPI 53
3.2.2 Superexponential decay of the BS-CMI on local, translation-invariant spin chains . 54
3.2.3 A MPO reconstruction of Gibbs states on local, translation-invariant spin chains . 55
3.2.4 Open questions and future work 57
3.3 Results and discussion of [Cap+24] 58
3.3.1 Additive approximate tensorisation of conditional relative entropy 58
3.3.2 The partition of the lattice 59
3.3.3 A mixing time improvement from uniform local gap and MCMI decay 61
3.3.4 From mixing time estimates to entropy contraction coefficients and gap 64
3.3.5 Open questions and future work 67

3.4 Results and discussion of [GMR24] . . . 68
3.4.1 Generation of Sobolev-preserving semigroups . . . 69
3.4.2 Improved estimates on norms of the semigroups . . . 73
3.4.3 Examples and discussion of perturbation theory . . . 75
3.4.4 Open questions and future work . . . 79
3.5 Results and discussion of [GR24b] . . . 80
3.5.1 Open questions and future work . . . 84

A Published Articles **95**
A.1 Article: Audenaert et al. (2025) . . . 96
A.2 Article: Bluhm et al. (2024) . . . 106
A.3 Article: Gondolf et al. (2024) . . . 156

B Submitted articles **225**
B.1 Article: Alhambra et al. (2024) . . . 226
B.2 Article: Capel et al. (2024) . . . 271

List of Acronyms

C_0-semigroup strongly continuous semigroup. 17–20, 36, 38, 39, 69–71

AFW Alicki-Fannes-Winter. 23, 24

BS Belavkin-Staszewski. ix, xviii, 2, 6, 28, 29, 44, 52–54

CE conditional entropy. ix, xviii, 2, 5–7, 23, 24, 43, 47–49, 51, 52

CMI conditional mutual information. ix, xviii, 2, 5, 6, 21, 24, 28, 29, 33, 44, 52, 53, 59

CMLSI complete modified logarithmic Sobolev inequality. xviii, 2, 14, 15, 17, 22, 31, 32, 34, 35, 66–68

CP completely positive. 4, 5, 8, 11, 12, 19, 20, 26, 69–72, 75

CPTP completely positive and trace preserving. 4–6, 10–12, 18, 27, 47, 66, 82

CSS Calderbank-Shor-Steane. 22, 29, 34, 64, 68

DB detailed balanced. 14

DPI data processing inequality. 6, 10, 14, 26, 28, 35, 40, 44, 53, 58, 66, 82

EP entropy production. 16

GKSL Gorini-Kossakowski-Sudarshan-Lindblad. ix, xviii, 2, 12, 13, 16, 18, 36, 68, 75

GNS Gelfand-Naimark-Segal. xvii, 2, 8, 13, 14, 31, 33, 34, 36, 66, 67

GR geometric Rényi. 6, 52, 82

HS Hilbert-Schmidt. ix, xvii, 4, 5, 8–10, 12–14, 16, 30, 31, 35, 41, 42, 44, 52–54, 65, 68, 76, 80, 81

JH Jordan-Hahn. 2, 24, 43, 47–50, 66

KL Kullback-Leibler. 5, 6

KMS Kubo-Martin-Schwinger. xvii, 2, 9, 13, 14, 17, 31, 33, 64, 65, 67, 68

LSI logarithmic Sobolev inequality. 15, 16

MCMI matrix-valued conditional mutual information. ix, xviii, 2, 59, 61, 63, 64, 67

MI mutual information. xviii, 5–7, 21, 24, 51

MLSI modified logarithmic Sobolev inequality. xviii, 2, 14–17, 22, 32, 34, 35, 44, 45, 58, 64, 66–68, 79

MPDO matrix product density operator. 25, 27, 28, 58

MPO matrix product operator. ix, 2, 22, 25, 27, 29, 44, 52–58

OU Ornstein-Uhlenbeck. 36, 38, 46, 76, 77

PDE partial differential equation. 1

PR Petz-Rényi. 6, 52, 82

QECC quantum error correction code. 29

QMS quantum Markov semigroup. ix, 2, 3, 11–18, 36–39, 46, 64–66, 68, 72, 74, 76–79

QSS quantum Sobolev space. ix, xvii, xviii, 3, 17, 38, 39, 69–72

SR sandwiched Rényi. ix, xviii, 2, 6, 24, 43, 49, 51, 52, 82

SSA strong subadditivity. 1, 35, 58

TD trace distance. ix, xviii, 14, 23, 26, 44, 52, 58, 66

List of Symbols

Symbol	**Description**
A. Fundamental sets and structures	
$\mathbb{N}\,(\mathbb{N}_0), \mathbb{Z}, \mathbb{R}, \mathbb{C}$	Natural (with zero), integer, real and imaginary numbers.
$\mathbb{R}_+$	Non-negative real numbers.
$\mathbb{R}[x], \mathbb{C}[x], \mathbb{R}[x,y], \mathbb{C}[x,y]$	Real and complex polynomials in one and two (non-commutative) variables respectively.
$\lvert S\rvert$	Cardinality (number of elements) of the set $S \subset \mathbb{C}$.
$\Lambda \Subset \mathbb{Z}^D$	Λ is a finite subset of $\mathbb{Z}^D$ (see section 1.3.4).
$A \sqcup B$	Disjoint union of $A, B \subseteq \mathbb{Z}^D$.
$\partial A, A\partial$	Set boundary relative to Λ, union with set boundary (see eq. (1.76)).
$\mathrm{dist}(x,y), \mathrm{dist}(A,B)$	Euclidean distance between $x, y \in \mathbb{Z}^D$, $A, B \subset \mathbb{Z}^D$ (see section 1.3.4).
B. Hilbert spaces, operator spaces, and operators	
$\mathcal{H}, \mathcal{H}_{AB}, \mathcal{H}_{ABC}, \ldots, \mathcal{K}$	Finite-dimensional (multipartite) Hilbert spaces.
$\mathcal{F}$	Hilbert space of square integrable functions from $\mathbb{R}$ to $\mathbb{C}$ ($L^2(\mathbb{R})$).
$\mathcal{X}$	Banach space.
$\overline{\mathcal{A}}$	Topological closure of $\mathcal{A} \subseteq \mathcal{X}$ in the Banach space $\mathcal{X}$.
$\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathcal{F}), \ldots$	Bounded linear operators on a finite/infinite-dimensional Hilbert space.
$\mathcal{T}(\mathcal{F})$	Trace-class operators on $\mathcal{F}$ (see eq. (1.3)).
$\mathcal{T}_f(\mathcal{F})$	Operators that have finite rank in the Fock basis (see eq. (1.93)).
$\mathcal{W}^k$	QSS of order k (see theorem 1.4.1).
$\mathcal{B}(\mathcal{X})$	Bounded linear operators on the Banach space $\mathcal{X}$.
$\mathcal{B}_{\mathbb{Z}^D}$	Algebra of quasi local observables (see eq. (1.45)).
$\mathcal{N}, \mathcal{M}$	Von Neumann subalgebras (see section 1.3.2).
$\mathcal{A}'$	Commutant of the set of operators $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ (see eq. (1.18)).
$a, a^*, \mathbf{N}$	Annihilation, creation, and number operator on $\mathcal{F}$.
$\mathbb{1}$	Identity on $\mathcal{H}$ or $\mathcal{F}$.
$\iota, \iota_A, \ldots$	The maximally mixed state on $\mathcal{H}, \mathcal{H}_A, \ldots$, i.e., $\mathbb{1}_{\mathcal{H}}/d_{\mathcal{H}}, \mathbb{1}_A/d_A, \ldots$
id	Identity on $\mathcal{B}(\mathcal{H})$ or $\mathcal{B}(\mathcal{F})$.
$(\mathcal{O}, \mathcal{D}(\mathcal{O}))$	(Unbounded) linear operator on a Banach space.
$(\overline{\mathcal{O}}, \mathcal{D}(\overline{\mathcal{O}}))$	Closure of a closable $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ (see section 1.2).
X^*	Hilbert space adjoint of X.
$\mathcal{O}^\dagger$	HS-adjoint of $\mathcal{O}$.
$R(\lambda, \mathcal{O})$	Resolvent of $\mathcal{O}$ for $\lambda \in \mathbb{C}$ in its resolvent set (see section 1.3.3).
C. Operations on operators	
Tr	Trace.
tr_A	Partial trace in the system A, i.e., $\mathcal{H}_A$.
$\bigcirc_{n=1}^{N} \mathcal{O}_n = \mathcal{O}_N \ldots \mathcal{O}_1$	Ordered composition (w.r.t. order on $\mathbb{Z}$) of linear maps.
$\langle \psi, \phi \rangle$	Hilbert space inner product on $\mathcal{H}$ or $\mathcal{F}$.
$\langle X, Y \rangle_{\mathrm{HS}}$	HS-inner product (see eq. (1.6)).
$\langle X, Y \rangle_\sigma$	Gelfand-Naimark-Segal (GNS)-inner product (see eq. (1.30)).
$\langle X, Y \rangle_{\sigma,1/2}$	Kubo-Martin-Schwinger (KMS)-inner product (see eq. (1.29)).
D. Norms	
$\lVert\cdot\rVert_{\mathcal{H}}, \lVert\cdot\rVert_{\mathcal{F}}$	Hilbert space norms.
$\lVert\cdot\rVert_{\mathcal{X}}$	Banach space norm.
$\lVert\cdot\rVert_\infty$	Norm on $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{F})$ respectively.

$\|\cdot\|_{\mathcal{B}(\mathcal{X})}$	Norm on bounded operators on a Banach space (see eq. (1.2)).
$\|\cdot\|_p$	Schatten p-norm (see eq. (1.4)).
$\|\cdot\|_{\mathcal{W}^k}$	Norm on $\mathcal{W}^k$ (see theorem 1.4.1).
$\|\cdot\|_{\mathcal{W}^{k'}\to\mathcal{W}^k}$	Norm on linear operators between QSS (see eq. (1.98)).
$\|\cdot\|_{1\to 1}$	Norm on operators between $\mathcal{T}(\mathcal{F})$ and itself or $\mathcal{B}(\mathcal{H})$ and itself respectively (see eq. (1.7))
E. Quantum states, dynamics, and related concepts	
$\mathcal{S}(\mathcal{H})$	Quantum states on $\mathcal{H}$.
$H_\Lambda, H_A, \dots$	Local Hamiltonian (see eq. (1.47)).
$\mathcal{L}$	Lindbladien on $\mathcal{B}(\mathcal{H})$, $\mathcal{D}(\mathcal{L}) \subseteq \mathcal{T}(\mathcal{F})$ or $\mathcal{D}(\mathcal{L}) \subseteq \mathcal{W}^k$ in GKSL-form and Schrödinger picture.
$\mathcal{L}_\Lambda, \mathcal{L}_A, \dots$	Global, and local Davies Lindbladien (see eqs. (1.74), (1.79)).
$\lambda(\mathcal{L}), \alpha(\mathcal{L}), \alpha_c(\mathcal{L})$	Gap, modified logarithmic Sobolev inequality (MLSI), complete modified logarithmic Sobolev inequality (CMLSI) of $\mathcal{L}$ (see eqs. (1.31), (1.38), and (1.40)).
$t_{\mathrm{mix}}(\mathcal{L};\varepsilon)$	Mixing time (see eq. (1.33)).
F. Information theoretic measures	
$S(\rho)$	Von Neumann entropy (see eq. (1.8)).
$T(\rho,\sigma)$	TD (see eq. (1.32)).
$D(\rho\\|\sigma)$	Umegaki relative entropy (see eq. (1.10)).
$\widehat{D}(\rho\\|\sigma)$	BS entropy (see eq. (1.12)).
$\widetilde{D}_\alpha(\rho\\|\sigma), \widetilde{Q}_\alpha(\rho,\sigma)$	SR divergence and its constituting functional (see eq. (1.13)).
$D_h^\varepsilon(\rho\\|\sigma)$	Hypothesis testing relative entropy (see eq. (1.100)).
$\mathbb{D}(\rho\\|\sigma)$	General divergence (see section 1.3.1).
$\mathbb{D}(\rho\\|\,\mathcal{C})$	Divergence distance to a set of states $\mathcal{C}$ (see eq. (1.16)).
$D_A(\rho\\|\sigma)$	Conditional relative entropy on $A \subseteq \Lambda \Subset \mathbb{Z}^d$ (see eq. (1.87)).
$S(A\|B)_\rho$	CE in A given B (see section 1.3.1).
$\widetilde{S}_\alpha^\uparrow(A\|B)_\rho$	SR-CE in A given B (see eq. (1.55))
$I(A:B)_\rho$	Mutual information (MI) of A and B (see section 1.3.1).
$I(A:B\|C)_\rho$	CMI of A and B given C (see eq. (1.9)).
$\widehat{I}(A;B\|C)_\rho$	BS-CMI of A against B given C (see eq. (1.72)).
$\mathbf{I}(A:C\|D)_\sigma$	MCMI of A and C given D (see eq. (3.21))
G. Asymptotic notation (Landau symbols)	
O	Landau symbol for upper bound (see eq. (1.49)).
Ω	Landau symbol for lower bound (see eq. (1.50)).
Θ	Landau symbol for exact scaling (see eq. (1.51)).

Chapter 1

Introduction

People have the mistaken impression that mathematics is just equations. In fact, equations are just the boring part of mathematics.
— attributed to Stephen Hawking.

Functional inequalities constitute one of the cornerstones of modern analysis, partial differential equation (PDE) theory, probability theory, statistical physics, and, more recently, information and quantum information theory. They are nearly ubiquitous in the study of dynamical systems and in the form of Hölder's and Young's inequalities, well known to every undergraduate studying real and functional analysis. In PDE theory, the Sobolev, Gagliardo-Nirenberg, and Nash inequalities provide crucial insights into the existence and regularity of solutions, enabling conclusions about global existence or blow-up behaviour. Markov's, Chebyshev's, and Jensen's inequalities are staples in the statistician's toolbox, while Poincaré and log-Sobolev inequalities, together with hypercontractivity, bring joy to probabilists and physicists alike. The latter have also found their way into information and quantum information theory, where they are joined by inequalities of entropy functionals and their generalisations to non-commutative operator spaces. Among the most prominent results in this realm are Lieb's concavity theorem—better known in its equivalent form as the strong subadditivity (SSA) of entropy or the Lieb-Ruskai inequality—as well as the Golden-Thompson, Araki-Lieb-Thirring, and, last but not least, Fannes' inequality, which provides a uniform continuity bound for the von Neumann entropy and serves as a comforting reminder that entropy, at the very least, does not change too abruptly.
While the present thesis does not claim the depth or genius of these celebrated inequalities, it nonetheless explores some of the aforementioned fields while sharing functional inequalities as the main object of interest. It consists of six projects at the intersection of quantum information theory and the quantum mechanics of many-body systems, with each project leaning more towards one domain or the other. Beyond some subtler connections, which will be explored in the remainder of this introduction and in the project discussions, a unifying theme across all of them is the role of one or more functional inequalities. These are either deployed directly to analyse the static and dynamic properties of many-body quantum systems and information-theoretic tasks, as in [GMR24; Cap+24; Alh+24] and [GR24b], or expected to refine existing relations with potential future applications, as in [Blu+24; Aud+25]. The inequalities mentioned, along with their immediate consequences, will be our primary focus, as they carry the majority of insights. At times, this comes at the cost of omitting detailed discussions of proofs, generalisations, and examples, which can be found in the respective papers underlying this thesis.

1.1 Overview

The thesis begins with a chapter on basic notation (section 1.2), in which we largely follow the nomenclature of the referenced papers, occasionally deviating to unify concepts and to clarify distinctions between certain objects. For example, we differentiate between finite-dimensional Hilbert spaces, denoted by $\mathcal{H}$ or $\mathcal{K}$, and the infinite-dimensional space $\mathcal{F} \cong L^2(\mathbb{R})$.
This is followed by the preliminaries, presented in sections 1.3 and 1.4, which are divided into two sections. The first, section 1.3, covers foundational material relevant to multiple projects, while the second, section 1.4, focuses on material specific to each individual project.

The reason for this subdivision is to allow the thesis to be read either in its entirety or selectively. Readers interested in only a subset of the projects can, after reviewing the basic notation (section 1.2) and general preliminaries (section 1.3), proceed directly to the relevant sections. Taken together, each project's preliminary, objective, and result sections form a self-contained unit, with their association clearly indicated in the section titles.
The general preliminaries (section 1.3) begin with a discussion of entropy functionals, which play a central role in all projects except [GMR24]. They are introduced in section 1.3.1, followed by a structure result of von Neumann subalgebras and their conditional expectations (section 1.3.2). These concepts are particularly relevant for [Cap+24] and [GR24b]. As we do not consider their infinite-dimensional counterparts, our discussion remains restricted to their definitions and properties in the setting of complex matrices. In section 1.3.3, we begin by introducing QMSs in finite dimensions, along with their one-to-one correspondence with generators in GKSL-form. Subsequently, we revisit inner products, such as KMS and GNS, present further the gap and MLSI and CMLSI and then demonstrate their use in studying mixing time; this lays the foundation for [Cap+24]. The discussion is then extended to generation theory on general Banach spaces, which provides an essential foundation for [GMR24]. The final general section explores Hamiltonians on finite lattice systems, introduces the concept of interactions, and their classification, which play a role in both [Cap+24] and [Alh+24].
For the project-specific preliminaries (section 1.4), we begin with an introduction to continuity bounds (section 1.4.1) including previous results relevant to [Aud+25; Blu+24]. This is followed by section 1.4.2, where we discuss connections between CMI, recoveries and the reconstruction of Gibbs states, which provides context for the results in [Alh+24]. We then examine results on the mixing times of Davies semigroups (section 1.4.3), the subject of [Cap+24] focusing mostly on gap, MLSI and CMLSI and theoretical framework around them. Staying in the realm of semigroups, we then introduce open Bosonic systems and previous work on Gaussian semigroups and the so-called cat code generators. We then define what we call quantum Sobolev spaces and the concept of a Sobolev preserving semigroup, which is the framework used in [GMR24]. The preliminaries are concluded with a discussion of hypothesis testing, Stein's lemma, and the generalisation to von Neumann subalgebra resources studied in [GR24b] (section 1.4.5).
Following these preliminaries, the project-specific objectives are presented in chapter 2, centred around the aforementioned key inequalities. These are then complemented by the corresponding results and discussion sections in chapter 3, which either prove these inequalities or apply them in various contexts.
We start again with the [Aud+25; Blu+24] bundled into section 2.1. For [Aud07] this central inequality is an upper bound on the distance of von Neumann entropy involving the Jordan-Hahn (JH) decomposition of the input states and a correction by the binary entropy in their trace distance. For [Blu+24] the central inequalities are continuity bounds for the SR divergence distance to a convex compact set with at least one positive-definite element, using two strategies from the literature that we build upon and generalise. In the results and discussion section of this project (section 3.1), we first present the novel bound for the von Neumann entropy and then derived from it the well-known Audenaert bound, a bound on the CE for states with equal marginals, as well as continuity bounds for the relative entropy, both for fixed and variable second arguments. We then go on to discuss the two approaches used to prove the SR divergence distance bounds whilst also presenting a mixture of both, behaving favourably in all regimes of $\alpha \in (1, \infty)$.
In section 2.2 the main inequalities are an upper bound on the trace error of a single reconstruction step by BS-CMI and a complementary upper bound on the latter quantity that quantifies the BS-recovery condition. The proof of both results is presented in section 3.2, together with proves of additional technical details leading to an efficient MPO reconstruction of Gibbs states for translation-invariant, local interactions on a spin chain. Derived from this result, we further describe an efficient algorithm for obtaining such a MPO from measurement data.
For section 2.3 the central inequality is an approximate tensorisation of the relative entropy with an additive correction, we call MCMI. Combining this result with the assumption of a uniform bound on the local gap, a suitable covering of the lattice and a sufficient decay of this, MCMI one can derive an improved mixing time estimate for the Davies semigroup for commuting local interactions. We further show that from the mixing time one can derive a discrete large-time contraction coefficient for the relative entropy and a global spectral gap of the semigroup, with all results presented in section 3.3.
The objective of section 2.4 is to demonstrate that, under a sufficient condition given by a sequence of

inequalities relating the formal generator of the open bosonic system to powers of the bosonic number operator, the formal generator serves as the core of the generator of a QMS that additionally is Sobolev-preserving. In section 3.4, the strategy used to prove the result is presented and further complemented by strengthened properties of the semigroup, which arise under stronger relations between the formal generator and the bosonic number operator. After establishing the sufficient condition for Gaussian semigroups and certain cat code gates, the section concludes with a discussion of the advantages of the QSS framework, particularly in the context of perturbation theory, where its usefulness is illustrated through concrete examples.
Lastly, we present the central inequality of [GR24b] in section 2.5 which is an operator inequality for a specific Stinespring decomposition of a conditional expectation, that we proof exists in section 3.5. Given this decomposition and its properties, we then can prove the generalised quantum Stein's lemma for subalgebra resource theories by reducing it to the previously established result in [HT16].
The thesis ends with appendix chapters containing the accepted versions and preprints of the submitted manuscripts, which this thesis is based on.

1.2 Basic notation

Although standard notation is used in most places, some conventions involve subtleties; therefore, a concise summary of the basic notation employed is provided here.
We denote finite-dimensional Hilbert spaces by $\mathcal{H}$ or $\mathcal{K}$, with their dimension given by d, or more explicitly as $d_{\mathcal{H}}$ respectively $d_{\mathcal{K}}$ when clarity is required. On any Hilbert space considered (finite or infinite-dimensional), we denote the inner product by $\langle\cdot,\cdot\rangle$ and the induced norm by $\|\cdot\|$. Subscripts, such as $\|\cdot\|_{\mathcal{H}}$ for a space $\mathcal{H}$ or $\|\cdot\|_{\mathcal{F}}$ for the space $\mathcal{F}$ (introduced below), will be used when necessary for clarity or emphasis. A general Banach space, of which Hilbert spaces are a special case, is denoted by $(\mathcal{X},\|\cdot\|_{\mathcal{X}})$.
The only infinite-dimensional Hilbert space considered in this thesis is $\mathcal{F}\cong L^2(\mathbb{R})$, equivalently defined as the closure of the span of an orthonormal basis $\{|n\rangle : n\in\mathbb{N}_0\}$:

$$\mathcal{F}\equiv\overline{\operatorname{span}\{|n\rangle : n\in\mathbb{N}_0\}}\,, \tag{1.1}$$

We primarily work with the latter representation, where $|n\rangle$ denotes the Fock basis. Intimately tied to this basis are the annihilation operator a ($a\,|n+1\rangle=\sqrt{n+1}\,|n\rangle$), creation operator a^* ($a^*\,|n\rangle=\sqrt{n+1}\,|n+1\rangle$), and, consequently, the number operator $\mathbf{N}=a^*a$. These operators are examples of unbounded linear operators acting on respective domains $\mathcal{D}(a)$, $\mathcal{D}(a^*)$, and $\mathcal{D}(\mathbf{N})$, which are dense subspaces of $\mathcal{F}$.
More generally, for a linear operator $\mathcal{O}$ on a Banach space $(\mathcal{X},\|\cdot\|_{\mathcal{X}})$, we denote its domain by $\mathcal{D}(\mathcal{O})$ and the operator-domain pair as $(\mathcal{O},\mathcal{D}(\mathcal{O}))$. Furthermore, we call $\{\mathcal{O}(X) : X\in\mathcal{D}(\mathcal{O})\}$ its range. If the domain $\mathcal{D}(\mathcal{O})$ is dense in $\mathcal{X}$ (that is, $\overline{\mathcal{D}(\mathcal{O})}=\mathcal{X}$), the operator $\mathcal{O}$ is called densely defined. A densely defined operator $(\mathcal{O},\mathcal{D}(\mathcal{O}))$ is called closed if its domain together with the associated graph norm $(\mathcal{D}(\mathcal{O}),\|\cdot\|_{\mathcal{X}}+\|\mathcal{O}(\cdot)\|_{\mathcal{X}})$ is a Banach space. For a densely defined closed $(\mathcal{O},\mathcal{D}(\mathcal{O}))$ we call $(\mathcal{O}',\mathcal{D}(\mathcal{O}'))$ a core, if $\mathcal{D}(\mathcal{O}')\subseteq\mathcal{D}(\mathcal{O})$, both operators agree on $\mathcal{D}(\mathcal{O}')$ and $\mathcal{D}(\mathcal{O}')$ is dense in $(\mathcal{D}(\mathcal{O}),\|\cdot\|_{\mathcal{X}}+\|\mathcal{O}(\cdot)\|_{\mathcal{X}})$. Conversely, if there exists a closed operator for which the densely defined operator $(\mathcal{O},\mathcal{D}(\mathcal{O}))$ serves as a core, we say that $(\mathcal{O},\mathcal{D}(\mathcal{O}))$ is closable and denote the closure by $(\overline{\mathcal{O}},\mathcal{D}(\overline{\mathcal{O}}))$.
The space of bounded linear operators on a Banach space $(\mathcal{X},\|\cdot\|_{\mathcal{X}})$ is denoted by

$$\mathcal{B}(\mathcal{X})\equiv\{\mathcal{O}:\mathcal{X}\to\mathcal{X} : \|\mathcal{O}\|_{\mathcal{B}(\mathcal{X})}<\infty\}\,,$$

with the corresponding norm defined as,

$$\|\mathcal{O}\|_{\mathcal{B}(\mathcal{X})}\equiv\sup_{X\in\mathcal{X}\setminus 0}\frac{\|\mathcal{O}(X)\|_{\mathcal{X}}}{\|X\|_{\mathcal{X}}}\,. \tag{1.2}$$

Notably, $(\mathcal{B}(\mathcal{X}),\|\cdot\|_{\mathcal{B}(\mathcal{X})})$ forms a Banach space in its own right. In particular, the spaces of bounded linear operators $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{F})$ on the Hilbert spaces $(\mathcal{H},\|\cdot\|_{\mathcal{H}})$ and $(\mathcal{F},\|\cdot\|_{\mathcal{F}})$, respectively, are defined in the same way. In these cases, however, we abbreviate $\|\cdot\|_{\mathcal{B}(\mathcal{H})}$ and $\|\cdot\|_{\mathcal{B}(\mathcal{F})}$ by $\|\cdot\|_\infty$, referring to this simply as the operator norm. For $\mathcal{O}\in\mathcal{B}(\mathcal{X})$ we use $\ker\mathcal{O}$ to denote the kernel of $\mathcal{O}$ and $\operatorname{Eig}\mathcal{O}$ for the set of eigenvalues.

For any operator $X \in \mathcal{B}(\mathcal{H})$ or $\mathcal{B}(\mathcal{F})$, its adjoint $X^* \in \mathcal{B}(\mathcal{H})$ respectively $X^* \in \mathcal{B}(\mathcal{F})$ is uniquely defined by $\langle X^*\phi, \psi\rangle = \langle \phi, X\psi\rangle$ for all $\phi, \psi \in \mathcal{H}$ or $\mathcal{F}$. A bounded operator is self-adjoint if $X = X^*$.
More generally, for an (unbounded) densely defined $(X, \mathcal{D}(X))$ on $\mathcal{F}$, its adjoint $(X^*, \mathcal{D}(X^*))$ is defined as follows: The domain $\mathcal{D}(X^*)$ consists of all $\eta \in \mathcal{F}$ for which the linear functional $\psi \mapsto \langle \eta, X\psi\rangle$ (defined for $\psi \in \mathcal{D}(X)$) is bounded. Riesz representation theorem then guarantees the existence of a unique vector, denoted $X^*\eta$, satisfying $\langle X^*\eta, \psi\rangle = \langle \eta, X\psi\rangle$ for all $\psi \in \mathcal{D}(X)$. Analogous to the bounded setting, we call a densely defined operator $(X, \mathcal{D}(X))$ self-adjoint if $X = X^*$. This, however, requires both that the domains coincide, $\mathcal{D}(X) = \mathcal{D}(X^*)$, and that $X\psi = X^*\psi$ for all ψ in this common domain. The operators $a, a^*, \mathbf{N}$ (all with their respective domain) mentioned above are closed, but only $\mathbf{N}$ is self-adjoint, while a and a^* are adjoints of each other.
Quantum states on $\mathcal{H}$ or $\mathcal{F}$, denoted by $\mathcal{S}(\mathcal{H})$ or $\mathcal{S}(\mathcal{F})$ respectively, are defined as operators possessing the properties of being self-adjoint, having a non-negative spectrum, and eigenvalues summing to one. For states on $\mathcal{F}$ (and trivially for ones on $\mathcal{H}$), these are necessarily bounded operators and belong to the set of trace-class operators $\mathcal{T}(\mathcal{F}) \subset \mathcal{B}(\mathcal{F})$. The trace of an operator is represented by $\mathrm{Tr}[\,\cdot\,]$ in both $\mathcal{B}(\mathcal{H})$ and $\mathcal{T}(\mathcal{F})$ and $\mathcal{T}(\mathcal{F})$ is defined as

$$\mathcal{T}(\mathcal{F}) = \{X \in \mathcal{B}(\mathcal{F}) : \mathrm{Tr}[|X|] < \infty\} \tag{1.3}$$

with $|X| := (X^*X)^{1/2}$ the operator absolute value. We call $\|\cdot\|_1 = \mathrm{Tr}[|X|]$ the trace norm and note that $\mathcal{T}(\mathcal{F})$ together with this norm forms a Banach space. All definitions also apply on $\mathcal{B}(\mathcal{H})$, and since $\mathcal{H}$ is finite-dimensional, we have $\mathcal{T}(\mathcal{H}) = \mathcal{B}(\mathcal{H})$. More generally we can define for $X \in \mathcal{B}(\mathcal{H})$ or $X \in \mathcal{B}(\mathcal{F})$, the Schatten-$p$ norms as

$$\|X\|_p \equiv \mathrm{Tr}[|X|^p]^{1/p}\,, \quad p \in [1, \infty]\,, \tag{1.4}$$

where for $X \in \mathcal{B}(\mathcal{F})$, these norms can be infinite if $p \in [1, \infty)$. A fundamental property connecting them is Hölder's inequality. For any $X, Y \in \mathcal{B}(\mathcal{H})$ and any pair of conjugate exponents $p, q \in [1, \infty]$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, it states that

$$|\,\mathrm{Tr}[X^*Y]| \le \|X\|_p \|Y\|_q\,. \tag{1.5}$$

This inequality also holds in the infinite-dimensional setting, provided the norms on the right-hand side are finite, and the trace is well-defined. Note that for $p = \infty$, indeed $\|\cdot\|_\infty$ from eq. (1.2) with $\mathcal{X} = \mathcal{H}$ or $\mathcal{F}$ and eq. (1.5) coincide. To conclude this discussion of norms, we highlight the special case $p = 2$, as $\|\cdot\|_2$ corresponds to the norm induced by the HS-inner product,

$$\langle X, Y\rangle_{\mathrm{HS}} := \mathrm{Tr}[X^*Y]\,, \tag{1.6}$$

for $X, Y \in \mathcal{B}(\mathcal{H})$ which turns $\mathcal{B}(\mathcal{H})$ into a Hilbert space itself. Note that for $p = q = 2$, Hölder's inequality (1.5) reduces precisely to the Cauchy-Schwarz inequality: $|\langle X, Y\rangle_{\mathrm{HS}}| \le \|X\|_2 \|Y\|_2$.
Multipartite systems are denoted as $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \cdots \equiv \mathcal{H}_{ABC\ldots}$, with the associativity of the tensor product allowing us to restrict attention to the bipartite case. In this setting, the partial trace tr_A is a map $\mathrm{tr}_A : \mathcal{B}(\mathcal{H}_{AB}) \to \mathcal{B}(\mathcal{H}_B)$, which may be embedded into $\mathcal{B}(\mathcal{H}_{AB})$ via the extension $\mathbb{1}_A \otimes \mathrm{tr}_A$, or in normalised form as $\iota_A \otimes \mathrm{tr}_A$, where we use $\iota_A \equiv \mathbb{1}_A / d_A$ to denote the maximally mixed state. To avoid redundancy, we often omit identity operators, and when an operator acts non-trivially only on a specific subsystem, we indicate this with a subscript—for instance, $X_A \equiv X_A \otimes \mathbb{1}_B \in \mathcal{B}(\mathcal{H}_{AB})$.
For self-adjoint operators $X, Y \in \mathcal{B}(\mathcal{H})$, we use the Loewner-order:

$$X \ge Y \quad \text{if and only if} \quad \langle \psi, (X - Y)\psi\rangle \ge 0 \quad \forall \psi \in \mathcal{H} \setminus \{0\}\,,$$

with $X > Y$ defined analogous. We call a linear map $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ hermiticity-preserving if $\Psi(X^*) = \Psi(X)^*$ for all $X \in \mathcal{B}(\mathcal{H})$ and further positive if $X \ge 0$ implies $\Psi(X) \ge 0$. A map is completely positive (CP) if its extension

$$\mathrm{id}_n \otimes \Psi : \mathcal{B}(\mathbb{C}^n \otimes \mathcal{H}) \to \mathcal{B}(\mathbb{C}^n \otimes \mathcal{K})$$

is positive for all $n \in \mathbb{N}_0$. A completely positive and trace preserving (CPTP) map or quantum channel is a linear CP map $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ that is also trace-preserving. For maps on $\mathcal{B}(\mathcal{H})$, this means $\mathrm{Tr}[\Psi(X)] = \mathrm{Tr}[X]$ for all $X \in \mathcal{B}(\mathcal{H})$. A related concept are unital linear maps $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, defined

by the property $\Psi(\mathbb{1}_{\mathcal{H}}) = \mathbb{1}_{\mathcal{K}}$. The HS-adjoint of a linear map Ψ, denoted $\Psi^\dagger$, is defined with respect to the HS-inner product (see eq. (1.6)) and connects these properties among others. In particular, a linear map $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is trace-preserving if and only if its HS-adjoint $\Psi^\dagger$ is unital.
Note that the notion of operator order is the same on $\mathcal{B}(\mathcal{F})$ while notions of positive, unital, CP and CPTP maps analogously carry over to bounded maps $\Psi : \mathcal{T}(\mathcal{F}) \to \mathcal{T}(\mathcal{F})$, that is bounded w.r.t. to the norm

$$\|\Psi\|_{1\to 1} \equiv \sup_{X\in\mathcal{T}(\mathcal{F})\setminus\{0\}} \frac{\|\Psi(X)\|_1}{\|X\|_1} . \tag{1.7}$$

1.3 General preliminaries

1.3.1 Entropy functionals

A cornerstone of quantum information theory are entropy functionals, which characterise optimal convergence rates in hypothesis testing (quantum Stein's Lemma [HP91; NO00]), purity [NC12], separability [HHH96; Per96], and, more generally, resource content [GS08; BG15], as well as the compression rate of memoryless channels [Sch95]. These functionals also arise naturally as bounds on optimal rates in quantum key distribution [KGR05; DW05], as well as in numerous other contexts. The number of such functionals is substantial and continues to grow (see, e.g., [HT24]), and it would be unfeasible to provide a comprehensive overview of all of them here. Instead, we shall present the widely used instances, along with those particularly relevant to the works addressed in this thesis.
Arguably the most prominent example is the von Neumann entropy, introduced by John von Neumann as early as 1932 [Neu96], predating yet generalising the well-known Shannon entropy [Sha48] from classical information theory. For a quantum state $\rho \in \mathcal{S}(\mathcal{H})$, it is defined as

$$S(\rho) \equiv -\operatorname{Tr}[\rho\log(\rho)] , \tag{1.8}$$

where the convention $0\log 0 = 0$ is used. It further serves as the building block for various other quantities such as the CE: $S(A|B)_\rho \equiv S(\rho_{AB}) - S(\rho_B)$ for $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$, the MI: $I(A:B)_\rho \equiv S(\rho_A) + S(\rho_B) - S(\rho_{AB})$, and, last but not least, the CMI, given for $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_{ABC})$ by

$$I(A:B|C)_\rho \equiv S(\rho_{AC}) + S(\rho_{BC}) - S(\rho_C) - S(\rho_{ABC}) . \tag{1.9}$$

These latter quantities, although rooted in quantum information theory, have also attracted the attention of mathematical physicists due to their usefulness in the study of Gibbs states on lattice systems. They enable the quantification of (conditional) independence and, if they exhibit decay with respect to lattice distance, allow the decomposition of a large system—or functionals of the global Gibbs state—into smaller, manageable patches. Since these patches effectively reduce the problem size (allow improving the scaling with system size), certain algorithms, estimates, or reconstructions can be applied efficiently at the local level, and then lifted to global conclusions about the system or its associated functionals [BK18; Sca+21; Ono+23; FFS24; Bar+24]. We employ strategies along these lines in both [Cap+24] and [Alh+24], albeit using different measures of independence, which we will introduce in sections 1.4.2 and 3.3.1 respectively.
To delve deeper into the structure of the CE, MI and CMI and connect to the independence measures employed in [Cap+24; Alh+24], we introduce the Umegaki relative entropy (or simply relative entropy) [Ume62]. This quantity not only allows for a reformulation of all three, i.e., CE, MI and CMI (eq. (1.11) and above), but also serves as a basis for related functionals, such as the one used in [Alh+24] which arises when replacing the relative entropy with a different divergence. The relative entropy for $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ is defined as

$$D(\rho\|\sigma) \equiv \begin{cases} \operatorname{Tr}[\rho\log\rho - \rho\log\sigma] & \text{if } \ker\sigma \subseteq \ker\rho\,, \\ +\infty & \text{otherwise}\,, \end{cases} \tag{1.10}$$

using again the convention that $0\log 0 = 0$ for ρ with $\ker\rho \neq \emptyset$. It generalises the well-known Kullback-Leibler (KL) divergence from classical statistics [KL51], reducing to it when $[\rho,\sigma] = 0$, as the eigenvalues can then be interpreted as classical probability distributions.
Alongside the von Neumann entropy, it is one of the most central functionals in quantum information theory,

underpinning key properties of CE, MI, and CMI through their reformulation in relative entropy. It appears in quantum hypothesis testing (to be explored further in section 1.4.5), resource theories, and satisfies the chain rule for conditional expectations (introduced in section 1.3.2). The aforementioned rewritings of CE, MI and CMI are given by $S(A|B)_\rho = -D(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B)$, $I(A:B)_\rho = D(\rho_{AB}\|\rho_A \otimes \rho_B)$ and finally

$$\begin{aligned} I(A:B|C)_\rho &= D(\rho_{ABC} \| \mathbb{1}_A \otimes \rho_{BC}) - D(\rho_{AC} \| \mathbb{1}_A \otimes \rho_C) \\ &= D(\rho_{ABC}\|\rho_A \otimes \rho_{BC}) - D(\rho_{AC}\|\rho_A \otimes \rho_C)\,. \end{aligned} \tag{1.11}$$

An alternative generalisation of the KL divergence to the quantum setting is given by the BS entropy, which plays a central role in the measure of independence used in [Alh+24]. For quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, it is defined as:

$$\widehat{D}(\rho\|\sigma) \equiv \begin{cases} \mathrm{Tr}\big[\rho \log\big(\rho^{1/2}\sigma^{-1}\rho^{1/2}\big)\big] & \text{if } \ker\sigma \subseteq \ker\rho\,, \\ +\infty & \text{otherwise}\,, \end{cases} \tag{1.12}$$

where σ^{-1} denotes the Moore-Penrose pseudo inverse. Its foundational use to define a measure of independence and the connection to recovery maps will be detailed in more depth in section 1.4.2.

We now turn to other entropy functionals relevant to both [GR24b] and [Blu+24], and examine common structural features among them. Just as the BS and Umegaki relative entropy are generalisations of the KL divergence, the SR divergence [ML+13] generalises the classical Rényi divergences introduced by Alfred Rényi in 1961 [Rén61]. For $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and $\alpha \in [1/2, 1) \cup (1, \infty]$ they are given as

$$\widetilde{D}_\alpha(\rho\|\sigma) \equiv \begin{cases} \frac{1}{\alpha-1} \log \widetilde{Q}_\alpha(\rho,\sigma) & \text{if } \ker\sigma \subseteq \ker\rho\,, \\ +\infty & \text{otherwise}\,, \end{cases} \quad \text{with} \quad \widetilde{Q}_\alpha(\rho,\sigma) \equiv \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]. \tag{1.13}$$

Notable special cases include $\alpha = 1/2$, where the quantity is the logarithm of the Uhlmann fidelity [Uhl76]; the limit $\alpha \to 1$, which recovers the relative entropy; and $\alpha = \infty$, yielding the so-called max Rényi divergence [Dat09], with the closed and variational forms

$$D_\infty(\rho\|\sigma) \equiv \begin{cases} \log \left\|\sigma^{-1/2}\rho\sigma^{-1/2}\right\|_\infty & \text{if } \ker\sigma \subseteq \ker\rho\,, \\ +\infty & \text{otherwise} \end{cases} = \log\inf\{\lambda : \rho \leq \lambda\sigma\}\,.$$

Due to the non-commutative nature of $\mathcal{S}(\mathcal{H})$, this family is only one among many quantum extensions of Rényi's classical equivalent, including, for instance, the geometric Rényi (GR) and Petz-Rényi (PR) divergences. What unifies all of these, and what are widely accepted as the defining properties of a divergence map $\mathbb{D} : \bigcup_{\mathcal{H}} \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \to \mathbb{R} \cup \{\infty\}$ [Gou24], are:

1. Normalisation: $\mathbb{D}(1\|1) = 0$,

2. Invariance under Hilbert space isomorphisms, and most notably

3. the data processing inequality (DPI), which states that for all $\mathcal{H}, \mathcal{K}$, $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, and CPTP maps $\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{K})$,
$$\mathbb{D}(\Phi(\rho)\|\Phi(\sigma)) \leq \mathbb{D}(\rho\|\sigma)\,. \tag{1.14}$$

From these (and especially from the DPI), one immediately obtains non-negativity $\mathbb{D} \geq 0$, and that $\mathbb{D}(\rho\|\rho) = 0$ [Gou24]. For all the divergences introduced so far, the stronger condition of faithfulness holds, i.e., $\mathbb{D}(\rho\|\sigma) = 0$ implies $\rho = \sigma$, while they are also additive under tensor products,

$$\mathbb{D}(\rho \otimes \rho'\|\sigma \otimes \sigma') = \mathbb{D}(\rho\|\sigma) + \mathbb{D}(\rho'\|\sigma')\,,$$

and can be extended to general non-negative operators. This allows one to formulate the additional property of anti-monotonicity in the second argument, which is satisfied by all divergences introduced explicitly so far: for $X, Y, Z \in \mathcal{B}(\mathcal{H})$, with $X \geq 0$ and $Y \geq Z \geq 0$, one has

$$\mathbb{D}(X\|Y) \leq \mathbb{D}(X\|Z)\,. \tag{1.15}$$

Throughout this thesis, whenever we use $\mathbb{D}$, we will explicitly specify the subset of divergences it refers to. With these properties at hand, we can now explore another perspective on divergences: their role as distance measures. Here, 'measure' is used colloquially rather than in the mathematical sense. For a divergence $\mathbb{D}$ (specifically, those introduced explicitly), such quantities express how far a quantum state $\rho \in \mathcal{S}(\mathcal{H})$ is from another state σ, or more generally, from a convex, compact set $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ containing at least one positive-definite state, via

$$\mathbb{D}(\rho \| \mathcal{C}) \equiv \inf_{\sigma \in \mathcal{C}} \mathbb{D}(\rho \| \sigma) \,. \tag{1.16}$$

The convexity of $\mathcal{C}$ ensures several useful properties while also being physically motivated by classical probabilistic combination of ensembles (quantum states). This, together with the requirement that $\mathcal{C}$ contains a full-rank state, arise naturally in resource theories [Gou24] and in hypothesis testing scenarios [BP10a], where such divergence-based distances are employed. As we will later see, these constraints on $\mathcal{C}$ also allowed us to prove uniform continuity bounds for certain divergences in [Blu+24].
It is noteworthy that the CE and MI can be written in the form of eq. (1.16) as

$$\begin{aligned} S(A|B)_\rho &= - \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D(\rho_{AB} \| \iota_A \otimes \sigma_B) + \log(d_A) \,, \\ I(A:B)_\rho &= \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D(\rho_{AB} \| \rho_A \otimes \sigma_B) \end{aligned} \tag{1.17}$$

i.e., the minimizer in the optimization is $\sigma_B = \rho_B$. This property—proven through non-negativity of the relative entropy and the chain rule (see eq. (1.20))—fails for other divergences [Blu+23a; Tom16] and underpins the special role of relative entropy among all other divergences.

1.3.2 Von Neumann subalgebras of matrices and their conditional expectations

The previous section concluded with the variational characterisation of the CE (eq. (1.17)), interpreting it as a distance measure to the convex set $\iota_A \otimes \mathcal{S}(\mathcal{H}_B)$. This set is precisely the set of states within the algebra $\mathbb{1}_A \otimes \mathcal{B}(\mathcal{H}_B)$. One can readily verify that this algebra is the commutant of $\mathcal{B}(\mathcal{H}_A) \otimes \mathbb{1}_B$, where the commutant of a set $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ is defined as

$$\mathcal{A}' \equiv \{X \in \mathcal{B}(\mathcal{H}) : [X, Y] = 0 \; \forall Y \in \mathcal{A}\} \,, \tag{1.18}$$

with $[X, Y] \equiv XY - YX$ denoting the commutator. The set $\mathbb{1}_A \otimes \mathcal{B}(\mathcal{H}_B)$ is also closed under the adjoint operation, forms an algebra, and contains the identity operator. These properties follow from the fact that it is the commutant of a set which is itself closed under the adjoint operation. Motivated by this observation one establishes the following definition[1]: A subset $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ is called a von Neumann subalgebra of $\mathcal{B}(\mathcal{H})$ if it is the commutant of a set $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ that is closed under the adjoint operation (that is, if $Y \in \mathcal{A}$, then $Y^* \in \mathcal{A}$).
It is a known that every finite-dimensional von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ can be decomposed as

$$\mathcal{N} = U \bigoplus_{k=1}^{K} \mathbb{1}_{m_k} \otimes \mathcal{B}(\mathbb{C}^{d_k}) U^*$$

with $m_k, d_k \in \mathbb{N}$ and $U : \bigoplus_{k=1}^{K} \mathbb{C}^{m_k} \otimes \mathbb{C}^{d_k} \to \mathcal{H}$ an isomorphism [Wol12]. For simplicity and readability, and since the subsequent results are unaffected, we will henceforth assume $\mathcal{H} = \bigoplus_{k=1}^{K} \mathbb{C}^{m_k} \otimes \mathbb{C}^{d_k}$, effectively setting the isomorphism U to the identity operator. As shown in [Wol12], there exist positive linear maps that project onto $\mathcal{N}$. It is also shown that any projection onto $\mathcal{N}$ that is additionally a positive map has the following structure:

$$E(X) = \bigoplus_{k=1}^{K} \mathbb{1}_{m_k} \otimes \mathrm{tr}_{\mathbb{C}^{m_k}} [P_k X P_k \pi_k \otimes \mathbb{1}_{\mathbb{C}^{d_k}}] \tag{1.19}$$

[1] Note that, throughout this thesis, we restrict attention to von Neumann subalgebras of the finite-dimensional $\mathcal{B}(\mathcal{H})$, making this definition sufficient and more consistent with the remainder of the discussion.

where P_k is the orthogonal projection onto the k-th subspace $\mathbb{C}^{m_k} \otimes \mathbb{C}^{d_k}$ within the direct sum $\mathcal{H} = \bigoplus_{k=1}^{K} \mathbb{C}^{m_k} \otimes \mathbb{C}^{d_k}$, and each $\pi_k \in \mathcal{S}(\mathbb{C}^{m_k})$ is a positive-semidefinite quantum state. This existence and structural result motivates the designation of a positive projection onto a von Neumann subalgebra as a 'conditional expectation'. It is relatively straightforward to deduce their basic properties, namely:

1. $E(X) = X$ for $X \in \mathcal{N}$,
2. $VE(X)W = E(VXW)$ for $V, W \in \mathcal{N}$,
3. E is an unital CP map.

Building on these properties, one can establish several notable interrelations between the map E, its HS-adjoint $E^\dagger$, and the state $\pi \equiv E^\dagger(\iota) = \frac{1}{d}\bigoplus_{k=1}^{K} \pi_k \otimes \mathbb{1}_{\mathbb{C}^{d_k}}$. These relationships will later underpin both the chain rule [OP93] and an entropy factorisation result for the relative entropy—results whose significance was already highlighted in the identification of eq. (1.17).
Although these results are known in the literature, they are often dispersed across different sources and have wider or more limited scope. For completeness and applicability in the projects presented in this thesis, we therefore summarise them here and provide detailed proofs. We begin with a lemma that captures a basic property of E, outlines properties of the state π, and presents a key characterisation of $E^\dagger$ in terms of π and E

Lemma 1.3.1 *Let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ be a conditional expectation (positive projection onto the von Neumann subalgebra $\mathcal{N}$), and let $\pi = E^\dagger(\iota) = \frac{1}{d}\bigoplus_{k=1}^{K} \pi_k \otimes \mathbb{1}_{\mathbb{C}^{d_k}}$. Then*

1. *E maps positive-definite operators to positive-definite operators.*
2. *$\langle E(X), Y\pi\rangle_{\mathrm{HS}} = \langle X, E(Y)\pi\rangle_{\mathrm{HS}}$ for all $X, Y \in \mathcal{B}(\mathcal{H})$. For $\pi > 0$ we will later call this GNS-symmetry (see eq. (1.30)).*
3. *$\pi \in \mathcal{N}'$, that is, π is in the commutant of $\mathcal{N}$.*
4. *For $\pi > 0$, E commutes with the modular operator of π, i.e., $[E, \Delta_\pi] = 0$, where $\Delta_\pi(\cdot) = \pi(\cdot)\pi^{-1}$.*
5. *$E^\dagger(\,\cdot\,\pi) = E(\cdot)\pi$ and for $\pi > 0$ $E^\dagger(\cdot) \equiv \pi^{1/2}E(\pi^{-1/2}\cdot\pi^{-1/2})\pi^{1/2}$.*

Proof. The first condition is a straightforward consequence of unitality. Assuming that $X \in \mathcal{B}(\mathcal{H})$ is positive-definite implies that there exists a $c > 0$ such that $X \geq c\,\mathbb{1}$. Hence, by the positivity of the conditional expectation, $E(X) \geq cE(\mathbb{1}) = c\,\mathbb{1} > 0$.
For the second claim, we use the second basic property of the conditional expectation and the fact that E preserves hermiticity to conclude:

$$\begin{aligned}\langle E(X), Y\pi\rangle_{\mathrm{HS}} &= \frac{1}{d}\operatorname{Tr}\big[E(X)^* Y E^\dagger(\mathbb{1})\big] = \frac{1}{d}\operatorname{Tr}[E(E(X)^* Y)]\\ &= \frac{1}{d}\operatorname{Tr}[E(X^* E(Y))] = \operatorname{Tr}[X^* E(Y)\pi] = \langle X, E(Y)\pi\rangle_{\mathrm{HS}}\ .\end{aligned}$$

To prove the third claim, let's assume $X \in \mathcal{B}(\mathcal{H})$ and $Y' \in \mathcal{N}$. Then, by the first and second basic properties of the conditional expectation and similar manipulations as above, together with the cyclicity of the trace:

$$\begin{aligned}\langle X, Y'\pi\rangle_{\mathrm{HS}} &= \operatorname{Tr}[X^* Y'\pi] = \frac{1}{d}\operatorname{Tr}[E(X)^* Y'] = \frac{1}{d}\operatorname{Tr}[Y' E(X)^*]\\ &= \frac{1}{d}\operatorname{Tr}[E(Y' X^*)] = \operatorname{Tr}[Y' X^* \pi] = \operatorname{Tr}[X^* \pi Y'] = \langle X, \pi Y'\rangle_{\mathrm{HS}}\ .\end{aligned}$$

As X was arbitrary, the Riesz representation theorem implies $Y'\pi = \pi Y'$. Since $Y' \in \mathcal{N}$ was also arbitrary, $\pi \in \mathcal{N}'$.
For the fourth claim note that by the second property of this lemma, the cyclicity of the trace, and the hermiticity-preservation of the conditional expectation, one has for arbitrary $X, Y \in \mathcal{B}(\mathcal{H})$:

$$\begin{aligned}\langle X, E(\Delta_\pi(Y))\pi\rangle_{\mathrm{HS}} &= \langle E(X), \Delta_\pi(Y)\pi\rangle_{\mathrm{HS}} = \langle Y^*, E(X^*)\pi\rangle_{\mathrm{HS}}\\ &= \langle E(Y)^*, X^*\pi\rangle_{\mathrm{HS}} = \langle X, \Delta_\pi(E(Y))\pi\rangle_{\mathrm{HS}}\ .\end{aligned}$$

As X was arbitrary, we get by the Riesz representation theorem $E(\Delta_\pi(Y))\pi = \Delta_\pi(E(Y))\pi$, which, due to the invertibility of π and the arbitrary choice of Y, immediately yields $[E, \Delta_\pi] = 0$. Let us briefly comment on the fact that this commutativity is solely implied by $\langle E(X), Y\pi\rangle_{\mathrm{HS}} = \langle X, E(Y)\pi\rangle_{\mathrm{HS}}$ for arbitrary $X, Y \in \mathcal{B}(\mathcal{H})$. This, in particular, means that if such a relation holds for another $\sigma \in \mathcal{S}(\mathcal{H})$ with $\sigma > 0$, one obtains $[E, \Delta_\sigma] = 0$ by the same argument.
To prove the last claim, we first note that the second claim, $\langle E(X), Y\pi\rangle_{\mathrm{HS}} = \langle X, E(Y)\pi\rangle_{\mathrm{HS}}$, can be rewritten using the definition of the HS-adjoint $E^\dagger$. Specifically, $\langle E(X), Y\pi\rangle_{\mathrm{HS}} = \left\langle X, E^\dagger(Y\pi)\right\rangle_{\mathrm{HS}}$ for arbitrary $X, Y \in \mathcal{B}(\mathcal{H})$. By the Riesz representation theorem, this yields $E(Y)\pi = E^\dagger(Y\pi)$ for all $Y \in \mathcal{B}(\mathcal{H})$, which can be written as $E(\cdot)\pi = E^\dagger(\,\cdot\,\pi)$. Now, assuming that π is invertible allows to rewrite this as $E^\dagger(\cdot) = E(\cdot\pi^{-1})\pi$, while commutativity with the modular operator, that is, the fourth fact in this lemma, finally gives

$$E^\dagger(\cdot) = E(\cdot\pi^{-1})\pi = E(\Delta_\pi(\pi^{-1/2}\cdot\pi^{-1/2}))\pi = \Delta_\pi(E(\pi^{-1/2}\cdot\pi^{-1/2}))\pi = \pi^{1/2}E(\pi^{-1/2}\cdot\pi^{-1/2})\pi^{1/2}\,.$$

□

Notably, many of these properties only hold when $\pi > 0$, that is, when it is positive-definite. Therefore, we want to provide equivalent conditions of when this is the case.

Lemma 1.3.2 *Let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ a conditional expectation (positive projection onto the von Neumann subalgebra $\mathcal{N}$) and $\pi = E^\dagger(\iota)$, then the following are equivalent:*

1. *$\pi > 0$.*
2. *$E^\dagger$ maps positive-definite elements to positive-definite elements.*
3. *There exist $\sigma \in \mathcal{S}(\mathcal{H})$, with $\sigma > 0$ such that $E^\dagger(\sigma) = \sigma$.*
4. *There exist $\sigma \in \mathcal{S}(\mathcal{H})$, with $\sigma > 0$ such that $\left\langle E(X), \sigma^{1/2}Y\sigma^{1/2}\right\rangle_{\mathrm{HS}} = \left\langle X, \sigma^{1/2}E(Y)\sigma^{1/2}\right\rangle_{\mathrm{HS}}$ for all $X, Y \in \mathcal{B}(\mathcal{H})$. We will later call this property KMS-symmetry (see eq. (1.29) and after).*

Proof. We begin with the fourth claim and then show all implications in reverse, ending with the implication of the fourth by the first.
4. implies 3.: For $X = \mathbb{1}$ and Y arbitrary one has due to the unitality of the conditional expectation

$$\langle\sigma, Y\rangle_{\mathrm{HS}} = \left\langle E(\mathbb{1}), \sigma^{1/2}Y\sigma^{1/2}\right\rangle_{\mathrm{HS}} = \left\langle \mathbb{1}, \sigma^{1/2}E(Y)\sigma^{1/2}\right\rangle_{\mathrm{HS}} = \langle\sigma, E(Y)\rangle_{\mathrm{HS}} = \left\langle E^\dagger(\sigma), Y\right\rangle_{\mathrm{HS}}\,.$$

Now as Y was arbitrary one has by Riesz representation theorem that $\sigma = E^\dagger(\sigma)$ with $\sigma > 0$ by assumption.
3. implies 2.: If $\sigma > 0$, for $X \in \mathcal{B}(\mathcal{H})$ with $X > 0$ there exists $c > 0$ such that $X \geq c\sigma$, hence by the positivity of $E^\dagger$:

$$E^\dagger(X) \geq cE^\dagger(\sigma) = c\sigma > 0\,.$$

2. implies 1.: As $E^\dagger$ maps positive-definite operators to positive-definite operators, and $\iota > 0$ one in particular has $\pi > 0$.
1. implies 4.: Let $\sigma = \pi > 0$, then through the second and last property of theorem 1.3.1 one has for arbitrary $X, Y \in \mathcal{B}(\mathcal{H})$:

$$\begin{aligned}\left\langle E(X), \pi^{1/2}Y\pi^{1/2}\right\rangle_{\mathrm{HS}} &= \left\langle E(X), \Delta_\pi^{-1/2}(Y)\pi\right\rangle_{\mathrm{HS}} = \left\langle X, E(\Delta_\pi^{-1/2}(Y))\pi\right\rangle_{\mathrm{HS}}\\ &= \left\langle X, \Delta_\pi^{-1/2}(E(Y))\pi\right\rangle_{\mathrm{HS}} = \left\langle X, \pi^{1/2}E(Y)\pi^{1/2}\right\rangle_{\mathrm{HS}}\,.\end{aligned}$$

Note that although the second property only gives $[E, \Delta_\pi] = 0$ this immediately yields $[E, p(\Delta_\pi)] = 0$ for $p : \mathbb{R} \to \mathbb{R}$ a polynomial and consequently $[E, f(\Delta_\pi)] = 0$ for $f : \mathbb{R} \to \mathbb{R}$ continuous by the Stone-Weierstraß approximation theorem. Hence, commutativity, specifically $[E, \Delta_\pi^{-1/2}] = 0$ which we used above, holds in particular for the function $x \mapsto x^{-1/2}$ under exclusion of an interval around zero fit to the spectrum of Δ_π. □

With these lemmas in place, we now turn to the two central inequalities: the chain rule [OP93] for the relative entropy and a corresponding entropy factorisation result. We emphasise once more that both results—as with the preceding lemmas—are known in the literature, though often presented under different names and in either more restrictive or more general frameworks. To adapt them to our specific setting, and in the interest of completeness and clarity regarding their scope of applicability, we restate and reprove them here.

Theorem 1.3.3 (Chain rule and entropy factorisation) *Let $E_{\mathcal{M}} : \mathcal{B}(\mathcal{H}) \to \mathcal{M}$ and $E_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ be conditional expectation (that is positive projections) onto $\mathcal{N}$ and $\mathcal{M}$ von Neumann subalgebras of $\mathcal{B}(\mathcal{H})$ respectively. Furthermore, let $\pi_{\mathcal{N}} = E_{\mathcal{N}}^\dagger(\iota) > 0$ (see theorem 1.3.2 for equivalent constraints). Then it holds for $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ that*

$$D(\rho \| E_{\mathcal{N}}^\dagger(\sigma)) = D(\rho \| E_{\mathcal{N}}^\dagger(\rho)) + D(E_{\mathcal{N}}^\dagger(\rho) \| E_{\mathcal{N}}^\dagger(\sigma))\,, \tag{1.20}$$

and as a consequence also

$$D(\rho \| E_{\mathcal{N}}^\dagger E_{\mathcal{M}}^\dagger(\sigma)) \leq D(\rho \| E_{\mathcal{N}}^\dagger(\rho)) + D(\rho \| E_{\mathcal{M}}^\dagger(\sigma))\,. \tag{1.21}$$

Proof. For the proof of the first claim we set $\pi = \pi_{\mathcal{N}}$ and $E = E_{\mathcal{N}}$ for better readability. We further momentarily assume $\rho, \sigma > 0$, that is both positive-definite. Using the final property established in theorem 1.3.1, we can express $E^\dagger(\rho) = \pi^{1/2} X \pi^{1/2}$ and $E^\dagger(\sigma) = \pi^{1/2} Y \pi^{1/2}$, where $X = E(\pi^{-1/2}\rho\pi^{-1/2})$ and $Y = E(\pi^{-1/2}\sigma\pi^{-1/2})$. Both X and Y are positive-definite, due to the assumption $\pi > 0$ and the first statement in theorem 1.3.1. Furthermore, due to $X, Y \in \mathcal{N}$, they commute with π (as noted in the second statement of the same lemma), and hence also with any polynomial in π. By the Stone-Weierstraß theorem, they consequently commute with $\pi^{1/2}$ as well. It therefore follows that

$$\begin{aligned}
D(\rho \| E^\dagger(\sigma)) &= \mathrm{Tr}[\rho \log \rho - \rho \log Y - \log \pi] \\
&= \mathrm{Tr}[\rho \log \rho - \rho \log X + \rho \log X - \rho \log Y - \log \pi] \\
&= \mathrm{Tr}\Big[\rho\big(\log \rho - \log(\pi^{1/2} X \pi^{1/2})\big)\Big] + \mathrm{Tr}[\rho(\log X - \log Y)] \\
&= \mathrm{Tr}\big[\rho\big(\log \rho - \log E^\dagger(\rho)\big)\big] + \mathrm{Tr}[\rho E(\log X - \log Y)] \\
&= \mathrm{Tr}\big[\rho\big(\log \rho - \log E^\dagger(\rho)\big)\big] + \mathrm{Tr}\Big[E^\dagger(\rho)\big(\log(\pi^{1/2} X \pi^{1/2}) - \log(\pi^{1/2} Y \pi^{1/2})\big)\Big]\,.
\end{aligned}$$

In the second to last line we use that $\log X, \log Y \in \mathcal{N}$ as $\mathcal{N}$ is a closed algebra, hence also closed under the operations that define the logarithm. Therefore, $\log X - \log Y = E(\log X - \log Y) \in \mathcal{N}$. We then concluded by taking the HS-adjoint in the last line.
For arbitrary $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and, $\varepsilon > 0$ we set $\rho_\varepsilon = \frac{1}{1+\varepsilon}(\rho + \varepsilon\pi)$ and σ_ε analogously. As both ρ_ε and σ_ε are positive-definite, one can apply the above and then cancel the common normalisation to obtain:

$$D(\rho + \varepsilon\pi \| E^\dagger(\sigma) + \varepsilon\pi) = D(\rho + \varepsilon\pi \| E_{\mathcal{N}}^\dagger(\rho) + \varepsilon\pi) + D(E^\dagger(\rho) + \varepsilon\pi \| E^\dagger(\sigma) + \varepsilon\pi)\,.$$

Since the above functionals are of the form $D(\rho' + \varepsilon X \,\|\, \sigma' + \varepsilon X)$, with $\rho', \sigma' \in \mathcal{S}(\mathcal{H})$ and $X \in \mathcal{B}(\mathcal{H})$, $X > 0$, therefore are upper semi-continuous in ε by [HM17, Proposition 3.8 (ii)], we may take the limit $\varepsilon \searrow 0$ of the equality to conclude the first claim.
The second claim is now just an application of the first using $\sigma = E_{\mathcal{M}}^\dagger(\sigma)$ together with the fact that $E_{\mathcal{N}}^\dagger$ is a CPTP map allowing for application of the DPI for the relative entropy [Tom16], i.e., $D(E_{\mathcal{N}}^\dagger(\rho) \| E_{\mathcal{N}}^\dagger(E_{\mathcal{M}}^\dagger(\sigma))) \leq D(\rho \| E_{\mathcal{M}}^\dagger(\sigma))$. □

Let us end this section, noting that the conditional expectations for a given von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ are all but unique. This non-uniqueness is illustrated by considering $\mathcal{N} = \mathbb{1}_A \otimes \mathcal{B}(\mathcal{H}_B)$ within $\mathcal{B}(\mathcal{H}_{AB}$. The map $E_1(X) = \iota_A \otimes \mathrm{tr}_A[X]$ is a conditional expectation onto $\mathcal{N}$. However, for any state $\rho_A \in \mathcal{S}(\mathcal{H}_A)$, the map $E_2(X) = \mathbb{1}_A \otimes \mathrm{tr}_A[(\rho_A \otimes \mathbb{1}_B)X]$ is also a conditional expectation onto the same algebra $\mathcal{N}$, but clearly $E_1 \neq E_2$ if $\rho_A \neq \iota_A$. This leads to the conclusion that the conditional expectation onto a given von Neumann algebra $\mathcal{N}$ is unique only up to the choice of the state $\pi = E^\dagger(\iota)$ associated with it.
Von Neumann subalgebras and their associated conditional expectations are crucial for understanding the structure of fixed point sets of CPTP maps [Wol12], particularly for semigroups in general and those considered in [Cap+24] in particular. We will therefore see them in the following section again.

1.3.3 A short introduction to semigroup theory for quantum systems

This section is devoted to the discussion of semigroups in both finite- and infinite-dimensional systems. We structure the discussion accordingly, as the infinite-dimensional setting raises issues that do not arise in the finite-dimensional case, whereas the latter—despite being less technically demanding—often allows for stronger results and a more developed theoretical framework.

Quantum Markov semigroups in finite dimensions—analysis of their asymptotics

We begin our discussion of finite-dimensional systems with one of the most prominent equations in physics, and arguably the most fundamental one: the Schrödinger equation, which is given as follows:

$$\frac{d}{dt}\rho = -i[H, \rho]\,, \qquad \rho(0) \in \mathcal{S}(\mathcal{H})\,. \tag{1.22}$$

Here, $\rho : \mathbb{R} \to \mathcal{S}(\mathcal{H})$ describes the (differentiable) time evolution of an initial state of the system $\rho(0)$, and $H \in \mathcal{B}(\mathcal{H})$ is a self-adjoint operator known as the Hamiltonian. This equation models the behaviour of a closed quantum system and, in doing so, has had profound impacts across almost all major fields of physics. One can readily write down the evolution operator $e^{-it[H,\cdot]} \equiv e^{-itH} \cdot e^{itH}$, which maps an initial state $\rho(0)$ to its time-evolved state $\rho(t)$, thereby uniquely solving eq. (1.22). The family $(e^{-it[H,\cdot]})_{t\in\mathbb{R}}$ forms a continuous one-parameter group of CPTP maps which does not distinguish between forward or backward propagation in time.
Although theoretically sensible for completely isolated systems, this ideal scenario is hard to realise in practice. A system in the laboratory, even if intended to be isolated, inevitably interacts with its environment ('the rest of the world'). Therefore, it is observed to eventually converge towards a thermal equilibrium state at an inverse temperature β in a time-irreversible process, determined by external factors such as the degree of isolation, cooling mechanisms and others. This motivates the need for a generalisation of the Schrödinger equation—one that reduces to the standard form for isolated systems, but inherently captures time-irreversibility by permitting only physically meaningful forward-time evolutions in the presence of external coupling (that is dissipation).
Starting from the solution to the Schrödinger equation (the unitary group $(e^{-it[H,\cdot]})_{t\in\mathbb{R}}$) and introducing minimal modifications to account for time-irreversibility leads to the concept of a QMS. This is a family of CPTP maps $(\mathcal{P}_t : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}))_{t\in\mathbb{R}_+}$ that additionally forms a continuous one-parameter semigroup, that is, satisfying:

1. $\mathcal{P}_0 = \mathrm{id}$,
2. $\mathcal{P}_s\,\mathcal{P}_t = \mathcal{P}_{s+t}$ for all $t, s \in \mathbb{R}_+$ (Semigroup property),
3. $t \mapsto \mathcal{P}_t$ is continuous on $\mathbb{R}_+$ (Continuity[2]).

The term 'Markov' in QMS refers to the fact that the future evolution depends only on the present state, not its past history.
Just as the Schrödinger equation describes evolution via the Hamiltonian H, there exists a one-to-one correspondence between such a QMS and a linear operator $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, called the generator, defined through the linear differential equation:

$$\frac{d}{dt}\rho = \mathcal{L}(\rho)\,, \qquad \rho(0) \in \mathcal{S}(\mathcal{H})\,. \tag{1.23}$$

The unique solution $\rho : \mathbb{R}_+ \to \mathcal{S}(\mathcal{H})$ for an initial state $\rho(0) \in \mathcal{S}(\mathcal{H})$ is given by $\rho(t) = \mathcal{P}_t(\rho(0))$, where the QMS can be shown to be equal to the exponential of the generator $\mathcal{P}_t = e^{t\,\mathcal{L}}$. Note that the exponential series suggests a natural extension of the semigroup to all of $\mathbb{R}$. However, the resulting group for negative times is no longer CP and thus does not map states to states, thereby stripping it of any interpretation as a physically viable process. Note further, that when considering QMS on infinite-dimensional systems, where the exponential series $e^{t\,\mathcal{L}}$ may be ill-defined, we will still retain this notation to denote the semigroup.

[2]Note that we will not discuss discrete time processes here, which are in principle possible, however.

Notably, in finite dimensions, the generator $\mathcal{L}$ can be expressed in a standard form, as demonstrated by Vittorio Gorini, Andrzej Kossakowski, and George Sudarshan [GKS76], and independently by Göran Lindblad [Lin76]. They showed that the generator $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ of a QMS can always be written as

$$\mathcal{L}(\rho) = -i[H, \rho] + \sum_{j=1}^{J} \left(L_j \rho L_j^* - \frac{1}{2}\{L_j^* L_j, \rho\} \right), \tag{1.24}$$

where $\{X, Y\} = XY + YX$ denotes the anti-commutator for $X, Y \in \mathcal{B}(\mathcal{H})$, the operators $L_j \in \mathcal{B}(\mathcal{H})$ (referred to as Lindblad or jump operators) are arbitrary, and $H \in \mathcal{B}(\mathcal{H})$ is self-adjoint. They also proved the converse: every operator $\mathcal{L}$ of the form eq. (1.24) generates a QMS. In recognition of this result, the expression is referred to as the GKSL form (also known as the Lindblad form) of the generator.

Even before this general characterisation, E.B. Davies, in a series of papers [Dav72; Dav73; Dav74; Dav76], explored specific, physically motivated scenarios. He considered a system coupled to a large heat bath (modelling the environment), with the combined entity evolving unitarily according to the Schrödinger equation. Davies demonstrated that, under certain assumptions about the bath and system-bath interaction (including a weak-coupling limit involving time rescaling), the resulting reduced dynamics of the system alone follows QMS dynamics. Despite the many underlying approximations, Davies dynamics remain significant to this day because they model the irreversible approach of a system towards its thermal equilibrium state (Gibbs state). It has been argued [KB16] that this dynamics even offer insights into thermalisation processes in nature, although this remains a topic of discussion and depends heavily on the validity of the approximations for a given physical system.

Independent of their role in thermalisation, these dynamics are of significant interest in quantum computing due to their application as state preparation mechanisms. If the semigroup dynamics can be implemented, they offer a means of preparing the Gibbs state at a given temperature. Setting aside the question of the usefulness of such Gibbs state preparation—which is discussed, for instance, in [Lin25] and will not be addressed here—the central concern for implementation is: How quickly does the system converge to its fixed point? This convergence rate has direct implications for the practical viability of such state preparation methods [Lin25]. Providing answers to this question for Davies dynamics (see section 1.4.3 for the definition), when applied to a particular subclass of Hamiltonians, is the focus of [Cap+24] and will be discussed at length in the associated sections 1.4.3, 2.3 and 3.3.

For now, we however, want to discuss these convergence questions in a broader framework of QMS $(e^{t\,\mathcal{L}})_{t\in\mathbb{R}_+}$ on $\mathcal{B}(\mathcal{H})$. Yet before we can discuss its speeds, we must first identify what the system converges to, that is, its fixed point(s) or invariant state(s), that is $\{X \in \mathcal{B}(\mathcal{H}) : e^{t\,\mathcal{L}}(X) = X \text{ for all } t \geq 0\}$.

For this purpose, it is useful to introduce the Heisenberg picture of the evolution. This is described by the adjoint semigroup $\left(e^{t\,\mathcal{L}^\dagger}\right)_{t\in\mathbb{R}_+} = \left(e^{t\,\mathcal{L}^\dagger}\right)_{t\in\mathbb{R}_+}$, that is the HS-adjoint, which transforms the semigroup of CPTP maps acting on states into a semigroup of unital CP maps acting on observables. The equivalent view is motivated by the fact that in quantum mechanics, access to system information is possible only through measurements, represented by observables (self-adjoint operators $M \in \mathcal{B}(\mathcal{H})$). The time evolution of the expectation value of such an observable—the sole quantity predicted by quantum mechanics—is then:

$$\mathbb{E}_{\rho_t}[M] = \mathrm{Tr}\big[e^{t\,\mathcal{L}}(\rho)M\big] = \mathrm{Tr}\Big[\rho e^{t\,\mathcal{L}^\dagger}(M)\Big]. \tag{1.25}$$

The choice between viewing the state evolving (Schrödinger picture) or the observable evolving (Heisenberg picture) is partly a matter of convention or convenience; often one picture (or a combination) simplifies the analysis, which is why we also use both perspectives.

As an example, analysing the fixed points is often facilitated by switching between pictures. First, note that the Cesàro mean, $\lim_{T\to\infty} \frac{1}{T}\int_0^T e^{t\,\mathcal{L}} dt$, defines a CPTP map that projects onto the fixed point set of the semigroup $(e^{t\,\mathcal{L}})_{t\in\mathbb{R}_+}$ [Wol12, Chapter 6.4]. This immediately guarantees the existence of at least one invariant state for any QMS in finite dimensions.

Unfortunately, it is not guaranteed that this fixed point is positive-definite. Thus, to further analyse the structure of this set and develop tools for studying convergence rates, we assume (for this finite-dimensional setting) the existence of a full-rank invariant state σ for the Schrödinger evolution ($e^{t\,\mathcal{L}}(\sigma) = \sigma$). An assumption that is naturally satisfied by the semigroups discussed in [Cap+24].

This existence of a full-rank invariant state σ of the Schrödinger dynamics implies that the set of fixed points of the Heisenberg evolution, $\{X \in \mathcal{B}(\mathcal{H}) : e^{t\,\mathcal{L}^\dagger}(X) = X \text{ for all } t \geq 0\}$, forms a von Neumann subalgebra of $\mathcal{B}(\mathcal{H})$. Furthermore, this subalgebra coincides with the commutant of the operators appearing in the GKSL-form, respectively the kernel subspace of $\mathcal{L}^\dagger$, that is

$$\{X \in \mathcal{B}(\mathcal{H}) : e^{t\,\mathcal{L}^\dagger}(X) = X \text{ for all } t \geq 0\} = \{H, L_k, L_k^\dagger : k = 1, \ldots, K\}' = \ker \mathcal{L}^\dagger \tag{1.26}$$

(see e.g., [Wol12, Theorem 7.2]). It is also straightforward to identify the matching conditional expectation onto this Neumann subalgebra—bearing in mind that conditional expectations are generally non-unique (see the discussion in section 1.3.2). In this case, it is given by the Cesàro mean of the Heisenberg evolution:

$$E = \lim_{T\to\infty} \frac{1}{T} \int_0^T e^{t\,\mathcal{L}^\dagger}\, dt\,, \tag{1.27}$$

clearly satisfying $Ee^{t\,\mathcal{L}^\dagger} = e^{t\,\mathcal{L}^\dagger}E = E$ for all $t \in \mathbb{R}^+$ equivalent to the analogous equalities for their HS-adjoints. The required fixed point condition immediately gives

$$E^\dagger(\sigma) = \lim_{T\to\infty} \frac{1}{T} \int_0^T e^{t\,\mathcal{L}}(\sigma)\, dt = \sigma\,, \tag{1.28}$$

meaning that all the properties established in theorem 1.3.2 apply, and consequently the relations in theorem 1.3.1 also hold. The fixed point σ can furthermore be used to define various inner products on $\mathcal{B}(\mathcal{H})$, in addition to the HS-inner product, thereby enabling the application of Hilbert space methods to the analysis of the semigroup. In this context, the literature frequently uses the terms symmetric and symmetry to refer to self-adjointness with respect to one of these inner products; we will adopt both terminologies throughout.

For our purposes, we concentrate on the two most important examples—both of which were already introduced implicitly in the previous section (section 1.3.2). For a more detailed treatment, including other types and properties of such inner products, we refer the interested reader to [CM17]. The first, known as the KMS-inner product, is defined as:

$$\langle X, Y\rangle_{\sigma,1/2} \equiv \mathrm{Tr}\Big[X^*\sigma^{1/2}Y\sigma^{1/2}\Big] \quad \text{with corresponding norm} \quad \|X\|_{\sigma,1/2} \equiv \sqrt{\langle X, X\rangle_{\sigma,1/2}}\,. \tag{1.29}$$

The second, the GNS-inner-product, is given by:

$$\langle X, Y\rangle_{\sigma} \equiv \mathrm{Tr}[X^*Y\sigma] \quad \text{with corresponding norm} \quad \|X\|_{\sigma} \equiv \sqrt{\langle X, X\rangle_{\sigma}}\,, \tag{1.30}$$

where both definitions hold for all $X, Y \in \mathcal{B}(\mathcal{H})$.

Before turning to their applications, let us clarify that when referring to KMS- or GNS-symmetry of a generator $\mathcal{L}$ of a QMS in the Schrödinger picture, we actually mean that this property holds for its Hilbert-Schmidt adjoint $\mathcal{L}^\dagger$ in the Heisenberg picture. We will aim for precision in this distinction but also will regularly call a QMS $\mathcal{L}$ KMS- or GNS-symmetric, despite meaning its HS-adjoint.

Now the significance of the KMS-inner product lies in the possibility of symmetrising the generator $\mathcal{L}^\dagger$ whilst ensuring that corresponding Schrödinger evolution still is a QMS. The corresponding symmetry requirement (KMS-symmetry) furthermore is generally weaker than the GNS one (as we will discuss below), which allows for more flexibility, for example, when constructing Gibbs sampling semigroups (see [Che+23; CKG23; DLL25b]). If not already present in $\mathcal{L}^\dagger$ one can perform the aforementioned symmetrisation by defining the generator:

$$\widehat{\mathcal{L}}^\dagger = \frac{1}{2}(\mathcal{L}^\dagger + \Gamma_\sigma^{-1}\,\mathcal{L}\,\Gamma_\sigma)\,,$$

where $\Gamma_\sigma(X) \equiv \sigma^{1/2}X\sigma^{1/2}$ and $\mathcal{L}$ denotes the HS-adjoint of $\mathcal{L}^\dagger$, i.e., the generator in Schrödinger picture. It can be readily verified that $\Gamma_\sigma\,\mathcal{L}^\dagger\,\Gamma_\sigma^{-1}$ generates a QMS by checking the defining properties for $(e^{t\Gamma_\sigma\,\mathcal{L}^\dagger\,\Gamma_\sigma^{-1}})_{t\in\mathbb{R}_+}$. Consequently, $\widehat{\mathcal{L}}$, as a convex combination of two generators in GKSL form, possesses the GKSL-form and thus generates a QMS, which by construction has a HS-adjoint $\widehat{\mathcal{L}}^\dagger$that is KMS-symmetric.

This means both $\widehat{\mathcal{L}}^\dagger$ and $\widehat{\mathcal{L}}$ are in particular diagonalisable with real, non-positive eigenvalues. Their spectral gap, denoted $\lambda(\widehat{\mathcal{L}}^\dagger)$, which quantifies the slowest decay rate towards the fixed point space, can be studied via the variational characterisation:

$$\lambda(\widehat{\mathcal{L}}^\dagger) = \inf_{\substack{X \in \mathcal{B}(\mathcal{H}),\\ (\mathrm{id} - \widehat{E})(X) \neq 0}} \frac{\left\langle X, -\widehat{\mathcal{L}}^\dagger(X) \right\rangle_{\sigma,1/2}}{\left\| (\mathrm{id} - \widehat{E})(X) \right\|_{\sigma,1/2}^2} . \tag{1.31}$$

The above definition clearly extends beyond $\widehat{\mathcal{L}}^\dagger$ to KMS-symmetric (and hence GNS-symmetric) generators in general as in those cases $\widehat{\mathcal{L}}^\dagger = \mathcal{L}^\dagger$ and $\widehat{E} = E$.
The condition of KMS-symmetry—often also referred to as the detailed balanced (DB) condition—is frequently assumed or explicitly constructed [Che+23; CKG23; DLL25b], and thus need not be artificially imposed by symmetrising the generator. For instance, the Davies generator considered in [Cap+24] inherently satisfies the DB condition and, in fact, exhibits the stronger property of GNS-symmetry. This stronger form of symmetry entails that if $\mathcal{L}^\dagger$ (the HS-adjoint of a QMS generator $\mathcal{L}$) as in the case of Davies generators, is GNS-symmetric with respect to a positive-definite state σ, then it is necessarily KMS-symmetric with respect to the same state. Moreover, it commutes with the modular operator $\Delta_\sigma(X) = \sigma X \sigma^{-1}$, for all $X \in \mathcal{B}(\mathcal{H})$, as discussed in the proof of theorem 1.3.1. This commutation relation has particularly far-reaching implications for the analysis of MLSI and CMLSI, as all known results in the literature (to the best of the author's knowledge) rely on this structural property. We will revisit these inequalities, namely, MLSI and CMLSI, later in this section, after introducing the concept of mixing time, which provides the primary motivation for their study.
For semigroups generated by either GNS- or KMS-symmetric generators $\mathcal{L}^\dagger$, we may now derive a preliminary estimate on the rate of convergence to the fixed point using their spectral gap. Combined with Grönwall's inequality, $\lambda(\mathcal{L}^\dagger)$ (as defined in eq. (1.31)) yields an exponential decay estimate in the KMS-norm:

$$\left\| (e^{t\,\mathcal{L}^\dagger} - E)(X) \right\|_{\sigma,1/2} \leq e^{-\frac{t\lambda(\mathcal{L}^\dagger)}{2}} \left\| (\mathrm{id} - E)(X) \right\|_{\sigma,1/2} .$$

Although of independent interest, the above decay behaviour often serves primarily as a tool to estimate a more relevant quantity in this context: the mixing time of the QMS. The mixing time is defined in terms of the trace distance

$$T(\rho,\sigma) \equiv \frac{1}{2}\|\rho - \sigma\|_1 \tag{1.32}$$

as follows:

$$t_{\mathrm{mix}}(\mathcal{L};\varepsilon) \equiv \inf\left\{ t \geq 0 : \sup_{\rho \in \mathcal{S}(\mathcal{H})} T(e^{t\,\mathcal{L}}(\rho), E^\dagger(\rho)) \leq \varepsilon \right\} . \tag{1.33}$$

By relating the KMS-norm decay to the TD decay (using standard norm inequalities like Hölder's inequality), one can derive the mixing time bound:

$$t_{\mathrm{mix}}(\mathcal{L};\varepsilon) \leq \frac{2}{\lambda(\mathcal{L}^\dagger)} \log \frac{\left\|\sigma^{-1/2}\right\|_\infty}{\varepsilon} . \tag{1.34}$$

Note that since GNS-symmetry implies KMS-symmetry, the spectral gap defined through the KMS-inner product (see eq. (1.31)) and hence also the above estimate works in both scenarios which motivates the use of the KMS framework, as it provides a unified approach for discussing both KMS- and GNS-symmetric QMS simultaneously.
Let us briefly comment on the use of the TD in the definition of mixing time, and highlight one important consequence of this choice. The trace distance is employed due to its operational interpretation—rooted in the measurement formalism of quantum mechanics (see, e.g., [Tom16, Chapter 7])—as well as its status as a norm, its monotonicity under quantum channels (DPI), and its role as the quantum analogue of the classical total variation distance. The property we now present illustrates how mixing times for different powers of a target accuracy relate, thereby formalising what one would naturally expect. As we could not find a suitable reference, we provide a self-contained proof:

Lemma 1.3.4 *Let $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be the generator of a QMS, and let $\varepsilon > 0$. Then, for every $n \in \mathbb{N}$, one has*

$$t_{\mathrm{mix}}(\mathcal{L}; \varepsilon^n) \leq n \cdot t_{\mathrm{mix}}(\mathcal{L}; \varepsilon)\,.$$

Proof. Let $\rho \in \mathcal{S}(\mathcal{H})$, and set $\rho' = e^{(n-1)t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho)$. Consider the decomposition

$$\rho' - E^\dagger(\rho) = (\rho' - E^\dagger(\rho))_+ - (\rho' - E^\dagger(\rho))_-\,,$$

where $(\cdot)_\pm$ denote the positive and non-positive parts of a self-adjoint operator. Since $E^\dagger$ is a projection $(E^\dagger)^2 = E^\dagger$ and satisfies $E^\dagger e^{(n-1)t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}} = E^\dagger$ (following both from eq. (1.27)), we obtain

$$E^\dagger((\rho' - E^\dagger(\rho))_+) = E^\dagger((\rho' - E^\dagger(\rho))_-)\,,$$

since

$$E^\dagger((\rho' - E^\dagger(\rho))_+ - (\rho' - E^\dagger(\rho))_-) = E^\dagger(\rho' - E^\dagger(\rho)) = 0\,.$$

Let $\delta = T(\rho', E^\dagger(\rho))$, and write $(\rho' - E^\dagger(\rho))_+ = \delta\mu$, $(\rho' - E^\dagger(\rho))_- = \delta\nu$, whereby construction $\mu, \nu \in \mathcal{S}(\mathcal{H})$. Then:

$$\begin{aligned}
T(e^{nt_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho), E^\dagger(\rho)) &= \tfrac{1}{2}\left\| e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho' - E^\dagger(\rho))\right\|_1 \\
&= \tfrac{\delta}{2}\left\| e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\mu - \nu - E^\dagger(\mu - \nu))\right\|_1 \\
&\leq \delta \cdot \max\left\{ T(e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\mu), E^\dagger(\mu)),\, T(e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\nu), E^\dagger(\nu))\right\} \\
&\leq \delta \cdot \varepsilon\,.
\end{aligned}$$

The final inequality follows from the definition of the mixing time $t_{\mathrm{mix}}(\mathcal{L}; \varepsilon)$ for all states in $\mathcal{S}(\mathcal{H})$. Now iterating this argument n times yields

$$T(e^{nt_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho), E^\dagger(\rho)) \leq \varepsilon^n\,.$$

Since $\rho \in \mathcal{S}(\mathcal{H})$ was arbitrary, the claim follows. □

Returning to our estimate of the mixing time by the spectral gap, we note that despite its relative simplicity, this approach has a disadvantage: the resulting bound depends logarithmically on $\left\|\sigma^{-1}\right\|_\infty$, the inverse of the smallest eigenvalue of σ. This dependence is problematic, as even the best case scaling of the bound (achieved for $\sigma = \iota$) is $\log d_{\mathcal{H}}$. To achieve better dimensional scaling, researchers employ techniques based on logarithmic Sobolev inequalities (LSIs), specifically the MLSI and CMLSI.
To translate these MLSI and CMLSI, which characterise contraction in relative entropy, into an estimate on the mixing time, one must bound the trace distance in terms of the relative entropy. This is achieved using Pinsker's inequality [Tom16], which asserts that for any states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$:

$$\|\rho - \sigma\|_1 \leq \sqrt{2D(\rho\|\sigma)}\,. \tag{1.35}$$

Now if one can establish an exponential decay for the relative entropy towards the relevant fixed point $E^\dagger(\rho)$, such as

$$D(e^{t\,\mathcal{L}}(\rho)\|E^\dagger(\rho)) \leq e^{-\alpha t} D(\rho\|E^\dagger(\rho)) \tag{1.36}$$

for all $t \geq 0$ and some $\alpha > 0$, Pinsker's inequality leads to an improved mixing time bound. Specifically, combining (1.35) and (1.36) with the bound $D(\rho\|E^\dagger(\rho)) \leq D(\rho\|\sigma) \leq \log\left\|\sigma^{-1}\right\|_\infty$, where the first inequality follows from the chain-rule (eq. (1.20)), one gets:

$$\begin{aligned}
\left\| e^{t\,\mathcal{L}}(\rho) - E^\dagger(\rho)\right\|_1 &\leq \sqrt{2D(e^{t\,\mathcal{L}}(\rho)\|E^\dagger(\rho))} \\
&\leq e^{-\alpha t/2}\sqrt{2D(\rho\|E^\dagger(\rho))} \\
&\leq e^{-\alpha t/2}\sqrt{2\log\|\sigma^{-1}\|_\infty}\,.
\end{aligned}$$

This yields a mixing time estimate:

$$t_{\mathrm{mix}}(\mathcal{L};\varepsilon) \leq \frac{2}{\alpha} \log \frac{\sqrt{\log \|\sigma^{-1}\|_\infty}}{\varepsilon}\,, \tag{1.37}$$

that has a better double-logarithmic dependence on $\left\|\sigma^{-1}\right\|_\infty$ compared to the single logarithm from the spectral gap method. Now an inequality of the form (1.36), or equivalently, its differential form

$$\alpha D(\rho\|E^\dagger(\rho)) \leq -\operatorname{Tr}\left[\mathcal{L}(\rho)(\log\rho - \log E^\dagger(\rho))\right] \equiv \mathrm{EP}_{\mathcal{L}}(\rho)\,, \tag{1.38}$$

holding uniformly for all positive-definite[3] $\rho \in \mathcal{S}(\mathcal{H})$ with some $\alpha > 0$ first appeared in [KT13] and is called a MLSI. The optimal constant satisfying this inequality is the MLSI constant defined as

$$\alpha(\mathcal{L}) = \inf_{\rho\in\mathcal{S}(\mathcal{H}):\rho>0} \frac{\mathrm{EP}_{\mathcal{L}}(\rho)}{D(\rho\|E^\dagger(\rho))}\,,$$

and the term $\mathrm{EP}_{\mathcal{L}}(\rho) \equiv -\operatorname{Tr}[\mathcal{L}(\rho)(\log\rho - \log E(\rho))]$ is known as the entropy production (EP). The addition 'modified' serves to distinguish it from the standard LSI considered in [TPK14], which we do not discuss here, as the MLSI offers advantages in the quantum setting (see, for example, the discussion in [Rou24]).

Let us briefly comment on the gauge freedom of the EP, as this is an important property not only used in our work [Cap+24], but also featured widely throughout the literature. We again could not find a suitable form of this statement in the literature, so we want to give a self-contained proof as well.

Lemma 1.3.5 (Gauge freedom of the entropy production) *Let $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be the generator of a QMS with a positive-definite invariant state. Then, for $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ both positive-definite,*

$$\mathrm{EP}_{\mathcal{L}}(\rho) = -\operatorname{Tr}\left[\mathcal{L}(\rho)\left(\log\rho - \log\left(E^\dagger(\sigma)\right)\right)\right]\,, \tag{1.39}$$

that is the entropy production is invariant under the choice of reference state in the second logarithmic argument, provided it is composed with $E^\dagger = \lim_{T\to\infty} \frac{1}{T}\int_0^T e^{t\,\mathcal{L}}\,dt$, the HS-adjoint of the conditional expectation $E = \lim_{T\to\infty} \frac{1}{T}\int_0^T e^{t\,\mathcal{L}^\dagger}\,dt$ onto $\ker \mathcal{L}^\dagger$.

Proof. First note that from $E^\dagger$ having a positive-definite fixed point, we can conclude (theorem 1.3.2) that it maps positive-definite operators to positive-definite operators, ensuring that both functionals in eq. (1.39) are well-defined in the first place. By theorem 1.3.1, one may write $E^\dagger(\rho) = \pi^{1/2}X\pi^{1/2}$, $E^\dagger(\sigma) = \pi^{1/2}Y\pi^{1/2}$, where $\pi = E^\dagger(\iota)$, $X = E(\pi^{-1/2}\rho\pi^{-1/2}) > 0$, and $Y = E(\pi^{-1/2}\sigma\pi^{-1/2}) > 0$. Since π commutes with the image of E, that is with the von Neumann subalgebra $\mathcal{N}$ (theorem 1.3.1), we obtain

$$\begin{aligned}\mathrm{EP}_{\mathcal{L}}(\rho) &= -\operatorname{Tr}\left[\mathcal{L}(\rho)\left(\log\rho - \log\left(E^\dagger(\rho)\right)\right)\right]\\ &= -\operatorname{Tr}\left[\mathcal{L}(\rho)\left(\log\rho - \log\left(E^\dagger(\sigma)\right)\right)\right] + \operatorname{Tr}[\mathcal{L}(\rho)(\log X - \log Y)]\,.\end{aligned}$$

Now, taking the HS-adjoint of $\mathcal{L}$ and using that it preserves hermiticity yields

$$\operatorname{Tr}[\mathcal{L}(\rho)(\log X - \log Y)] = \operatorname{Tr}\left[\rho\,\mathcal{L}^\dagger(\log X - \log Y)\right]\,.$$

Since $\log X - \log Y \in \mathcal{N}$, and $\mathcal{N}$ is the commutant of the operators appearing in the GKSL representation of $\mathcal{L}$ (eq. (1.26)), we find

$$\mathcal{L}^\dagger(\log X - \log Y) = (\log X - \log Y)\,\mathcal{L}^\dagger(\mathbb{1}) = 0\,,$$

which concludes the proof. □

[3] This is because a QMS that has a full-rank invariant state σ preserves positive definiteness, even in the limit $t \to \infty$. This can be shown using that for a given $X > 0$ there exists $c > 0$ s.t. $X \geq c\sigma$ hence $e^{t\,\mathcal{L}}(X) \geq ce^{t\,\mathcal{L}}(\sigma) = c\sigma > 0$. Having the decay eq. (1.36) for all positive-definite states and hence in particular for properly normalised $\rho + \varepsilon\sigma$ for $\varepsilon > 0$ one can lift it to all states by upper semicontinuity of $\varepsilon \mapsto D(\rho + \varepsilon\sigma\|E^\dagger(\rho) + \varepsilon\sigma)$ [HM17, Proposition 3.8 (ii)] taking the limit $\varepsilon \searrow 0$.

Finally, building upon the MLSI, we introduce the concept of the CMLSI, whose constant $\alpha_c(\mathcal{L})$ is defined via a 'completion' procedure:

$$\alpha_c(\mathcal{L}) \equiv \inf_{n\in\mathbb{N}} \alpha(\mathrm{id}_n \otimes \mathcal{L})\,. \tag{1.40}$$

Here, $\mathrm{id}_n \otimes \mathcal{L}$ acts on the extended space $\mathcal{B}(\mathbb{C}^n \otimes \mathcal{H})$. This type of 'completion', involving tensor products with identity maps on ancillary systems, is common in quantum information theory. It is often necessary to ensure that quantum objects behave analogously to their classical counterparts. For instance, the classical MLSI constant satisfies the property $\alpha(\mathcal{L}+\mathcal{L}') \geq \min\{\alpha(\mathcal{L}), \alpha(\mathcal{L}')\}$. However, this inequality fails for the MLSI constant $\alpha(\mathcal{L})$ for generators of QMS (see [Rou24]). By using the complete constant $\alpha_c(\mathcal{L})$, this desirable property is recovered:

$$\alpha_c(\mathcal{L}+\mathcal{L}') \geq \min\{\alpha_c(\mathcal{L}), \alpha_c(\mathcal{L}')\}\,.$$

In contrast, for the spectral gap of a KMS-symmetric semigroup, no such completion is necessary: the standard definition $\lambda(\mathcal{L}^\dagger)$ and its completed counterpart $\inf_{n\in\mathbb{N}_0} \lambda(\mathrm{id}_n \otimes \mathcal{L}^\dagger)$ coincide, as the spectral properties remain unchanged under tensorisation with the identity.
The remaining objective is to find $\lambda(\mathcal{L}^\dagger)$ and $\alpha(\mathcal{L})$, or leveraging a combination of both, as we will explore in [Cap+24], to derive mixing time estimates for specific models like the Davies semigroups.

Quantum Markov semigroups in infinite dimensions—generation theory

We now switch gears and discuss the generation theory for QMS on infinite-dimensional systems as promised before. Specifically, we consider systems described by trace-class operators on an infinite-dimensional, separable Hilbert space. Since all such Hilbert spaces are isometrically isomorphic to $\mathcal{F}$, our discussion, while appearing focused on $\mathcal{F}$, applies generally to this class of systems. It should be noted, however, that some foundational results, such as the Hille-Yosida [HP12; Yos48] and Lumer-Phillips [LP61] theorems, were originally derived in the more general context of Banach spaces and when stating them we will also do it in all generality.
Analogous to the finite-dimensional case discussed previously, we begin with the Schrödinger equation involving a densely defined linear operator H with domain $\mathcal{D}(H) \subseteq \mathcal{F}$. We deliberately exclude the simpler case where H is a bounded operator ($H \in \mathcal{B}(\mathcal{F})$), as the corresponding semigroup can then be directly defined via its exponential series. For the more general case of an unbounded generator, the solution to the abstract Cauchy problem:

$$\frac{d}{dt}\rho = -i[H,\rho], \quad \rho(0) \in \mathcal{D}([H,\cdot]) \subseteq \mathcal{T}(\mathcal{F}) \tag{1.41}$$

in the form of a differentiable map $\rho : \mathbb{R} \to \mathcal{D}([H,\cdot])$ remaining in the domain $\mathcal{D}([H,\cdot]) = \{X : X \in \mathcal{T}(\mathcal{F}),\ [H,X] \in \mathcal{T}(\mathcal{F})\}$, is fully characterised by Stone's theorem. This theorem states that an operator $(-iH, \mathcal{D}(H))$ generates a strongly continuous one-parameter unitary group $(e^{-itH})_{t\in\mathbb{R}}$ on $\mathcal{F}$ if, and only if, $(H, \mathcal{D}(H))$ is a self-adjoint operator [Sto30; Sto32]. Although Stone's theorem is formulated on the Hilbert space $\mathcal{F}$, the resulting unitary group can be directly lifted to the space of trace-class operators $\mathcal{T}(\mathcal{F})$ by defining the corresponding evolution superoperator as $(e^{-it[H,\cdot]} \equiv e^{-itH}(\cdot)e^{itH})_{t\in\mathbb{R}}$. This lifted group retains the essential properties (strong continuity, group structure and complete positivity) and provides the solution to the abstract Cauchy problem eq. (1.41), with $(-i[H,\cdot], \mathcal{D}([H,\cdot]))$ serving as its generator.
Stone's theorem has been generalised in several directions. Most relevant to our discussion are the Hille-Yosida and Lumer-Phillips theorems, which address the generation of strongly continuous semigroups (C_0-semigroups) associated with linear differential equations on general Banach spaces. As mentioned earlier, we will briefly adopt this more general Banach space framework because it allows for a unified treatment covering the Hilbert space $(\mathcal{F}, \|\cdot\|_{\mathcal{F}})$, the space of trace-class operators $(\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$, and the QSSs discussed in [GMR24].
Let us first clarify some terminology, including a precise definition of a C_0-semigroup on a Banach space $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$. This definition encompasses the notion of a strongly continuous one-parameter unitary group in positive time, mentioned previously. A family of bounded linear maps $(\mathcal{P}_t : \mathcal{X} \to \mathcal{X})_{t\in\mathbb{R}_+} \subset \mathcal{B}(\mathcal{X})$ is called a C_0-semigroup if it satisfies the following properties:

1. $\mathcal{P}_0 = \mathrm{id}$,
2. $\mathcal{P}_s \mathcal{P}_t = \mathcal{P}_{s+t}$ for all $t, s \in \mathbb{R}_+$ (Semigroup property),
3. $t \mapsto \mathcal{P}_t(X)$ is continuous as map $\mathbb{R}_+ \to \mathcal{X}$ for all $X \in \mathcal{X}$ (Strong continuity).

Analogous to the finite-dimensional setting, every C_0-semigroup is uniquely associated with a densely defined closed linear operator $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$, known as its generator [Nag00] where the C_0-semigroup provides the unique solution to the abstract Cauchy problem:

$$\frac{d}{dt} X = \mathcal{O}(X)\,, \quad X(0) \in \mathcal{D}(\mathcal{O})\,, \tag{1.42}$$

as a differentiable function $X : \mathbb{R}_+ \to \mathcal{D}(\mathcal{O}) \subseteq \mathcal{X}$. This connection of the generator and the associated semigroup is established by the celebrated Hille-Yosida theorem (see [Nag00]), which provides necessary and sufficient conditions on $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ for it to generate a C_0-semigroup. These conditions involve the resolvent operator $R(\lambda, \mathcal{O}) = (\lambda\,\mathrm{id} - \mathcal{O})^{-1}$, which is defined for $\lambda \in \mathbb{C}$ such that the inverse to $(\lambda - \mathcal{O}, \mathcal{D}(\mathcal{O}))$ exists and is a bounded operator on $\mathcal{X}$. The set of such λ is called the resolvent set of $\mathcal{O}$. The Hille-Yosida theorem states that a linear operator $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ generates a C_0-semigroup $(e^{t\,\mathcal{O}})_{t\in\mathbb{R}_+}$ satisfying[4] $\left\|e^{t\,\mathcal{O}}\right\|_{\mathcal{B}(\mathcal{X})} \leq Me^{\omega t}$ for some constants $M \geq 0$ and $\omega \in \mathbb{R}$ if and only if:

1. $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ is densely defined and closed.
2. The resolvent set contains the interval (ω, ∞), and for all $\lambda > \omega$ and all integers $n \geq 1$:

$$\|R(\lambda, \mathcal{O})^n\|_{\mathcal{B}(\mathcal{X})} \leq \frac{M}{(\lambda - \omega)^n}\,.$$

Subsequently, the Lumer-Phillips theorem (see [Nag00]) provided conditions which are often easier to verify in practice, particularly for contractive or quasi-contractive semigroups (where the bound involves $M = 1$) like those studied in our [GMR24]. It states that a densely defined $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ is the core to a generator of a C_0-semigroup satisfying $\left\|e^{t\overline{\mathcal{O}}}\right\|_{\mathcal{B}(\mathcal{X})} \leq e^{\omega t}$ if and only if:

1. There exists an $\omega \in \mathbb{R}$ such that the operator $(\mathcal{O} - \omega\,\mathrm{id})$ is dissipative, (meaning, for all $\lambda > 0$ and $X \in \mathcal{D}(\mathcal{O})$, it holds $\lambda\|X\|_{\mathcal{X}} \leq \|(\lambda\,\mathrm{id} - (\mathcal{O} - \omega\,\mathrm{id}))(X)\|_{\mathcal{X}}$).
2. The range of $(\lambda_0\,\mathrm{id} - \mathcal{O})$ is dense in $\mathcal{X}$ for some $\lambda_0 > \omega$.

Although these theorems fully characterise the generation theory of semigroups, applying them to concrete settings still requires explicit verification of their conditions. Ideally, exploiting additional structure, such the GKSL-form or even further assumptions on the involved Hamiltonian and jump operators, would allow these abstract conditions to be translated into simpler, more direct criteria. The work in [GMR24] investigates this precise question, specifically within the context of trace-class operators on the bosonic system $\mathcal{F}$. The goal is to identify readily verifiable conditions under which an operator in GKSL-form—such as in eq. (1.24), but involving an unbounded Hamiltonian and jump operators constructed from the annihilation and creation operators $a, a^\dagger$, defined on a suitable domain—serves as a core for the generator of a C_0-semigroup of CPTP maps, i.e., a QMS.
We are aware of two main lines of work that have leveraged generalisations of the GKSL-form to the unbounded setting, in an attempt to show that this form is already sufficient to conclude that the operator itself or its closure are a generator of a QMS. The first line of work, developed by researchers including Fagnola and Chebotarev [CF98; Fag18], addresses the problem within the Heisenberg picture. The second line, initiated by Davies [Dav77] and subsequently applied by others [ASR16], works in the Schrödinger picture, which is the framework adopted for this discussion. Both lines of work encounter complications related to the loss of probability in the semigroups they obtain, unless they make further assumptions beyond just a suitable generalisation of the GKSL-form to the unbounded setting. We will discuss this problem further below.

[4]This indeed is not a constraint but satisfied by all C_0-semigroups [Nag00, Proposition 5.5.]

Consistent with the methodology in [GMR24], and reflecting a preference for the Schrödinger picture, we will follow Davies' strategy up to the emergence of the issue with the probability loss in the so called 'minimal semigroup', while streamlining some explanations. The resolution of this minimal semigroup problem and the inspiration that lead to its resolution in our specific context (see [GMR24]) is deferred to sections 1.4.4, 2.4 and 3.4.

Now in [Dav77], Davies studies the well-posedness of a linear differential equation also known as the quantum Fokker-Planck equation. Specifically, Davies considers a generator of the form:

$$\mathcal{L}(\rho) = G\rho + \rho G^* + \sum_{j=1}^{J} L_j \rho L_j^* \equiv \mathcal{G}(\rho) + \Sigma(\rho)\,, \tag{1.43}$$

and implicitly defines the domain of $\mathcal{L}$ as $\mathcal{D}(\mathcal{L}) = \operatorname{span}\{|\psi\rangle\langle\phi| : \psi, \phi \in \mathcal{D}(G) \subseteq \mathcal{F}\}$. The above form of the generator can be related to eq. (1.24) by identifying $G = -iH - \frac{1}{2}\sum_{j=1}^{J} L_j^* L_j$.

One of his key assumptions is that the pair $(G, \mathcal{D}(G))$ already generates a contractive, C_0-semigroup $(e^{tG})_{t\in\mathbb{R}+}$ on $\mathcal{F}$, meaning in particular $\left\|e^{tG}\right\|_\infty \leq 1$. Under this assumption he then lifts this C_0-semigroup to a contractive, CP, C_0-semigroup on $\mathcal{T}(\mathcal{F})$, defined by

$$\left(e^{t\,\mathcal{G}}(\cdot) \equiv e^{tG}(\cdot)e^{tG^*}\right)_{t\in\mathbb{R}_+},$$

and generated by $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$, with $(\mathcal{G}, \mathcal{D}(\mathcal{L}))$ serving as a core. We omit the detailed discussion of the domain $\mathcal{D}(\mathcal{G})$ and the core property of $(\mathcal{G}, \mathcal{D}(\mathcal{L}))$, referring the reader to [Dav77] for further details.

To verify the conditions of the Hille-Yosida or Lumer-Phillips theorem, Davies then introduces a perturbed generator parameterised by $\delta \in [0, 1)$:

$$\mathcal{L}_\delta(\rho) \equiv \mathcal{G}(\rho) + \delta\,\Sigma(\rho)\,, \tag{1.44}$$

with domain $\mathcal{D}(\mathcal{L}_\delta) = \mathcal{D}(\mathcal{G})$. To establish the well-posedness of this generator, he crucially exploits properties of $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$. More precisely, he employs the following identity on $\mathcal{D}(\mathcal{G})$:

$$\lambda\,\mathrm{id} - \mathcal{L}_\delta = (\mathrm{id} - \delta\,\Sigma\,R(\lambda, \mathcal{G}))(\lambda\,\mathrm{id} - \mathcal{G}), \quad \lambda > 0\,,$$

where $\mathcal{A}_\lambda \equiv \Sigma R(\lambda, \mathcal{G})$ is a bounded and contractive operator on $\mathcal{T}(\mathcal{F})$. Boundedness and contractivity follow from eq. (1.43), with the fact that $(\mathcal{G}, \mathcal{D}(\mathcal{L}))$ forms a core of $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$ combined with complete positivity of $R(\lambda, \mathcal{G})$ and $(\Sigma, \mathcal{D}(\mathcal{L}))$ and hence $\mathcal{A}_\lambda$. Since $\delta \in [0, 1)$ and $\|\mathcal{A}_\lambda\|_{1\to1} \leq 1$, the operator $(\mathrm{id} - \delta\,\mathcal{A}_\lambda)$ is invertible with a bounded inverse given by the convergent von Neumann series $\sum_{n=0}^{\infty}(\delta\,\mathcal{A}_\lambda)^n$. The resolvent of the perturbed generator is therefore given by

$$R(\lambda, \mathcal{L}_\delta) = R(\lambda, \mathcal{G})(\mathrm{id} - \delta\,\mathcal{A}_\lambda)^{-1} = R(\lambda, \mathcal{G}) \sum_{n=0}^{\infty} \delta^n\,\mathcal{A}_\lambda^n\,.$$

Moreover, as both $R(\lambda, \mathcal{G})$ and $\mathcal{A}_\lambda$ are CP maps, so is the resolvent $R(\lambda, \mathcal{L}_\delta)$ which directly implies the same property for its associated semigroup (see e.g., [Nag00] for the resolvent representation of semigroups), whose existence we have yet to prove. Ignoring domain and closure subtleties, we finally observe that $\mathrm{Tr}[\mathcal{L}(\rho)] = 0$ implies $\mathrm{Tr}[\mathcal{L}_\delta(\rho)] \leq 0$ for all $\rho \in \mathcal{D}(\mathcal{L}_\delta)$, and hence, for $\lambda > 0$,

$$\lambda\|\rho\|_1 = \lambda\,\mathrm{Tr}[\rho] \leq \mathrm{Tr}[(\lambda - \mathcal{L}_\delta)(\rho)] \leq \|(\lambda - \mathcal{L}_\delta)(\rho)\|_1\,.$$

Combined with the complete positivity of the resolvent, this yields the bound $\|R(\lambda, \mathcal{L}_\delta)\|_{1\to1} \leq \frac{1}{\lambda}$. We reiterate that this discussion omits several technical details, which are thoroughly treated in both [Dav77] and [GMR24].

This settled the last requirement for the application of the Hille-Yosida theorem with $M = 1, \omega = 0$ (or Lumer-Phillips) for $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$. Therefore, for each $\delta \in [0, 1)$, this operator generates a contractive and CP C_0-semigroup on $\mathcal{T}(\mathcal{F})$ while further having $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}))$ forming a core. The remaining challenge, now, is that this result holds only for $\delta < 1$, whereas the physically relevant generator corresponds to $\delta = 1$.

By considering the limit as $\delta \to 1$, Davies constructs a limiting semigroup, he terms the 'minimal semigroup'

(or minimal solution). This minimal semigroup is guaranteed to be contractive and CP, but it is not necessarily trace-preserving. If trace preservation is violated, that is, if $\mathrm{Tr}\big[\mathcal{P}_t^{\min}(\rho)\big] < \mathrm{Tr}[\rho]$ for some $t > 0$ and state $\rho \in \mathcal{T}(\mathcal{F})$, then the original operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ does not have a unique closure that generates a C_0-semigroup. In such cases, there exist multiple trace-preserving extensions, which can be parametrised by a reset state (see [Dav77] and the discussion in section 3.4.3). Interestingly, in the same paper, Davies provides evidence (specifically, the divergence of the number operator in a pure birth process [Dav77, Example 3.3.]) that this failure of trace preservation coincides with the divergence of physical observables in finite time under the evolution generated by the minimal semigroup. This co-occurrence, among other considerations, motivated the specific assumption imposed in [GMR24], which ensures the stability of physical observables (in this case, the moments of the number operator) under the evolution and as a result, trace preservation of the minimal semigroup, as we will later see (section 3.4).
This concludes our overview of the more general semigroup generation theory. We will revisit these concepts when discussing the specific physical setup in section 1.4.4, clarify the objectives in section 2.4 and present the main results of [GMR24] in section 3.4.

1.3.4 Hamiltonians on lattice systems

In this section, we introduce Hamiltonians on lattice systems, i.e., on $\mathbb{Z}^D$. More precisely, we present the mathematical framework that allows one to treat systems of growing but finite size, where lower levels of this hierarchy are naturally embedded into higher ones, eventually approximating the infinite lattice $\mathbb{Z}^D$. This is motivated as follows: Consider a theoretical model built from elementary components of fixed size (atoms in a solid, for example, or the local checks of a quantum code) and construct systems of increasing size from these building blocks. The smaller systems should be naturally embedded within the larger ones, capturing the idea that one starts with a small system and enlarges it by successively adding components. The systems one typically aims to understand theoretically have relatively large system sizes, and a key goal is to gain insights into the scaling behaviour of their properties in terms of this size. Examples include the mixing time of Davies dynamics [Cap+24] or algorithms for predicting properties by modelling or reconstruction of system characteristics, such as the reconstruction of Gibbs states [Alh+24]. Note that one could alternatively adopt the perspective of Hamiltonians purely on finite subsets. This latter perspective is used in [Cap+24] bringing along with it some notational caveats and ambiguities, that we want to mitigate here by reinterpreting those results within the interaction framework.
We begin by introducing the notation. Let $\Lambda \Subset \mathbb{Z}^D$ denote a finite subset of the lattice $\mathbb{Z}^D$. Its cardinality (i.e., number of elements) is denoted by $|\Lambda|$, and its diameter by $\mathrm{diam}(\Lambda) \equiv \max\{d(x,y) : x, y \in \Lambda\}$, where $d(x,y)$ represents a suitable distance metric on $\mathbb{Z}^D$, in our case the Euclidean one. To each site $z \in \mathbb{Z}^D$, we associate a finite-dimensional Hilbert space $\mathcal{H}_z = \mathbb{C}^d$. From these, we construct the Hilbert space of the finite subset as $\mathcal{H}_\Lambda = \bigotimes_{z\in\Lambda} \mathcal{H}_z$. For any $\Lambda' \subseteq \Lambda \Subset \mathbb{Z}^D$, we can consider an operator $X \in \mathcal{B}(\mathcal{H}_{\Lambda'})$ as an element of $\mathcal{B}(\mathcal{H}_\Lambda)$ by tensoring with the identity operator on the complement, i.e., $X \otimes \mathbb{1}_{\Lambda\setminus\Lambda'} \in \mathcal{B}(\mathcal{H}_\Lambda)$. This identification is often kept implicit but may be indicated with a subscript if necessary (see also section 1.2). One also says that $X_{\Lambda'}$ is supported on Λ'.
Using this embedding, we define the algebra of quasi local observables on $\mathbb{Z}^D$ as

$$\mathcal{B}_{\mathbb{Z}^D} = \overline{\bigcup_{\Lambda\Subset\mathbb{Z}^D} \mathcal{B}(\mathcal{H}_\Lambda)}^{\|\cdot\|_\infty}, \tag{1.45}$$

where the bar denotes the closure with respect to the operator norm $\|\cdot\|_\infty$. Within this setup, we set an interaction to be a map

$$\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}, \qquad \Lambda \mapsto \Phi(\Lambda) \quad \text{with} \quad \Phi(\Lambda) = \Phi(\Lambda)^*, \tag{1.46}$$

where each $\Phi(\Lambda)$ is a self-adjoint operator supported on Λ. The Hamiltonian for any finite subset $\Lambda \Subset \mathbb{Z}^D$ is then defined as an element of $\mathcal{B}(\mathcal{H}_\Lambda)$ by summing the interaction terms supported within Λ:

$$H_\Lambda = \sum_{\Lambda'\subseteq\Lambda} \Phi(\Lambda'). \tag{1.47}$$

Various constraints can be imposed on the interaction Φ. We focus here only on those immediately relevant to the settings in [Cap+24] and [Alh+24], while we will not cover the generalisation to short-range interactions in [Alh+24]. For both setups, we therefore assume the interaction has finite range. This means, firstly, that the interaction strength is uniformly bounded:

$$J \equiv \sup\{\|\Phi(\Lambda)\|_\infty : \Lambda \Subset \mathbb{Z}^D\} < \infty\,,$$

and secondly, that the spatial range of the interaction terms is bounded too:

$$r \equiv \sup\{\mathrm{diam}(\Lambda) : \Phi(\Lambda) \neq 0, \Lambda \Subset \mathbb{Z}^D\} < \infty\,.$$

Let us remark that in [Alh+24] we require a different definition of the interaction strength to be finite, namely $\sup_{z\in\mathbb{Z}^D} \sum_{\Lambda\Subset\mathbb{Z}^D : z\in\Lambda} \|\Phi(\Lambda)\|_\infty$. As discussed in the same paper, both constraints on the interaction strength are equivalent up to constants that depend on r and the lattice dimension D. Since all results in [Alh+24] are stated using Landau notation (see eqs. (1.49)–(1.51)), with all hyperparameters except the system size absorbed into the asymptotic notation, this modification has no impact on the form of the statements.

The parameters r and J, together with the lattice structure, would be sufficient hyperparameters to derive the results in [Cap+24; Alh+24]. Yet, we want to introduce further parameters which can simplify arguments in places and potentially tighten bounds while further improving the potential for reuse of parts of the derivation in future work or analysis. All these additional parameters implicitly depend on D and r, meaning one can derive bounds on them involving only the lattice dimension D and the interaction range r. We begin with the maximum size of interacting sets:

$$\kappa \equiv \sup\{|\Lambda| : \Phi(\Lambda) \neq 0, \Lambda \Subset \mathbb{Z}^D\}\,,$$

and follow with the connectivity, or maximum overlap, of the interaction terms:

$$g \equiv \sup_{z\in\mathbb{Z}^D} |\{\Lambda : \Lambda \Subset \mathbb{Z}^D, \Phi(\Lambda) \neq 0, z \in \Lambda\}|\,.$$

We further introduce three properties of an interaction that do not come in parameter form:

1. We call an interaction Φ commuting if there exists a commuting algebra $\mathcal{A} \subset \mathcal{B}_{\mathbb{Z}^D}$ over $\mathbb{C}$ such that $\{\Phi(\Lambda) : \Lambda \Subset \mathbb{Z}^D\} \subseteq \mathcal{A}$.

2. Furthermore, we say that a commuting interaction Φ is marginal commuting if the commuting algebra $\mathcal{A}$ is also closed under partial traces, meaning that for all $\Lambda' \Subset \mathbb{Z}^D$, $\mathbb{1}_{\Lambda'} \otimes \mathrm{tr}_{\Lambda'}[\mathcal{A}] \subseteq \mathcal{A}$.

3. Finally, we define translation invariance specifically for the one-dimensional systems, that is $\mathbb{Z}$: An interaction $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}\} \to \mathcal{B}_{\mathbb{Z}}$ is called translation-invariant if there exists a translation distance $t \in \mathbb{Z}$ such that for any finite subset $\Lambda \Subset \mathbb{Z}$, the interaction term for the shifted set is the appropriately translated version of the original term. Formally, $\tau_t(\Phi(\Lambda)) = \Phi(t + \Lambda)$, where τ_t represents the automorphism on the quasi-local algebra $\mathcal{B}_{\mathbb{Z}}$ induced by the spatial shift $z \mapsto z + t$.

The marginal commuting property, perhaps the most subtle of these three restrictions, has far-reaching consequences; for example, it implies the existence of a 'strong-effective Hamiltonian' (terminology as used in cited works), which allows for the use of classical tools to prove properties such as the decay of the MI [BCPH24] and the CMI [Cap+24].

Its name becomes clearer when examining the Gibbs state of a local Hamiltonian on $\Lambda \Subset \mathbb{Z}^D$ at inverse temperature $\beta \in \mathbb{R}_+$:

$$\sigma^\Lambda \equiv \frac{1}{\mathrm{Tr}[e^{-\beta H_\Lambda}]} e^{-\beta H_\Lambda} \in \mathcal{B}(\mathcal{H}_\Lambda) \tag{1.48}$$

and its embedded marginal on $A \subset \Lambda$, given by $\sigma^\Lambda_A = \mathbb{1}_A \otimes \mathrm{tr}_A[\sigma^\Lambda] \in \mathcal{B}(\mathcal{H}_\Lambda)$. Since σ^Λ, as the rescaled exponential of the local Hamiltonian H_Λ lies in the closure of $\mathcal{A}$, its marginal σ^Λ_A must also belong to this closure. This follows from the fact that both the partial trace and tensoring with the identity are continuous operations under which $\mathcal{A}$ is closed by the marginal commuting property. Consequently, it follows that

$[\sigma_A^\Lambda, \sigma^\Lambda] = 0$, that is, the Gibbs state commutes with its marginal.
Note that in the definition of the Gibbs state (eq. (1.48)) and of its marginal, we usually use the shorthand notation σ and σ_A when the relevant region $\Lambda \Subset \mathbb{Z}^D$ is clear from the context.
One example of such marginal commuting interactions arises in Calderbank-Shor-Steane (CSS)-codes, or more generally, commuting interactions where all interaction terms $\Phi(\Lambda)$ are Pauli strings. A proof that marginals of Gibbs states of Hamiltonians associated with these CSS-codes commute was previously discovered by Sebastian Stengel. Here, however, we present an alternative proof that not only recovers this result but also establishes the stronger marginal commuting property.

Example 1. Assume $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}$ to be a commuting interaction that is composed of Pauli strings, meaning for all $\Lambda \Subset \mathbb{Z}^D$ where $\Phi(\Lambda) \neq 0$, it holds that

$$\Phi(\Lambda) = \lambda \bigotimes_{z\in\Lambda} P_z, \quad \lambda \in \mathbb{R},\ P_z \in \{\mathbb{1}, X, Y, Z\} \subset \mathcal{B}(\mathbb{C}^2)\,,$$

where $\{\mathbb{1}, X, Y, Z\}$ are the identity and Pauli matrices. Then Φ further is marginal commuting with the marginal commuting algebra $\mathcal{A}$ generated by all interaction terms $\{\Phi(\Lambda) : \Lambda \Subset \mathbb{Z}^D\}$.

Proof. By assumption, Φ is commuting, meaning the algebra $\mathcal{A}$ generated by the terms $\{\Phi(\Lambda) : \Lambda \Subset \mathbb{Z}^D\}$ is commuting. It remains to show that $\mathcal{A}$ satisfies $\mathbb{1}_{\Lambda'} \otimes \mathrm{tr}_{\Lambda'}[\mathcal{A}] \subseteq \mathcal{A}$ for all $\Lambda' \Subset \mathbb{Z}^D$. Note that the operation $\mathbb{1}_{\Lambda'} \otimes \mathrm{tr}_{\Lambda'}$ can be expressed as a composition of commuting single-site operations: $\bigcirc_{z\in\Lambda'}(\mathbb{1}_z \otimes \mathrm{tr}_z)$. Therefore, it suffices to demonstrate closure under $\mathbb{1}_z \otimes \mathrm{tr}_z$ for an arbitrary single site $z \in \mathbb{Z}^D$. Any element in $\mathcal{A}$ is a linear combination of finite products (monomials) of the interaction terms $\Phi(\Lambda)$. Since $\mathbb{1}_z \otimes \mathrm{tr}_z$ is a linear map, if it maps monomials of interaction terms back into $\mathcal{A}$, it will map any polynomial (that is, any element of $\mathcal{A}$) back into $\mathcal{A}$. The interaction terms are Pauli strings. Products of Pauli matrices are Pauli matrices (up to a phase factor $\pm 1, \pm i$), and thus products of Pauli strings are also Pauli strings. Therefore, any monomial in the interaction terms is a Pauli string P. We examine the action of $\mathbb{1}_z \otimes \mathrm{tr}_z$ on such a monomial P. The result depends on the Pauli operator P_z acting on site z. If P_z is traceless (that is, either X, Y, or Z), then $\mathrm{tr}_z[P_z] = 0$, and consequently $\mathbb{1}_z \otimes \mathrm{tr}_z[P] = 0$. If $P_z = \mathbb{1}$, then $\mathrm{tr}_z[P_z] = 2$, and $\mathbb{1}_z \otimes \mathrm{tr}_z[P]$ yields an operator proportional to P itself (as P already has $\mathbb{1}$ at site z). In both cases, the result (0 or a scalar multiple of an element P already representable within the algebra) is contained in $\mathcal{A}$. Therefore, $\mathbb{1}_z \otimes \mathrm{tr}_z$ maps monomials in $\mathcal{A}$ to elements in $\mathcal{A}$, and by linearity, $\mathbb{1}_z \otimes \mathrm{tr}_z[\mathcal{A}] \subseteq \mathcal{A}$. By composition, $\mathbb{1}_{\Lambda'} \otimes \mathrm{tr}_{\Lambda'}[\mathcal{A}] \subseteq \mathcal{A}$ holds for all $\Lambda' \Subset \mathbb{Z}^D$. □

Finally, we wish to introduce an important concept for dealing with scaling properties in the context of interactions: the Landau notation. This notation is particularly relevant to the discussions in [Alh+24] regarding error scaling of a MPO reconstruction, and in [Cap+24] concerning scaling of mixing times, gaps, MLSI, and CMLSI. The quantities that appear in these contexts can typically be regarded as functions mapping systems (i.e., subsets $\Lambda \Subset \mathbb{Z}^D$) to non-negative real numbers, that is, $g : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathbb{R}_+$, while only being dependent on the interaction and other global constants such as for example the inverse temperature β. We then write,

$$g(\Lambda) = O(f(|\Lambda|)) \tag{1.49}$$

for a function $f : \mathbb{R}_+ \to \mathbb{R}_+$, if there exist constants $c_1 > 0, c_2 \geq 0$ such that for all $\Lambda \Subset \mathbb{Z}^D$:

$$g(\Lambda) \leq c_1 f(|\Lambda|) + c_2\,.$$

These constants (and all subsequent constants in this context, unless specified otherwise) depend only on inverse temperature β, the interaction (characterised by locality parameters J, r, g, κ), the lattice dimension D, and the local dimension d. Analogously, we set

$$g(\Lambda) = \Omega(f(|\Lambda|)) \tag{1.50}$$

if there exists a constant $c > 0$ (with the dependencies specified above) such that for all $\Lambda \Subset \mathbb{Z}^D$:

$$g(\Lambda) \geq c f(|\Lambda|)\,.$$

Finally, we write

$$g(\Lambda) = \Theta(f(|\Lambda|)) \tag{1.51}$$

if there exists constants $c_1, c_2 > 0, c_3 \geq 0$ (with the dependencies specified above) such that for all $\Lambda \Subset \mathbb{Z}^D$:

$$c_1 f(|\Lambda|) \leq g(\Lambda) \leq c_2 f(|\Lambda|) + c_3 \,.$$

Furthermore, we also use notation such as $g = O(\mathrm{poly}(f(|\Lambda|))), \Omega(\mathrm{poly}(f(|\Lambda|)))$ or $\Theta(\mathrm{poly}(f(|\Lambda|)))$. This implies the existence of a polynomial $p : \mathbb{R}_+ \to \mathbb{R}_+$, whose coefficients and degree depend only on the parameters specified above, such that $g = O(p \circ f(|\Lambda|)), \Omega(p \circ f(|\Lambda|))$ or $\Theta(p \circ f(|\Lambda|))$ respectively.
This concludes our introductory discussion of Hamiltonians on lattices and further the general preliminaries. We will now proceed to discuss the background of specific projects, starting with a section on continuity bounds.

1.4 Project specific preliminaries

1.4.1 Continuity bounds for some entropy functionals ([Aud+25] and [Blu+24])

When dealing with entropy functionals of any kind, a central challenge in both theoretical and experimental analysis lies in the imperfections in the preparation, control, and measurement of quantum states. These imperfections naturally lead to the question of the stability of the functional under such perturbations. Qualitative knowledge in the form of continuity can be complemented by quantitative understanding via continuity bounds, which enable both approximability and theoretical analysis of such quantities (for example, in the characterisation of the stability of approximate quantum Markov chains [Sut18]).
Continuity bounds most commonly take the form of estimates involving the trace-distance (TD), which we recall as $T(\rho, \rho') = \frac{1}{2}\|\rho - \rho'\|_1$, owing to its operational significance. These bounds are then expressed as

$$|f(\rho) - f(\rho')| \leq g(T(\rho, \rho')) \,,$$

where $\rho, \rho' \in \mathcal{S}(\mathcal{H})$ (or a restricted subset of states), and $g : [0,1] \to [0, \infty)$ is a continuous function vanishing at zero, independent of the specific input states (subject to the aforementioned constraint on the set).
One of the earliest and most prominent examples of such a bound in quantum information theory is due to Fannes [Fan73], who established a continuity bound for the von Neumann entropy, later build upon by him and Alicki to extended to a bound on the CE in [AF04]. Subsequently, Audenaert [Aud07] improved the original von Neumann entropy bound to its yet tightest form by deriving the sharp inequality

$$|S(\rho) - S(\rho')| \leq \varepsilon \log(d-1) + h(\varepsilon) \tag{1.52}$$

with $\varepsilon = T(\rho, \rho')$, and $h(x) \equiv -x \log x - (1-x)\log(1-x)$ the binary entropy. The term 'sharp' indicates that the bound is saturated by certain states.
This result was later rederived within the framework of majorisation theory, and more specifically using majorisation flow, in [HD22] which also enabled a broader treatment, deriving continuity bounds for general Schur-concave and Schur-convex functions in the eigenvalues of the involved states.
Similarly, the Alicki-Fannes bound for the CE was significantly refined, culminating in Winter's work [Win16], where he derived an almost sharp continuity bound:

$$|S(A|B)_\rho - S(A|B)_{\rho'}| \leq \varepsilon \log d_A^2 + h\Big(\frac{\varepsilon}{1+\varepsilon}\Big) \,, \tag{1.53}$$

for $\rho, \rho' \in \mathcal{S}(\mathcal{H}_{AB})$ with $\varepsilon = T(\rho, \rho')$. This bound is almost sharp, due to its correct scaling with respect to all relevant parameters, yet falling short of the conjectured optimal bound (see [Wil20, Eq. (58)]):

$$|S(A|B)_\rho - S(A|B)_{\rho'}| \overset{?}{\leq} \varepsilon \log\big(d_A^2 - 1\big) + h(\varepsilon) \,. \tag{1.54}$$

The proof strategies leading to eq. (1.53) were unified by Shirokov through the so-called Alicki-Fannes-Winter (AFW)-method [Shi20; Shi23], which placed the treatment of continuity bounds for the von Neumann

entropy, CE, MI, and CMI on a common footing and included extensions to the infinite-dimensional setting [Shi20]. The method has since found further applications [SW23] and has been extended in [Blu+23a; Blu+23b]. Unlike the approach based on majorisation theory [HD22] which produces sharp bounds for functionals dependent only on the eigenvalues of the involved states, the AFW method, while often not sharp, applies to both basis-dependent entropy functionals and those depending solely on the eigenvalues of the input state.
The bound for the CE (eq. (1.53)) has also been extended to the setting of SR-CE, that is, continuity bounds for the mapping

$$\widetilde{S}^{\uparrow}_{\alpha}(A|B)_{\rho} = - \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \widetilde{D}_{\alpha}(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B) , \quad \rho \in \mathcal{S}(\mathcal{H}_{AB}) . \tag{1.55}$$

These functionals generalises the CE using the SR divergences (see eq. (1.13)) instead of the relative entropy in its variational formulation eq. (1.17). As with the SR divergence and the relative entropy, the SR-CE converges to the CE in the limit $\alpha \to 1$ (see, e.g., [Blu+24]).
In [MD22], Marwah and Dupuis proved that for $\rho, \rho' \in \mathcal{S}(\mathcal{H}_{AB})$ and $\varepsilon = T(\rho, \rho')$,

$$\begin{aligned} &|\widetilde{S}^{\uparrow}_{\alpha}(A|B)_{\rho} - \widetilde{S}^{\uparrow}_{\alpha}(A|B)_{\rho'}| \\ &\quad \leq \begin{cases} \log(1+\varepsilon) + \frac{1}{1-\alpha} \log\Big(1 + \varepsilon^{\alpha} d_A^{2(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\Big) & \alpha \in [1/2, 1) , \\ \log\big(1+\sqrt{2\varepsilon}\big) + \frac{\alpha}{1-\alpha} \log\Big(1 + (2\varepsilon)^{\frac{2\alpha-1}{2\alpha}} d_A^{\frac{2(1-\alpha)}{\alpha}} - \frac{\sqrt{2\varepsilon}}{(1+\sqrt{2\varepsilon})^{\frac{1-\alpha}{\alpha}}}\Big) & \alpha \in (1, \infty) . \end{cases} \end{aligned} \tag{1.56}$$

Using a different technique based on rewriting $\widetilde{S}^{\uparrow}_{\alpha}(A|B)_{\rho}$ in terms of a norm, Beigi and Goodarzi [BG23] derived a bound for $\alpha > 1$:

$$|\widetilde{S}^{\uparrow}_{\alpha}(A|B)_{\rho} - \widetilde{S}^{\uparrow}_{\alpha}(A|B)_{\rho'}| \leq \frac{\alpha}{\alpha - 1} \log\Big(1 + 2\varepsilon d_A^{\frac{2(\alpha-1)}{\alpha}}\Big) , \tag{1.57}$$

which exhibits better scaling for large α compared to eq. (1.56). However, unlike eq. (1.56), it does not converge to eq. (1.53) as $\alpha \to 1$, but instead diverges, giving superiority of eq. (1.56) for α close to one.
Except for Fannes' original proof [Fan73] and the proofs relying on majorisation theory, all the aforementioned strategies share a common tool: the JH-decomposition. It decomposes the difference between two states $\rho, \rho' \in \mathcal{S}(\mathcal{H})$ into a rescaled difference of two other states $\mu, \nu \in \mathcal{S}(\mathcal{H})$:

$$\rho - \rho' = \varepsilon(\mu - \nu) . \tag{1.58}$$

The scaling parameter is $\varepsilon = T(\rho, \rho')$, and μ and ν are given, respectively, by $\mu = \varepsilon^{-1}(\rho - \rho')_{+}$ and $\nu = \varepsilon^{-1}(\rho - \rho')_{-}$, where $(\cdot)_{\pm}$ denotes the positive/non-positive part of a self-adjoint operator. The combination of this decomposition with specific properties of the entropy functional is now the key to the majority of results discussed above. It is also used in the project [Aud+25] to improve upon eq. (1.52), with further implications for eq. (1.54). Additionally, our work [Blu+24] features it in a generalisation and strengthening of both eq. (1.56) and eq. (1.57). A more detailed account of the aims of each project can be found in section 2.1, while the corresponding results are presented in section 3.1.

1.4.2 Reconstruction of Gibbs states on spin chains ([Alh+24])

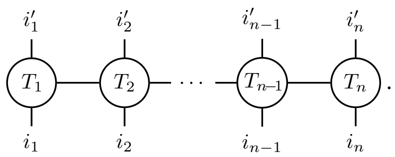
One significant challenge in the simulation, storage, and analysis of quantum theory is the exponential growth in the complexity of a system with the number of subsystems. More precisely, the total dimension of a combined Hilbert space $\bigotimes_{i=1}^{n} \mathbb{C}^d$ is d^n, meaning it grows exponentially with the number of subsystems, n. This 'curse of dimensionality', in particular is relevant in the simulation and storage of the thermal state, that is the Gibbs state (eq. (1.48)) of a system. When considering highly symmetric Hamiltonians, one would intuitively expect there to be a much more compressed representation. For Gibbs states one spin chains (that is interactions on $\mathbb{Z}$), which are further local and translation-invariant, for instance, one should be able to reconstruct from a limited number of smaller building blocks. These building blocks could then be measured and, due to their limited number and size, efficiently stored, giving an approximate representation of the global system's thermal state usable in further theoretical analysis. The process of reconstructing the Gibbs state from smaller building blocks (its marginals to be precise) and combining this with a measurement algorithm is called Gibbs learning and is the topic of [Alh+24].

A very successful tool used not only in the reconstruction of Gibbs states but further in simulation [VGRC04; Che+21], phase classification [SPGC11] and eigenvalue estimates [HV17] of spin chains are MPOs or more specifically when the considered operators are states matrix product density operators (MPDOs). The idea is that a huge tensor—e.g., the coefficient tensor of the Gibbs state in a local product basis, that is, $|i_1, \ldots, i_n\rangle$ with $i_j = 1, \ldots, d$ in our case—can be written as a chain of small tensors. Let us illustrate this on the aforementioned Gibbs state $\sigma \in \mathcal{S}(\bigotimes_{i=1}^n \mathbb{C}^d)$. Writing it in the given product basis, we notice that it can be represented by a tensor with $2n$ free ('physical') indices, each ranging from 1 to d, which form the coefficients in the basis decomposition:

$$\sigma = \sum_{i_1,\ldots,i_n,i'_1,\ldots,i'_n} \sigma_{i_1\ldots i_n}^{i'_1\ldots i'_n} |i_1 \ldots i_n\rangle\langle i'_1 \ldots i'_n| \,. \tag{1.59}$$

For an exact MPDO decomposition, one needs to find n tensors $T_1, \ldots, T_n$, or with explicit indices $(T_1)_{i_1}^{i'_1\alpha_1}$, $(T_2)_{i_2\alpha_1}^{i'_2\alpha_2}, \ldots, (T_n)_{i_n\alpha_{n-1}}^{i'_n}$, where $a_j = 1, \ldots, b_j$ are called 'bond' indices, such that the following equality holds:

$$\sigma_{i_1\ldots i'_n}^{i'_1\ldots i_n} = \sum_{\alpha_1,\ldots,\alpha_n} (T_1)_{i_1}^{i'_1\alpha_1} (T_2)_{i_2\alpha_1}^{i'_2\alpha_2} \ldots (T_n)_{i_n\alpha_{n-1}}^{i'_n} = \tag{1.60}$$



The right-hand-side shows the often preferred graphical notation, which also reveals the chain-like nature of a MPO or MPDO, respectively. The dangling legs represent the physical indices, while the connecting lines represent the bond indices, which are contracted (summed over). The maximum dimension of the latter, that is, $b = \max\{b_j : j = 1, \ldots, n-1\}$, is called the bond dimension and is an important parameter, as it gives an upper bound on the number of contractions to perform and the memory requirement to store the tensor. Clearly every state can be written as a MPDO while efficiency (one can conclude an advantage in storage, where naively d^{2n} numbers would need to be stored compared to a MPO with bond-dimension b where at most $n(db)^2$ numbers are necessary) is only guaranteed if b is smaller than exponential in n (what we call system size).

To achieve storage efficiency, we restrict the bond dimension, trading its size for reconstruction accuracy, and no longer require the reconstruction to be a valid quantum state, but only that it remains self-adjoint and positive semidefinite; such objects are referred to as MPOs[5]. The main challenge is now to find a reconstruction that offers a scalable and favourable trade-off between reconstruction accuracy and bond dimension size.

Previous work [KAA21] tackling this problem has relied on explicit knowledge of the Hamiltonian (that is, knowledge of the interactions) and found efficient representations (meaning a bond dimension that is sub-polynomial in the quotient of system size and reconstruction error). The problem with this approach is that reconstructing the Gibbs state from measurements would require structural knowledge of the interaction and Hamiltonian learning, which suffers from impractically large sample complexities for polynomial time algorithms [Ans+21; Bak+24; HKT24]. Furthermore, it is unclear how measurement errors propagate through the previously mentioned reconstructions.

Alternatively, one could learn the representation directly, applying the proposed algorithms from [Fan+23; Qin+24] for finitely correlated states (often seen as equivalent to MPO). However, this algorithm incurs an approximation error that depends on a bound for the singular values of the reconstruction map. Since such a bound is currently unknown, no conclusions about its efficiency can be drawn.

Arguably the most canonical path is attempting to reconstruct the Gibbs state using its marginals, which are

[5]In the literature, MPOs are not necessarily required to be self-adjoint and positive-semidefinite, but these properties have significant advantages, not only when trying to implement algorithms numerically but also in the proofs of efficient reconstruction, as we will see later (see section 3.2).

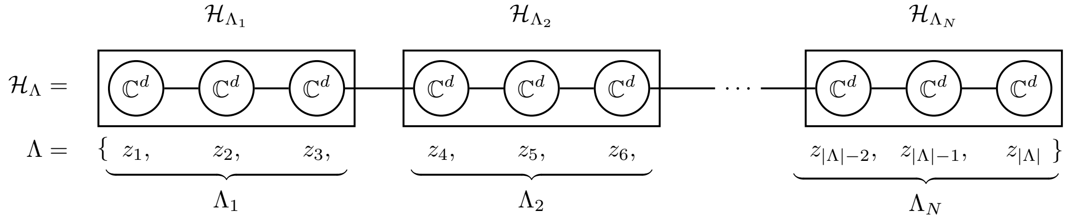


Figure 1.1: A schematic decomposition of the spin chain into disjoint, adjacent intervals of equal size.

easily accessible through measurement. This approach also has the advantage that, in doing so, one enters the well-studied realm of recovery maps [Pet88; Jun+18; SBT16; Sut18] and strengthened DPI [BLW15; Jun+18; SBT16; CV18] with much research and tools to rely on.

As the name suggests, recovery maps (approximately) recover a global state from its marginals. We begin with the toy example of a tripartite system $\mathcal{H}_{ABC}$, where each subsystem A, B, or C might itself be composed of multiple smaller systems (one should imagine a tripartition of a finite spin chain). A recovery map for $\sigma_{ABC} \in \mathcal{S}(\mathcal{H}_{ABC})$ is then a linear CP map $\mathcal{R}^\sigma_{B\to BC} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{BC})$ such that for the marginal σ_B

$$\mathcal{R}^\sigma_{B\to BC}(\sigma_B) = \sigma_{BC} \,. \tag{1.61}$$

That is, acting on the marginal of the middle subsystem, the recovery map exactly reconstructs the joint state of the middle and right subsystems. We can lift this map to $\mathrm{id}_A \otimes \mathcal{R}^\sigma_{B\to BC} : \mathcal{B}(\mathcal{H}_{AB}) \to \mathcal{B}(\mathcal{H}_{ABC})$ by tensoring with the identity map on system A. Ideally, this lifted map recovers σ_{ABC} exactly from σ_{AB}, but even an approximate recovery with a controllable error would suffice, that is

$$(\mathrm{id}_A \otimes \mathcal{R}^\sigma_{B\to BC})(\sigma_{AB}) \approx^\varepsilon \sigma_{ABC} \,.$$

Here and in the following, $\approx^\varepsilon$ means approximate equality up to ε in TD, that is

$$T(\sigma_{ABC}, (\mathrm{id}_A \otimes \mathcal{R}^\sigma_{B\to BC})(\sigma_{AB})) = \varepsilon \,. \tag{1.62}$$

With this intuitive understanding alone, one already can then devise an algorithm to reconstruct a state $\sigma_\Lambda \in \mathcal{S}(\mathcal{H}_\Lambda)$ (not necessarily a Gibbs state) on a finite spin chain $\Lambda \Subset \mathbb{Z}$ as follows:

1. Subdivide the chain into a disjoint union of adjacent sets (intervals) $\Lambda = \bigcup_{n=1}^N \Lambda_n$ of equal size, as in figure 1.1. The total Hilbert space is decomposed accordingly: $\mathcal{H}_\Lambda = \bigotimes_{n=1}^N \mathcal{H}_{\Lambda_n}$ where $\mathcal{H}_{\Lambda_n} = \mathbb{C}^{d^{|\Lambda_n|}}$.

2. Then use corresponding recovery maps to reconstruct from left to right. Beginning with the state $\sigma_{12} \equiv \sigma_{\Lambda_1\Lambda_2}$ (the first step is an exact recovery, i.e., $\mathcal{R}_1(\sigma_1) = \sigma_{12}$, which is why we skip this step from now on and start with σ_{12}) and reconstruct an approximation of the state on $\mathcal{H}_{1:3} \equiv \mathcal{H}_{\Lambda_1\Lambda_2\Lambda_3}$ using the recovery map $\mathcal{R}_2 \equiv \mathrm{id}_1 \otimes \mathcal{R}^\sigma_{2\to 1:3}$, which maps $\mathcal{B}(\mathcal{H}_{2:3}) \to \mathcal{B}(\mathcal{H}_{1:3})$. Note our shorthand notation $n : m \equiv \Lambda_n \dots \Lambda_m$. For the general n-th step, the recovery map is $\mathcal{R}_n \equiv \mathrm{id}_{1:n-1} \otimes \mathcal{R}^\sigma_{n\to n:n+1}$, mapping $\mathcal{B}(\mathcal{H}_{1:n}) \to \mathcal{B}(\mathcal{H}_{1:n+1})$. This gives the full reconstruction as an ordered composition (denoted by $\bigcirc$) of recovery steps, starting from the initial marginal σ_1:

$$\left(\bigcirc_{n=1}^{N-1} \mathcal{R}_n\right)(\sigma_1) = \left(\bigcirc_{n=2}^{N-1} \mathcal{R}_n\right)(\sigma_{12}) \approx^\delta \sigma_\Lambda \,. \tag{1.63}$$

Under the assumption that the composed maps $\bigcirc_{n=i}^{N-1} \mathcal{R}_n$ satisfy a uniform bound with respect to the TD, that is, there exists a constant c such that

$$T(\bigcirc_{n=i}^{N-1} \mathcal{R}_n(X), \bigcirc_{n=i}^{N-1} \mathcal{R}_n(Y)) \le c\, T(X,Y) \quad \forall X,Y \in \mathcal{B}(\mathcal{H}_{1:n}), X,Y \ge 0 \;\forall i = 2,\dots,N-1\,, \tag{1.64}$$

one employs a telescopic sum:

$$\delta \le \sum_{i=3}^N T(\bigcirc_{n=i}^{N-1} \mathcal{R}_n(\sigma_{1:i}), \bigcirc_{n=i-1}^{N-1} \mathcal{R}_n(\sigma_{1:i-1})) \le c \sum_{i=3}^N T(\sigma_{1:i}, \mathcal{R}_{i-1}(\sigma_{1:i-1}))\,, \tag{1.65}$$

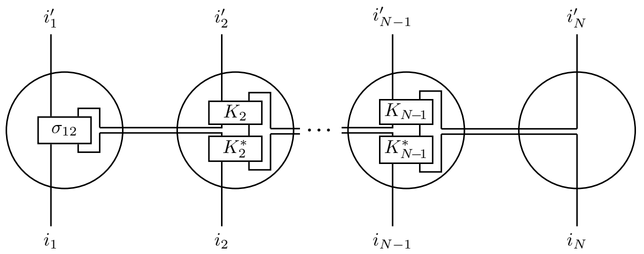
and thereby can get an estimate on the total error δ.

The above represents the foundational idea from [Alh+24] for the reconstruction of states (specifically, Gibbs states of local, translation-invariant Hamiltonians) on spin chains. Similar ideas have previously appeared, for instance in [BK18], in the context of quantum state preparation and using different recovery maps, though not in connection with MPO reconstructions for Gibbs learning.
The most important question remains: Which recovery map should one choose? And how to control the error of a single reconstruction step, that is $T(\sigma_{1:i}, \mathcal{R}_{i-1}(\rho_{1:i-1}))$ for that map.
One natural candidate is the Petz recovery [Pet88]:

$$\widetilde{\mathcal{R}}^{\sigma}_{B\to BC} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{BC}), \quad \widetilde{\mathcal{R}}^{\sigma}_{B\to BC}(X) \equiv \sigma_{BC}^{1/2}((\sigma_B^{-1/2} X \sigma_B^{-1/2}) \otimes \mathbb{1}_C)\sigma_{BC}^{1/2} . \tag{1.66}$$

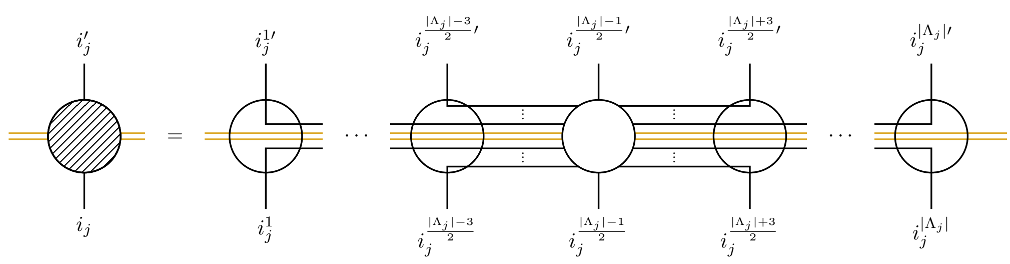
An advantage of this map is that it is CPTP, which ensures that the overall reconstruction map in eq. (1.63) is also CPTP, yielding the bound in eq. (1.64) with $c = 1$. This further implies that the reconstruction not only gives a MPO (as becomes clear from the following graphical decomposition):

$$(\bigcirc_{n=2}^{N} \widetilde{\mathcal{R}}_n(\sigma_{12}))^{i'_1 \dots i'_N}_{i_1 \dots i_N} =$$



(1.67)

with $K_n = \sigma_{n:n+1}^{1/2}(\sigma_n^{-1/2} \otimes \mathbb{1}_{n+1})$, but a MPDO. What remains is to estimate $T(\sigma_{1:i}, \widetilde{\mathcal{R}}_{i-1}(\sigma_{1:i-1}))$, a task we postpone briefly in order to comment on the bond dimension of the reconstruction described above, relative to the original system. This comparison is slightly skewed due to the regrouping carried out in the initial step, namely the decomposition $\mathcal{H}_\Lambda = \bigotimes_{n=1}^{N} \mathcal{H}_{\Lambda_n}$. As a result, the bond dimension b cannot be directly inferred from the regrouped bond dimension b_r, which only provides an upper bound on the number of contractions between pairs of tensors, as represented by the horizontal lines in eq. (1.67). However, we can conclude that the bond dimension, in relation to the original system, is upper-bounded by the product of the maximal physical dimension after regrouping $p_r = \max\{d^{|\Lambda_n|} : n = 1, \dots, N\}$ and the bond dimension after regrouping $b_r = \max\{d^{2|\Lambda_n|} : n = 1, \dots, N\}$, i.e., $b_r \cdot p_r$. This becomes clear when examining the constituents (circles in the diagram) of eq. (1.67) and decomposing each of them individually into a MPO with respect to the original systems. For these constituents, we cannot assume any further internal structure, so the following becomes optimal in terms of minimal bond dimension:



(1.68)

where $j = 1, \dots, N$, and $i'_j = (i_j^{1\prime}, \dots, i_j^{|\Lambda_j|\prime})$ and $i_j = (i_j^1, \dots, i_j^{|\Lambda_j|})$ are index tuples referring to the original chain, with $i_j^k, i_j^{k\prime} \in \{1, \dots, d\}$. In eq. (1.68), we have assumed that $|\Lambda_j|$ is odd; the case where $|\Lambda_j|$ is even leads to an asymmetric variant of the diagram. This is obtained by shortening the chain by one site and shifting the central circle either left or right. Importantly, this modification does not alter the bond dimension.
The resulting bond dimension can now be read off as local dimension d raised to the power of the maximum number of horizontal black lines entering or exiting the central circle (since all other components have fewer

such lines), multiplied by the bond dimension of the regrouping b_r (represented by the orange lines):

$$b = \max\{b_r \cdot d^{2\left\lfloor\frac{|\Lambda_j|-1}{2}\right\rfloor} : j = 1, \dots N\} \leq \max\{b_r \cdot d^{|\Lambda_j|} : j = 1, \dots N\} = b_r \cdot p_r \,.$$

After addressing this technicality, we now return to the Petz recovery map as a primitive for MPDO reconstruction.

One disadvantage of the Petz recovery is the difficulty obtaining a tight error estimate for individual recovery steps based on desired information-theoretic quantities. A ubiquitous [Sut18; BK18; Rou24] strategy would be to upper bound the recovery error, $T(\sigma_{1:i}, \widetilde{\mathcal{R}}_{i-1}(\sigma_{1:i-1}))$ in terms of the CMI. Particularly for Gibbs states of local, translation-invariant interactions[6] on a spin chain, the CMI has been shown to decay exponentially with the size of the separating system at every positive temperature [Kuw24]: There exist constants c_1, c_2, dependent only on the interaction strength J and range r, such that for the Gibbs state $\sigma^\Lambda \in \mathcal{S}(\mathcal{H}_\Lambda)$ on $\Lambda \Subset \mathbb{Z}$ at $\beta \in \mathbb{R}_+$ and any decomposition $\Lambda = A \sqcup B \sqcup C$,

$$I(A : C|B)_\sigma \leq e^{-c_1\,\mathrm{dist}(A,C)/\beta + c_2\beta\log(\beta\,\mathrm{dist}(A,C))} \,. \tag{1.69}$$

Note that the above further holds for $\Lambda = D_1 \sqcup A \sqcup B \sqcup C \sqcup D_2$, given that $\mathrm{dist}(D_1 \sqcup A, C \sqcup D_2) = \mathrm{dist}(A, C)$, i.e.,

$$I(A : C|B)_\sigma \leq I(AD_1 : CD_2|B)_\sigma \leq e^{-c_1\,\mathrm{dist}(A,C)/\beta + c_2\beta\log(\beta\,\mathrm{dist}(A,C))}$$

by the DPI for the relative entropy. Such a result is necessary in the context of our proposed reconstruction, as the error of a single recovery step is bounded by a CMI on a subsystem. The problem, however, is that despite intensive effort [BLW15; SBT16; Jun+18], it has not been possible to date to show an upper bound of the form

$$f(T(\mathrm{id}_A \otimes \widetilde{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB}), \sigma_{ABC})) \leq I(A : C|B)_\sigma$$

where $f : [0, 1] \to \mathbb{R}$ is a function vanishing continuously at zero and, crucially, is independent of the eigenvalues of the involved states. The independence is necessary, as such dependencies for Gibbs states of local interactions scale exponentially in system size (Λ), potentially outweighing the decay one gets from eq. (1.69).

One can sidestep this problem by using the rotated Petz recovery map, instead of the standard one, which we denote, overwriting the previous notation, as

$$\widetilde{\mathcal{R}}^\sigma_{B\to BC} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{BC})\,, \quad \widetilde{\mathcal{R}}^\sigma_{B\to BC}(X) \equiv \int_{\mathbb{R}} \beta_0(t)\sigma_{BC}^{\frac{1+it}{2}}((\sigma_B^{-\frac{1+it}{2}} X \sigma_B^{-\frac{1-it}{2}}) \otimes \mathbb{1}_C)\sigma_{BC}^{\frac{1-it}{2}}\, dt \,. \tag{1.70}$$

For this map, it was shown [Jun+18] that

$$(T(\mathrm{id}_A \otimes \widetilde{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB}), \sigma_{ABC}))^{1/2} \leq \widetilde{D}_{1/2}(\mathrm{id}_A \otimes \widetilde{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB}), \sigma_{ABC}) \leq I(A : C|B)_\sigma \,, \tag{1.71}$$

where the first inequality stems from Fuchs-van de Graaf [FG99] relating trace distance and fidelity combined with a standard inequality for the logarithm relating fidelity and the Rényi divergence $\widetilde{D}_{1/2}$. This means that the rotated Petz recovery yields a valid recovery map that, through the upper bound on the CMI and its decay for the specific interaction considered, even provides an estimate for the single-step recovery error. This resolves the issue of eigenvalue dependence and allows one to proceed in proving bounds using the CMI decay. However, at the time of writing of [Alh+24], only a sub-exponential decay of the CMI with the distance of A and C was rigorously known [KB19], whereas exponential decay, that is, eq. (1.69), was only expected (and later proven [Kuw24]). Another quantity, namely the BS-CMI, showed greater potential, despite lacking some relevant inequalities at the time. In [BC19], for example, it was shown that

$$\begin{aligned} \left(\frac{\pi}{4}\right)^4 \|Z^{ABC}\|_\infty^{-2} \left\|\sigma_{AB}^{1/2}\sigma_B^{-1/2}(Z^{BC})^{1/2}\sigma_B^{1/2} - (Z^{ABC})^{1/2}\sigma_{BC}^{1/2}\right\|_2^4 \\ \leq \widehat{D}(\sigma_{AB} \otimes \iota_C \|\sigma_{ABC}) - \widehat{D}(\sigma_B \otimes \iota_C\|\sigma_{BC}) \equiv \widehat{I}(A; C|B)_\sigma \end{aligned} \tag{1.72}$$

[6]The paper [Kuw24] indeed shows the result for exponentially decaying interactions, but since we did not introduce them here we only state it in this specific form.

with $Z^{ABC} = \sigma_{AB}^{-1/2}\sigma_{ABC}\sigma_{AB}^{-1/2}$ and $Z^{BC} = \sigma_B^{-1/2}\sigma_{BC}\sigma_B^{-1/2}$. Note that we exchange the roles of A and C here compared to [Alh+24] and refer to $\widehat{I}(A;C|B)_\sigma$ simply as the BS-CMI (rather than 'reversed') to maintain analogy with the CMI and align with the left-to-right recovery notion discussed. Although the lower bound in eq. (1.72) does not explicitly involve a recovery map, it has the potential to be reformulated into a meaningful bound on the single-step recovery error of a candidate recovery map competing with the rotated Petz recovery.
This result was further complemented in [BCPH22], where the authors established a super-exponential decay of $\left\|\sigma_{ABC}\sigma_{AB}^{-1}\sigma_B\sigma_{BC}^{-1} - \mathbb{1}\right\|_\infty$. Specifically, they proved that for a local, translation-invariant interaction and inverse temperature $\beta \in \mathbb{R}_+$ there exists $c = \Theta(1)$ such that for any finite interval $\Lambda \Subset \mathbb{Z}$, any decomposition $\Lambda = A \sqcup B \sqcup C$ into adjacent subintervals, with corresponding Gibbs state $\sigma \in \mathcal{S}(\mathcal{H}_\Lambda)$, one has

$$\left\|\sigma_{ABC}\sigma_{AB}^{-1}\sigma_B\sigma_{BC}^{-1} - \mathbb{1}\right\|_\infty \leq \frac{c}{\left(\left\lfloor \frac{\lfloor |B|/2 \rfloor}{r} \right\rfloor + 1\right)!} . \tag{1.73}$$

Here r is the range of the interaction, $(\,\cdot\,)!$ denotes the factorial, and $\lfloor\,\cdot\,\rfloor$ is the floor function, that is the greatest integer less than or equal to the argument. In the same paper, the authors conjectured this expression to be an upper bound on the BS-CMI due to its resemblance to the recovery condition (note the different phrasing, we are not talking about a recovery map here) of the BS-relative entropy [BCPH22]. If one could close this major gap, further relate the quantity bounded in eq. (1.72) to the error of a recovery map that preserves hermiticity and positivity (that is a map leading to a MPO reconstruction), and address secondary issues like the uniform bound for the chain of maps in eq. (1.64), then there would be hope for a better co-dependence between the approximation error and the necessary bond dimension, driven by the super-exponential decay shown in eq. (1.73). We emphasise again that even with the subsequent development (that is, the proof of exponential decay of CMI [Kuw24]), the result of [Alh+24] detailed in section 3.2 arguably leads to a stronger result than the potential path through the standard CMI and rotated Petz map, due to the faster (super-exponential) decay rate of a single-step recovery error. A summary and clarification of the objectives of the project can be found in section 2.2, while the results and further details—including those concerning reconstruction from measurements—are presented in section 3.2.

1.4.3 Convergence analysis for the Davies semigroup ([Cap+24])

The Davies semigroup, introduced superficially already in section 1.3.3, describes a semigroup approximation to the reduced dynamics of a system S weakly coupled to a continuous-variable heat bath B. The total evolution is according to a Hamiltonian of the form:

$$H_S + \varepsilon H_{SB} + H_B \,.$$

To then obtain the weak coupling approximation in the form of a semigroup on the systems, it is assumed that the initial state is $\rho_S \otimes \rho_B$, where ρ_B is the thermal state of the bath at inverse temperature β. An additional assumption is that the bath throughout the evolution remains in its thermal state. Rescaling time as $\tau = \varepsilon^2 t$ and taking the limit $\varepsilon \to 0$ then gives rise to what we today call Davies dynamics, named after E.B. Davies who rigorously derived what we heuristically described here in the 1970s [Dav74].
Alternatively, from a quantum information theoretic perspective, the Davies dynamics can be viewed as a method for implementing an efficient Gibbs sampler on a quantum computer, a concept first proposed in [KB16]. Finally, in connecting quantum information theory and physics, Davies dynamics also provide an error model for quantum error correction codes (QECCs), with particular relevance for CSSs codes. In this context, it is assumed that a memory state (that is a state intended for storage) is subjected to thermal error following Davies dynamics, with the system Hamiltonian defined as the sum of all parity-checks.
A crucial question, pertinent to the physics interpretation and central to both Gibbs sampling and the thermalisation of QECCs, is the speed of convergence, that is the mixing time (eq. (1.33)). A reformulation of this question with the different flavours of perspective becomes: How quickly does the algorithm converge, or how rapidly does the system thermalise, respectively? Specifically is it feasible to use the algorithm or is there sufficient time for active error correction before significant amounts of information are lost? This 'speed', or mixing time (eq. (1.33)), is, in this context, measured by its dependence on the system size. Three regimes are distinguished, particularly for the Davies-Lindbladian $\mathcal{L}_\Lambda : \mathcal{B}(\mathcal{H}_\Lambda) \to \mathcal{B}(\mathcal{H}_\Lambda)$ with $\Lambda \Subset \mathbb{Z}^D$ (see also section 1.3.4 for notation):

1. slow-mixing: $t_{\text{mix}}(\mathcal{L}_\Lambda; 1/2) = O(\text{poly}(\exp(|\Lambda|)))$
2. fast-mixing: $t_{\text{mix}}(\mathcal{L}_\Lambda; 1/2) = O(\text{poly}(|\Lambda|))$
3. rapid-mixing: $t_{\text{mix}}(\mathcal{L}_\Lambda; 1/2) = O(\text{poly}(\log(|\Lambda|)))$

Driven by the above applications, the analysis of the mixing time of the Davies dynamics of local commuting interactions is the primary motivation for [Cap+24]. The final goal is to reduce the question of rapid mixing to a suitable uniform decay of an explicit measure of correlation in the Gibbs state, which for example for high temperature can be shown to exhibit this decay. Although rooted in applications, we will analyse the mixing time from a sole mathematical perspective and will begin with a formal introduction to the Davies dynamics, before diving into previous work on the topic.

In our setting, we restrict ourselves to the case where the underlying interaction used to define the Davies dynamics is local and commuting (see section 1.3.4). That is, for a local commuting interaction $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}$, the Davies generator on $\Lambda \Subset \mathbb{Z}^D$ in the Schrödinger picture is defined as

$$\mathcal{L}_\Lambda(\rho) \equiv \sum_{z\in\Lambda} \mathcal{L}_k(\rho), \quad \text{with} \quad \mathcal{L}_z(\rho) \equiv \sum_{\omega\in\text{Bohr}(H_{z\partial})} \sum_{\alpha=1}^{d^2} \chi^{\beta,\omega} \left(L^\omega_{z,\alpha} \rho L^{\omega *}_{z,\alpha} - \frac{1}{2}\{L^{\omega *}_{z,\alpha} L^\omega_{z,\alpha}, \rho\}\right), \tag{1.74}$$

where we deliberately excluded the additional term $i[H_\Lambda, \cdot]$ from the generator, as it commutes with $\mathcal{L}_\Lambda$ and thus has no effect on the convergence analysis, yet adds technicality. By mentioning this excluded term, we already anticipate the Hamiltonian corresponding to the dynamics, namely the local Hamiltonian $H_\Lambda \in \mathcal{B}(\mathcal{H}_\Lambda)$ (see eq. (1.47)), derived from the interaction. The state stabilised by the dynamics is naturally its Gibbs state σ^Λ at inverse temperature β (see eq. (1.48)), that we, for ease of notion, often just denote by σ for all discussions of the Davies semigroup. With this settled, we can define $L^\omega_{z,\alpha}$. One begins by choosing a self-adjoint HS-basis $\{L_\alpha\}_{\alpha=1}^{d^2}$ of the local Hilbert space $\mathcal{B}(\mathbb{C}^d)$ (e.g., for $d = 2$, the Pauli matrices, or for general d, the generalised Gell-Mann matrices). Then, for every $z \in \Lambda$, one obtains $L_{z,\alpha} \equiv L_\alpha \otimes \mathbb{1}_{\Lambda\setminus\{z\}}$. The requirement of a HS-basis is to ensure that for the trivial interaction $\Phi = 0$ the Davies semigroup is depolarising ($\mathcal{L}_z = (\text{tr}_z - \text{id})$), while for a non-trivial interaction $\Phi \neq 0$ at $\beta = 0$, the semigroup still remains related to the depolarising one. Now the $L^\omega_{z,\alpha}$ are the 'Fourier coefficients' of the evolution $e^{-itH_\Lambda} L_{z,\alpha} e^{itH_\Lambda}$. Since we are dealing with a commuting interaction and $L_{z,\alpha}$ is only supported on z, many terms cancel, and one obtains:

$$\begin{aligned} e^{-itH_\Lambda} L_{z,\alpha} e^{itH_\Lambda} &= e^{-itH_{z\partial}} L_{z,\alpha} e^{itH_{z\partial}} \\ &= \sum_{E,E'\in\text{Eig}(H_{z\partial})} e^{it(E'-E)} P_E L_{\alpha,k} P_{E'} \\ &= \sum_{\omega\in\text{Bohr}(H_{z\partial})} e^{it\omega} \sum_{E,E'\in\text{Eig}(H_{z\partial}):E'-E=\omega} P_E L_{\alpha,k} P_{E'} \\ &\equiv \sum_{\omega\in\text{Bohr}(H_{z\partial})} e^{it\omega} L^\omega_{z,\alpha} \,. \end{aligned} \tag{1.75}$$

In the second step, we decomposed the local Hamiltonian $H_{z\partial}$ into its eigendecomposition, with E as the eigenvalues and P_E as the projections onto the eigenspaces. We then grouped the summands by their Bohr frequency, that is, the difference in eigenvalues

$$\text{Bohr}(H_{z\partial}) \equiv \{E - E' : E, E' \in \text{Eig}(H_{z\partial})\}$$

and set

$$L^\omega_{z,\alpha} \equiv \sum_{E,E'\in\text{Eig}(H_{z\partial}):E'-E=\omega} P_E L_{z,\alpha} P_{E'} \,.$$

One important definition, which further justifies the first step in eq. (1.75), remains: that of the set $z\partial$, which is shorthand for $\{z\}$ together with its boundary. We define the boundary of a set $A \subseteq \Lambda$ as

$$\partial A = \{z \in \Lambda \setminus A : \text{dist}(z, A) \leq r\} \tag{1.76}$$

and $A\partial$ as shorthand for $A \sqcup \partial A$. Although evident from the definition, we wish to emphasise here that the boundary is within Λ and not $\mathbb{Z}^D$. This definition is due to the fact that once Λ and hence the local Hamiltonian are fixed, all other quantities are defined relative to them. With this notation we decompose $H_\Lambda = (H_\Lambda - H_{z\partial}) + H_{z\partial}$, bundling all terms with non-trivial support on z into the local Hamiltonian $H_{z\partial}$. The final objects to define are the jump rates $\chi^{\beta,\omega} \in \mathbb{R}_+$, which we assume to satisfy

$$\min\{\chi^{\beta,\omega} : z \in \Lambda, \omega \in \mathrm{Bohr}(H_{z\partial})\} = \Omega(1) \,. \tag{1.77}$$

One could allow them to have explicit site (z) and basis (α) dependence, but to ensure a uniform single site local gap (that is $\min\{\lambda(\mathcal{L}_z^\dagger) : z \in \Lambda\} = \Omega(1)$, independent of Λ) and, consequently[7], a uniform single site local CMLSI, one would require uniform bounds, such as eq. (1.77) anyway. Thus, we have decided to omit these dependencies, as they primarily add technical complexity without significant conceptual value for the discussion. What remains consistent, whether locality indices are included or not, is that temperature appears exclusively in the definition of the jump rates. Furthermore, they must satisfy the KMS-condition, i.e.,

$$\chi^{\beta,-\omega} = e^{-\beta\omega}\chi^{\beta,\omega} \,. \tag{1.78}$$

This condition guarantees the GNS-symmetry, with respect to the GNS-inner product (see eq. (1.30)) defined using σ, for each $\mathcal{L}_z^\dagger$ (that is their HS-adjoint) individually, and consequently for $\mathcal{L}_\Lambda^\dagger$ as a whole. Note that when we talk just about GNS-symmetry in the context of the Davies semigroup without specifying the state, we always mean it as being defined with respect to the Gibbs state on Λ, i.e, $\sigma = \sigma^\Lambda$.
Analogous to the concept of a local Hamiltonian, we can define the local Davies-Lindbladian on $A \subseteq \Lambda$, relative to the global one in eq. (1.74), as

$$\mathcal{L}_A = \mathcal{L}_{A\subseteq\Lambda} = \sum_{z\in A} \mathcal{L}_z \,, \tag{1.79}$$

where the leftmost notation, $\mathcal{L}_A$, is a shorthand used when the global system Λ is unambiguous from the context, while $\mathcal{L}_{A\subseteq\Lambda}$ is employed to explicitly specify the local operator's dependence on the global system Λ. This is a notable difference to the local Hamiltonian on A, that is H_A (eq. (1.47)) which is defined independently of H_Λ.
A further difference from H_A which is supported (acts non-trivially) on $\mathcal{H}_A$ is that the local Lindbladien $\mathcal{L}_A$ and its dual are supported on $A\partial$. Hence, they act non-trivially on $\mathcal{B}(\mathcal{H}_{A\partial})$ within $\mathcal{B}(\mathcal{H}_\Lambda) = \mathcal{B}(\mathcal{H}_{A\partial}) \otimes \mathcal{B}(\mathcal{H}_{\Lambda\setminus A\partial})$. This is a consequence of $\mathcal{L}_z$ being supported on $z\partial$, as can be deduced from eq. (1.75).
Since $\mathcal{L}_A^\dagger$ is also GNS-symmetric, we can apply the theory described in section 1.3.3 to conclude that

$$E_A^\dagger \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T e^{t\,\mathcal{L}_A}\,dt = \lim_{t\to\infty} e^{t\,\mathcal{L}_A} \tag{1.80}$$

is the HS-adjoint of a conditional expectation E_A, the latter of which maps $\mathcal{B}(\mathcal{H}_\Lambda)$ onto

$$\{L_{z,\alpha}^\omega : z \in A, \alpha = 1, \dots, d^2, \omega \in \mathrm{Bohr}(H_{z\partial})\}' \,,$$

i.e., E_A and $E_A^\dagger$ are supported on $A\partial$ too. Furthermore, E_A inherits GNS-symmetry from $\mathcal{L}_A^\dagger$ and thereby fulfils the fourth condition in theorem 1.3.2, hence also all results from theorem 1.3.1 hold. Due to the localised support of $\mathcal{L}_A$ and hence E_A, these conditional expectation further factorisation over suitably distant sets. More precisely, for $A, B \subset \Lambda$ such that $A\partial \cap B\partial = \emptyset$, we have the identity:

$$E_{A\cup B} = E_A E_B = E_B E_A \,, \tag{1.81}$$

which by duality also holds for $E_{A\cup B}^\dagger$, $E_A^\dagger$, and $E_B^\dagger$. The proof follows directly from the characterisation in eq. (1.80), the decomposition $\mathcal{L}_{A\cup B} = \mathcal{L}_A + \mathcal{L}_B$, and the commutativity $[\mathcal{L}_A, \mathcal{L}_B] = 0$, being a consequence of the disjoint supports implied by $A\partial \cap B\partial = \emptyset$.
Besides the mathematical consequence we have seen above, the locality of the Lindbladien terms $\mathcal{L}_z$ is

[7] This connection is discussed later and boils down to eq. (1.85) from [Gao+22].

necessary for efficient simulation [KB16] and crucially relies on the underlying interaction to be commuting and local, which is why the Davies generator is primarily suited as a Gibbs sampler for local commuting interactions. The local non-commuting case has only recently seen breakthroughs in [Che+23; DLL25b], which have established quasi-local Lindbladians that, for certain parameter configurations, reduce to the Davies dynamics.
Let us now turn to the literature analysing the speed of convergence of the semigroup $(e^{t\,\mathcal{L}_\Lambda})_{t\in\mathbb{R}_+}$ to $E_\Lambda^\dagger$, defined through the mixing time $t_{\mathrm{mix}}(\mathcal{L}_\Lambda;1/2)$. We will focus here on strategies analysing the gap and MLSI, and will not delve into the newer approaches presented in [RFA25; RFA24] for analysing the samplers proposed in [Che+23], as they explicitly target the high-temperature regime, rather than universal decay conditions defined in terms of explicit or implicit correlation measures of the Gibbs state.
We will also separate the discussion of the gap and the MLSI and remind the reader that if both exhibit the same scaling (that is, $\lambda(\mathcal{L}_\Lambda^\dagger)=\Omega(f(|\Lambda|))$ and $\alpha(\mathcal{L}_\Lambda)=\Omega(f(|\Lambda|)))$, then the MLSI yields an exponentially faster mixing time with respect to system size. This is due to the logarithmic (eq. (1.37)), as opposed to linear (eq. (1.34)), dependence on $\log\left\|\sigma^{-1}\right\|_\infty=\Theta(|\Lambda|)$ ([Cap+24, Proof of Lemma B.6]).
Although the gap and MLSI are different quantities, the heuristic strategy to control them for generators composed of sums of local terms (such as the Davies generator, eq. (1.74)) largely follows paths paved by the classical literature on the analysis of Glauber dynamics (see [Mar99]), albeit with some quantum-specific caveats. Inspired by the so-called Dobrushin-Shlosman condition or strong analyticity [DS85; DS87; Ces01] (later weakened to a necessary condition by Martinelli [Mar99]), the analysis of the gap or MLSI for quantum systems can be broadly categorised into two approaches with similar characteristics. These approaches aim to derive the gap ($\lambda(\mathcal{L}_\Lambda^\dagger)$) or MLSI ($\alpha(\mathcal{L}_\Lambda)$) for the entire system Λ by:

1. Showing a splitting of the local quantities, i.e., gap or MLSI, where $\mu\in\{\lambda,\alpha\}$:

$$\mu(\mathcal{L}_{A\cup B}^{(\dagger)})\geq\frac{1}{1-f^\mu(A,B,A\cup B,\sigma)}\min\{\mu(\mathcal{L}_A^{(\dagger)}),\mu(\mathcal{L}_B^{(\dagger)})\}\tag{1.82}$$

   with an interpretable function $f^\mu(A,B,A\cup B,\sigma)$ that ideally depends only on the Gibbs state and the regions (a correlation measure on marginals of the Gibbs state). As apparent from the structure, such an inequality can only be established for suitably chosen sets $A,B\subseteq\Lambda$ such that $f^\mu(A,B,A\cup B,\sigma))<1$. This condition is typically met by showing (or assuming) that f^μ decays with the separation between $A\setminus B$ and $B\setminus A$ (e.g., $\mathrm{dist}(A\setminus B,B\setminus A)$), then choosing A and B accordingly. In the classical literature, this corresponds to the mentioned Dobrushin-Shlosman condition or strong analyticity, later refined by Martinelli [Mar99]. Together with a suitable covering of the lattice, such a decay allows for the subsequent decomposition of the gap or MLSI by eq. (1.82) until system size independent gap or MLSI (CMLSI) constants are reached, thereby closing the argument.

2. Relating the gap or MLSI at a global level to a proxy quantity, such as an alternative generator constructed from the sum of local ground state projectors. One then analyses this proxy by following the same strategy described above for the original quantities. That is, one attempts to find a suitable analogue of eq. (1.82) for this proxy. This, combined with a lattice covering, then allows for a reduction to local terms. Such strategy are particularly relevant in the quantum setting, where the direct approach might yield a function $f^\mu(A,B,A\cup B,\sigma)$ that only implicitly features the Gibbs state, making it difficult or impossible to analyse directly.

The initial work by [KB16], which proposed Gibbs sampling using Davies dynamics, analyses the gap following the first approach. The authors show $\lambda(\mathcal{L}_\Lambda^\dagger)=\Omega(1)$ for all $\Lambda\Subset\mathbb{Z}^D$ under the existence of a 'strong-clustering' condition: There exist constants c,ξ, both $\Theta(1)$, such that for all $\Lambda\Subset\mathbb{Z}^D$, $A,B\subseteq\Lambda$, and all $X\in\mathcal{B}(\mathcal{H}_\Lambda)$:

$$|\langle(\mathrm{id}-E_{A\cup B})E_A(X),(\mathrm{id}-E_{A\cup B})E_B(X)\rangle_{\sigma,1/2}|\leq c\,\langle X,X\rangle_{\sigma,1/2}\,e^{-\mathrm{dist}(B\setminus A,A\setminus B)/\xi}\,.\tag{1.83}$$

With this, they concluded that for sufficiently overlapping $A,B\subseteq\Lambda$ (i.e., large $\mathrm{dist}(B\setminus A,A\setminus B)$), one can split the local gap as:

$$\lambda(\mathcal{L}_{A\cup B}^\dagger)\geq\frac{1}{1-2\varepsilon}\min\{\lambda(\mathcal{L}_A^\dagger),\lambda(\mathcal{L}_B^\dagger)\}$$

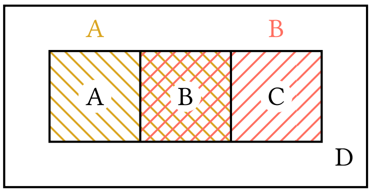


Figure 1.2: Schematic of the setup starting from $A, B \subseteq \Lambda$ and relabelling $A \setminus B \to A$, $A \cap B \to B$, $B \setminus A \to C$ and $\Lambda \setminus A \cup B \to D$ to obtain a disjoint decomposition $\Lambda = A \sqcup B \sqcup C \sqcup D$.

where $\varepsilon = ce^{-\operatorname{dist}(B\setminus A, A\setminus B)/\xi} < \frac{1}{2}$. Using a suitable (overlapping) covering, they then decomposed the global gap into a controllable coefficient term and a minimum over effectively single site local gaps. These single site local gaps are, by assumption, uniformly lower-bounded (see the preceding discussion regarding jump rates).

Unfortunately, beyond the case of commuting one-dimensional interactions, where decay can be demonstrated [KB16, Theorem 28, Proposition 29], eq. (1.83) remains difficult to establish in practice. The challenge arises primarily from the implicit nature of the Davies conditional expectations (i.e., the E_A). For example, it is expected that the condition holds at high temperatures for general D-dimensional systems, although a rigorous proof of this remains open.

While not in the context of Gibbs sampling and with more specific interactions in mind, some spectral analysis of the Davies semigroup has been carried out in [AFH09; TK15; Tem16; KLCT16; LPGPH23], partially even before [KB16], and following one of the general approaches outlined above. However, these results often depend heavily on the particular structure of the interaction in question and do not yield a general measure of correlation.

Despite initial focus on specific cases, this line of research culminated in a recent manuscript [LPGPH25], presented at a 2024 workshop in Tübingen. This work arguably offers the most comprehensive and unified gap analysis to date for the Davies semigroup associated with local commuting interactions. In this work, the authors follow the second strategy, using a proxy quantity they term the ‘purified Hamiltonian’. This is a linear operator $\mathcal{O}_\Lambda : \mathcal{B}(\mathcal{H}_\Lambda) \to \mathcal{B}(\mathcal{H}_\Lambda)$ that is GNS-symmetric, has σ as its ground state, and otherwise possesses a negative spectrum. They then show that

$$\lambda(\mathcal{L}_\Lambda^\dagger) \geq \lambda(\mathcal{O}_\Lambda) \min_{z\in\Lambda} \lambda(\mathcal{L}_z^\dagger)\,,$$

not only for the Davies generator, but for any Lindbladian that is a sum of local, frustration free terms, all of which are GNS-symmetric with respect to σ. Since $\mathcal{O}_\Lambda$ is GNS- and hence KMS-symmetric, the gap $\lambda(\mathcal{O}_\Lambda)$ in the above expression is defined just as eq. (1.31) with E_Λ being derived from $\mathcal{L}_\Lambda^\dagger$ (eq. (1.27) or analogously eq. (1.80)) projecting on the agreeing kernels $\ker \mathcal{L}_\Lambda^\dagger = \ker \mathcal{O}_\Lambda$.

To estimate the gap of this proxy, they use its specific structure as a sum of local projections and a lattice decomposition, together with a correlations measure, to unify the local sums into the ground state projection E_Λ. This measure, for two sets $A, B \subseteq \Lambda$, compares marginals of the Gibbs state and becomes easier to understand if one performs the following relabelling: $A \setminus B \to A$, $B \setminus A \to C$, $A \cap B \to B$ and $\Lambda \setminus (A \cup B) \to D$, thereby disjointly decomposing $\Lambda = A \sqcup B \sqcup C \sqcup D$ (a schematic is provided in figure 1.2). The measure is then given by:

$$\begin{aligned} &\Delta(A : C|D)_\sigma \\ &\equiv \sup\{\left|\operatorname{Tr}_{ADC}[(\sigma_{ADC} - \sigma_{AD}\sigma_D^{-1}\sigma_{DC})Q_{DC}^* R_{AD}]\right| : \|Q_{DC}\|_\sigma \leq 1, \|R_{AD}\|_\sigma \leq 1\}\,. \end{aligned} \tag{1.84}$$

One immediately notices common aspects with, for example, the CMI discussed in section 1.3.1. Furthermore, it reduces more straightforwardly to classical conditions (cf. [DS85; DS87; Ces01; Mar99]), while further being upper-bounded by $\frac{1}{2}(\left\|\mathbb{1} - (\sigma_{AD}\sigma_D^{-1}\sigma_{DC})\sigma_{ADC}^{-1}\right\|_\infty + \left\|\mathbb{1} - \sigma_{ADC}^{-1}(\sigma_{AD}\sigma_D^{-1}\sigma_{DC})\right\|_\infty)$. This bound can be shown to decay exponentially with the separation between A and C at high temperatures for marginal-commuting systems, via a slight modification of [Cap+24, Lemma D.1]. Moreover, for local, translation-invariant one-dimensional systems, we expect the bound to decay at all temperatures, even without the marginal-commuting assumption, based on a combination of the proof strategies in [BCPH22]

and theorem 3.2.3.
For the MLSI, not only is the correlation measure more complicated, but the very existence of MLSI/CMLSI constants for the local/global generators, is not trivial. Unlike the local/global gap, which is a spectral property of a self-adjoint operator and thereby is positive for finite-dimensional systems, the positivity of the CMLSI constant for GNS-symmetric generators was only proven in [GR22] and later improved in [Gao+22]. In both cases, the authors achieved this by lower-bounding the CMLSI by the gap and the so-called complete Pimsner-Popa index. The improved version states that for a generator $\mathcal{L}^\dagger : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ that is GNS-symmetric with respect to a full-rank invariant state, with E being the corresponding conditional expectation onto its kernel:

$$\alpha_c(\mathcal{L}) \geq \frac{\lambda(\mathcal{L}^\dagger)}{2\log(10C_c(E))}\,, \tag{1.85}$$

where the index is given as

$$C(E) = \inf\{\lambda : X \leq \lambda E(X), \forall X \in \mathcal{B}(\mathcal{H})\} \quad \text{and} \quad C_c(E) = \sup\{C(\mathrm{id}_n \otimes E) : n \in \mathbb{N}\}\,.$$

For our setting specifically, i.e., Davies generator derived from local commuting interactions, it holds that $\max\{\log C_c(E_A)/|A| : A \subseteq \Lambda\} = O(1)$ (see e.g., [Cap+24, Lemma B.9]). While this resolves the question of positivity of the local and global MLSI/CMLSI for Davies generators, the relation will reappear in the discussion (section 3.3) of the results of [Cap+24] as a necessary tool.
An initial attempt at a MLSI analysis now might involve translating the strategy from [KB16], using an analogue of the 'strong-clustering' condition, by replacing the inner product with the relative entropy. Such an inequality is often termed approximate tensorisation or approximate entropy factorisation and is formally given as: There exists a similarity measure f such that for all $\Lambda \Subset \mathbb{Z}^D$ and all $A, B \subseteq \Lambda$:

$$(1 - f^\alpha(E_{A\cup B}, E_A, E_B))D(\rho\|E^\dagger_{A\cup B}(\rho)) \leq D(\rho)\|E^\dagger_A(\rho)) + D(\rho\|E^\dagger_B(\rho))\,. \tag{1.86}$$

This measure should intuitively quantify how close $E_{A\cup B}$ (or $E^\dagger_{A\cup B}$) is to behaving like a composition of E_A and E_B (or $E^\dagger_A$ and $E^\dagger_B$), in analogy to eq. (1.83). For the case of $E_{A\cup B} = E_A E_B$, for example, eq. (1.21) yields the above inequality with $f^\alpha(E_{A\cup B}, E_A, E_B) = 0$. Ideally, one would hope for $f^\alpha(E_{A\cup B}, E_A, E_B)$ to serve as a correlation measure depending on the regions and the Gibbs state, but this connection remains entirely obscure due to the function's implicit form and the subtle dependence of the E_A, E_B and $E_{A\cup B}$ on the Gibbs state marginals.
Neglecting this problem for the moment and assuming that eq. (1.86) holds with $f^\alpha(E_{A\cup B}, E_A, E_B) < 1$, one can immediately deduce a relationship between the MLSI constants analogous to what we have seen for the gap:

$$\alpha(\mathcal{L}_{A\cup B}) \geq \frac{1}{1 - f^\alpha(E_{A\cup B}, E_A, E_B)} \min\{\alpha(\mathcal{L}_A), \alpha(\mathcal{L}_B)\}\,.$$

By just requiring a decay of this implicit $f^\alpha(E_{A\cup B}, E_A, E_B)$ exponential in $\mathrm{dist}(B \setminus A, A \setminus B)$, it is straightforward to translate the strategy of [KB16] to obtain a global MLSI constant $(\alpha(\mathcal{L}_\Lambda))$ from uniform single site CMLSI (which by eq. (1.85) is equivalent to uniform single site local gap).
This strategy, but with an explicit instead of an implicit $f^\alpha(E_{A\cup B}, E_A, E_B)$ that depends solely on the Gibbs state and the regions, was only recently realised for a specific class of interactions (namely, CSS-codes) by Sebastian Stengele and co-authors in as yet unpublished work. Interestingly, their measure can again be identified with corresponding measures found for the classical Glauber dynamics [Mar99].
All other previous literature analysing the MLSI of Davies semigroups employed proxies, analogously to the approach for the gap in [LPGPH25]. This includes the initial proof for non-interacting systems in [CLPG18], which was subsequently followed by proofs for translation-invariant one-dimensional systems in [Bar+24]. Both of these papers, did not replace the generator itself but rather the term $D(\rho\|E^\dagger_A(\rho))$ with a proxy. For this purpose, a so-called conditional relative entropy was introduced in [CLPG18], defined for a positive-definite state $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$ and marginals of the Gibbs state σ as

$$D_A(\rho\|\sigma) \equiv D(\rho\|\sigma) - D(\rho_{\Lambda\setminus A}\|\sigma_{\Lambda\setminus A})\,. \tag{1.87}$$

It relates to $D(\rho\|E^\dagger_A(\rho))$ simply by the inequality

$$D_A(\rho\|\sigma) \leq D(\rho\|E^\dagger_A(\rho)) \tag{1.88}$$

which follows from $E_A^\dagger(\sigma_A) = \sigma$, the DPI and the chain-rule (eq. (1.20)). On the global level, i.e., for $A = \Lambda$, we have $D(\rho \| E_\Lambda^\dagger(\rho)) = D(\rho\|\sigma) = D_\Lambda(\rho\|\sigma)$.
Following eq. (1.86), these papers then aimed for an explicit correlation measure g such that:

$$(1 - g(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D))D_{A\cup B}(\rho\|\sigma) \leq D_A(\rho\|\sigma) + D_B(\rho\|\sigma)\,, \tag{1.89}$$

where the measure g depends on marginals of the Gibbs state (using the relabelling from figure 1.2 to define the arguments of g).
Combined with a uniform single-site local CMLSI and a suitable covering, this once again reduces the task of establishing the existence and scaling of a global MLSI constant ($\alpha(\mathcal{L}_\Lambda)$) to verifying the decay of $g(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D)$. Unlike eq. (1.86), an inequality for this proxy would have implications beyond Davies generators, potentially allowing the same proxy and strategy to be used for other Gibbs samplers, such as those proposed in [CKG23]. At the time of writing, however, such a result (that is, eq. (1.89)) has only been established for the specific case where $A \sqcup B = \Lambda$, or for the trivial case where σ tensorises across all disjoint partition $\Lambda = A \sqcup B \sqcup C \sqcup D$ (note the use of the relabelled sets here); which were shown in [CLPG18]. In the latter case (trivial tensorisation), $g(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D) = 0$, and the inequality reduces to the DPI of the relative entropy or, relatedly, the SSA of the von Neumann entropy. The first case ($A \sqcup B = \Lambda$) was used in [Bar+24] to prove rapid mixing for spin chains with translation-invariant commuting interactions at all temperatures, and the second case (trivial tensorisation) was used in [CLPG18] for non-interacting systems. While the result in [CLPG18] demonstrated an $\Omega(1)$ MLSI constant and was therefore optimal, [Bar+24] initially found $\alpha(\mathcal{L}_\Lambda) = \Omega((\log|\Lambda|)^{-1})$, a scaling that was later improved to $\Omega(1)$ in [Koc+25].
The strategy employed in [Koc+25] builds upon the work in [CRF20], which was specifically tailored to 2-local interactions. Both works also employ a proxy quantity, but instead of entirely replacing $D(\rho\|E_A^\dagger(\rho))$, they substitute the argument $E_A^\dagger(\rho)$ within the relative entropy with $E_A^{S,\dagger}(\rho)$, where E_A^S is the so-called Schmidt conditional expectation. On the global level, these conditional expectations (respectively their HS-adjoints) agree, that is, $E_\Lambda^\dagger(\rho) = \mathrm{Tr}[\rho]\sigma = E_\Lambda^{S,\dagger}(\rho)$, while generally, it holds that $D(\rho\|E_A^{S,\dagger}(\rho)) \leq D(\rho\|E_{A\partial}^\dagger(\rho))$ (note the support $A\partial$ on the right-hand-side). This inequality allows one to relate results obtained for the Schmidt back to the Davies conditional expectation. An approximate tensorisation akin to eq. (1.86), but for these Schmidt conditional expectations (that is, using terms like $D(\rho\|E_A^{S,\dagger}(\rho))$) instead of the Davies ones, was shown and utilised in [CRF20], while culminating in the analysis of 2-local commuting interactions in [Koc+25]. There, the correlation measure from [CRF20] was related to an explicit 'strong local indistinguishability' measure in the Gibbs state. This enabled the demonstration of an $\Omega(1)$ MLSI constant for translation-invariant commuting spin chains at every temperature and an $\Omega(1)$ MLSI constant at high temperature for 2-local commuting systems in any dimension.
In [Cap+24], a crucial idea is to use the proxy $D_A(\rho\|\sigma)$ from eq. (1.87) and to weaken the requirement of an inequality like eq. (1.89). Specifically, we allow for an additive error term that involves an explicit correlation measure in the Gibbs state, rather than a purely multiplicative prefactor. Although such modification does not yield a direct MLSI inequality, it leads to a relative entropy decay bound similar to eq. (1.36) but with an additional additive error δ:

$$D(e^{t\,\mathcal{L}_\Lambda}(\rho)\|\sigma) \leq e^{-\alpha t} D(\rho\|\sigma) + \delta\,. \tag{1.90}$$

Assuming one can derive this inequality for $\delta < 1$, this then gives a mixing time:

$$t_{\mathrm{mix}}(\mathcal{L}_\Lambda; \sqrt{\delta}) \leq \frac{1}{\alpha} \log \frac{\sqrt{\log\|\sigma^{-1}\|_\infty}}{\delta}\,, \tag{1.91}$$

which can be 'lifted' to a gap estimate as well as to an entropy contraction at multiples of the mixing time. That is, for all $n \in \mathbb{N}$:

$$D(e^{n t_{\mathrm{mix}}(\mathcal{L}_\Lambda;\sqrt{\delta})\,\mathcal{L}_\Lambda}\rho\|\sigma) \leq \sqrt{\delta}^{\,n} D(\rho\|\sigma)\,. \tag{1.92}$$

Unfortunately, however, this does not directly yield a MLSI constant, which hints at a possible fundamental distinction between scaling of MLSI and scaling of mixing time. A clear and formal statement of the objectives, along with the assumptions imposed to derive the above results, can be found in section 2.3. The derivation of the results are presented in section 3.3, accompanied by additional discussion and further insights into interrelation between mixing time, spectral gap, and MLSI.

1.4.4 A short introduction to open bosonic systems ([GMR24])

In section 1.3.3, we already discussed the general difficulty of proving that a formal generator in GKSL-form indeed is a core to the generator of a QMS, or put differently that the formal Cauchy problem (eq. (1.42)) is well-posed, meaning for a given initial value in a dense set has a differentiable solution fulfilling the differential equation. In this section, we now provide a more detailed discussion and introduce the specific setting covered in [GMR24], along with its motivation and relevant previous work.
We first introduce the formal generators under consideration, where the term 'formal' here signifies that the given operator structurally resembles a generator (i.e., has GKSL-form) but in all generality will not be a generator to a QMS nor a core to one. Given that the action of polynomials (for non-commutative x, y denoted as $\mathbb{C}[x, y]$) in annihilation a and creation a^* operators is naturally defined on $\operatorname{span}\{|n\rangle : n \in \mathbb{N}\}$, the appropriate domain for a formal generator constructed from such operators is the set of finite-rank operators on $\mathcal{F}$:

$$\mathcal{T}_f(\mathcal{F}) \equiv \operatorname{span}\{|n\rangle\langle m| : n, m \in \mathbb{N}_0\} \subset \mathcal{T}(\mathcal{F})\,. \tag{1.93}$$

On this domain, we set a formal generator of an open bosonic system in the Schrödinger picture to be

$$\mathcal{L}(X) = -i[H, X] + \sum_{j=1}^{J} \Big(L_j X L_j^{\circledast} - \frac{1}{2}\{L_j^{\circledast} L_j, X\}\Big) \tag{1.94}$$

for $X \in \mathcal{T}_f(\mathcal{H})$, $p_j \in \mathbb{C}[x, y]$ for $j = 0, \dots K$, $H = p_0(a, a^*)$ and $L_j = p_j(a, a^*)$ for $k = 1, \dots, K$. In the above notation, $L_j^{\circledast}$ refers to the formal adjoint, defined such that $L_j^{\circledast} = p_j^{\circledast}(a, a^*)$. The polynomial $p_j^{\circledast}$ is obtained by complex conjugating all coefficients of p_j and swapping the exponents of x and y in each term (for example, cx^py^q becomes $\overline{c}x^qy^p$). H is required to be a formal Hamiltonian, that is, $H = H^{\circledast}$. The question now is whether $(\mathcal{L}, \mathcal{T}_f(\mathcal{F}))$ constitutes the core of a generator of a QMS.
With respect to the choice of Hilbert space and constituent operators, this setting is a restriction of the framework considered by Davies (discussed in section 1.3.3). However, in terms of initial assumptions, it is weaker, as it does not, a priori, require the operator $G = -iH - \frac{1}{2}\sum_{j=1}^{J} L_j^{\circledast} L_j$ to generate a contractive C_0-semigroup on $\mathcal{F}$.
The mathematical study of more specific open bosonic systems dates back to the last century, beginning with the study of quantum Ornstein-Uhlenbeck (OU) semigroups. Those are defined as $L_1 = \mu a^*$ and $L_2 = \mu a$ for $\mu, \nu \in \mathbb{R}$ and serve as theoretical models for laser and maser systems [FM19]. For these semigroups, the study of existence and convergence properties heavily relied on their unique, faithful normal invariant state, with respect to which the generator is GNS-symmetric [CM17]. Consequently, Hilbert space tools were available and could be used to not only sidestep the minimal semigroup problem (see the discussion in section 1.3.3) but even demonstrate their hypercontractivity [CS07].
More recently, these existence results and the parameter ranges for which such faithful normal invariant states exist have been generalised to Gaussian QMS [AFP21], where $H = p_0(a, a^*)$ with $p_0 \in \mathbb{C}[x, y]$ a quadratic, and $L_1 = p_1(a, a^*), L_2 = p_2(a, a^*)$ with $p_1, p_2 \in \mathbb{C}[x, y]$ linear polynomials. The tools used to go beyond minimal semigroups in this context were developed in [Fag18] and tailored to this specific setting. While Gaussian generators encompass a broad range of physical systems [EW07] and are relevant from a quantum information perspective [EW07; Wan+07; HHW10], there is growing interest in semigroups that go beyond the Gaussian class, particularly in the context of quantum error correction [Mir+14; GM19; Pur+19; Cha+22]. The so-called cat codes (more precisely, the QMS that implement them) theoretically stabilise a two-dimensional subspace—the code space—in $\mathcal{T}(\mathcal{F})$ and further implement a universal gate set on this logical qubit space. By engineering of the underlying dynamic, the code space and the universal gate set are protected against a dominant error in the physical system (boson loss), thereby simplifying the task of active error correction [Mir+14] via this self-correction mechanism. The single-mode gates with corresponding formal generator and dominant error are given as:

1. Identity-gate (stabilising dynamic): $L_1 = a^2 - \alpha^2$ for $\alpha \in \mathbb{R}$;
2. $Z(\theta)$-gate: $L_1 = a^2 - \alpha^2$ for $\alpha \in \mathbb{R}$, $H = a + a^*$;
3. Identity-gate and boson loss: $L_1 = a^2 - \alpha^2$ for $\alpha \in \mathbb{R}$, $L_2 = a$;

4. $Z(\theta)$-gate and boson loss: $L_1 = a^2 - \alpha^2$ for $\alpha \in \mathbb{R}$, $L_2 = a$, $H = a + a^*$.

We could introduce real parameters multiplying every L_i and the H, but as they do not alter the discussion, they are omitted. Note that we have excluded here the X, CNOT, and Toffoli gates, as we will not discuss the generalisations to time-dependent and multimodal generators from [GMR24]. While generally possible, these extensions primarily introduce technical complications while the underlying ideas remain largely the same.

Papers discussing these codes primarily focus on implementation and engineering aspects and their correction properties on the code space, while lacking the mathematical framework to describe the general dynamics [Mir+14; GM19; Pur+19; Cha+22]. Well-posedness is often considered a consequence of a derivation [GGK94] involving the system's mode weakly coupling to an external drive (acting as a bath or reservoir), such that the effective dynamics on the system mode follow the formal generator of either of the above gates. This strategy is comparable to that in [Dav74], but neglects the unbounded nature of the system operators. Other papers [Mir+14; Leg+15] employ a strategy they term adiabatic elimination to justify well-posedness of the described dynamics, presenting experimental measurements and numerical results that underpin their findings.

In [ASR15], the authors, for example, analyse the fixed point set of the identity-gate and its perturbation by boson loss via this method of 'adiabatic elimination'. However, they do not explicitly demonstrate the existence of either dynamics, which is what they build upon when doing 'adiabatic elimination' in their perturbation analysis. Instead, the authors prove that for any $k \geq 1$, there exists a μ_k such that

$$\operatorname{Tr}\left[\,\mathcal{L}[a^2-\alpha^2](\rho)\,\mathbf{N}^k\,\right] \leq -k\left(\operatorname{Tr}\left[\rho\,\mathbf{N}^k\,\right]\right)^{\frac{k+1}{k}} + \mu_k \tag{1.95}$$

for all quantum states $\rho \in \mathcal{T}_f(\mathcal{F})$, where $\mathcal{L}[L](\rho) \equiv L\rho L^{\circledast} - \frac{1}{2}\{L^{\circledast}L, \rho\}$ is used as a shorthand notation. They then argue that this guarantees the existence of the QMS, not only for the identity-gate but also for the perturbed operator, provided the prefactor of that perturbation is small. Furthermore, they claim

$$\operatorname{Tr}\left[e^{t\overline{\mathcal{L}[a^2-\alpha^2]}}(\rho)\,\mathbf{N}^k\,\right] \leq \max\left\{\left(\frac{\mu_k}{k}\right)^{\frac{k+1}{k}}, \operatorname{Tr}\left[\rho\,\mathbf{N}^k\,\right]\right\} \tag{1.96}$$

for all $\rho \in \mathcal{T}(\mathcal{F})$. The primary unresolved issue is whether $\mathcal{L}[a^2-\alpha^2]$ together with the domain $\mathcal{T}_f(\mathcal{F})$ indeed is the core to a generator of a QMS. Furthermore, it is not established that $\mathcal{T}_f(\mathcal{F})$ constitutes a core for this generator conformal with $\operatorname{Tr}\left[\,\cdot\,\mathbf{N}^k\,\right]$, meaning that eq. (1.95) by density can be extended to all states within the set

$$\bigcup_{t\in\mathbb{R}_+} e^{t\overline{\mathcal{L}[a^2-\alpha^2]}}(\mathcal{T}_f(\mathcal{F}))\,.$$

This, however, is a necessary requirement to affirm the differentiability of

$$t \mapsto \operatorname{Tr}\left[e^{t\overline{\mathcal{L}[a^2-\alpha^2]}}(\rho)\,\mathbf{N}^k\,\right],$$

for a state $\rho \in \mathcal{T}_f(\mathcal{F})$ and finally conclude eq. (1.96). All these aspects remain unaddressed in [ASR15] and remained open even after a follow-up paper by the same authors [ASR16].

Although it does not address these issues, this work remains, to the best of our knowledge, the only rigorous derivation of existence and convergence results for formal generators in the context of cat codes. In that work, the authors consider the 'l-legged' cat identity-gate—a generalisation of the '2-legged' version previously discussed, defined by $(\mathcal{L}[a^l-\alpha^l], \mathcal{T}_f(\mathcal{F}))$ for $l \geq 2$, $\alpha \in \mathbb{R}$. Deviating from their initial approach of using the number operator to argue for stability, they instead use an operator tailored to the generator specifically and adopt a strategy similar to that of Davies (see discussion in section 1.3.3), however considering two spaces: $(\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$ and further $(\mathcal{D}(S(l)\cdot S(l)), \|S(l)\cdot S(l)\|_1)$, where $S(l) = \mathbb{1} + L(l)^* L(l)$ with $L(l) \equiv a^l - \alpha^l$. They demonstrate that the latter is a Banach space which is, furthermore, compactly embedded into $(\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$, that is:

1. $\mathcal{D}(S(l)\cdot S(l)) \subset \mathcal{T}(\mathcal{F})$,
2. $\|\cdot\|_1 \leq \|S(l)\cdot S(l)\|_1$ on $\mathcal{D}(S(l)\cdot S(l))$,

3. bounded sets in $(\mathcal{D}(S(l)\cdot S(l)), \|S(l)\cdot S(l)\|_1)$ are precompact (closure is compact) in $(\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$.

Unlike Davies, who only obtained minimal semigroups on $(\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$, Azouit et al. establish that these are also contractive C_0-semigroups on $(\mathcal{D}(S(l)\cdot S(l)), \|S(l)\cdot S(l)\|_1)$. This property, via the compact embedding, permits them to take the limit $\delta \to 1$ in eq. (1.44) and to conclude that $(\mathcal{L}[a^l - \alpha^l], \mathcal{T}_f(\mathcal{F}))$ forms the core of a generator for a QMS a strategy akin to the one in [Che03]. Furthermore, they succeed in demonstrating convergence to the kernel of $(L(l)\cdot L(l)^*)$, in the sense that for states $\rho \in \mathcal{D}(S(l)\cdot S(l))$,

$$\operatorname{Tr}\Big[L(l)e^{t\overline{\mathcal{L}[a^l-\alpha^l]}}(\rho)L(l)^*\Big] \le e^{-l!t}\operatorname{Tr}[L(l)\rho L(l)^*]\,. \tag{1.97}$$

As obvious from the above summary this approach is quite specific to the identity gate of the 'l-legged' cat code and lacks applicability to the perturbative context investigated in [ASR15]. The 2015 work, in contrast, lacks the mathematical rigour of the 2016 study. The motivation behind [GMR24], therefore, was to put the heuristics of [ASR15] on a solid mathematical foundation and establish a framework capable of treating all formal generators associated with the cat code, the OU process, and Gaussian semigroups. At the same time one should be able to conduct perturbation analysis of these systems by adapting the approach from [ASR16] using compactly embedded spaces, but now based on the number operator, as in [ASR15].
To this end, we introduce the notation for QSSs, discuss some of their properties, and define the concept of Sobolev-preserving semigroups. Let us begin with the definition of the QSS of order k:

Definition 1.4.1 (Quantum Sobolev space of order k [GMR24]) For $k \in \mathbb{R}_+$, we call the Banach space $(\mathcal{W}^k, \|\cdot\|_{\mathcal{W}^k})$, with

$$\mathcal{W}^k = \{(\mathbf{N}+\mathbb{1})^{-k/2}X(\mathbf{N}+\mathbb{1})^{-k/2} : X \in \mathcal{T}(\mathcal{F})\} \quad \text{and} \quad \|X\|_{\mathcal{W}^k} = \left\|(\mathbf{N}+\mathbb{1})^{k/2}X(\mathbf{N}+\mathbb{1})^{k/2}\right\|_1$$

for $X \in \mathcal{W}^k$, the QSS of order k. For $k=0$, naturally $(\mathcal{W}^k, \|\cdot\|_{\mathcal{W}^k}) = (\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$.

The name was chosen due to the resemblance of classical Sobolev spaces: The number operator $\mathbf{N}$ is proportional to the Laplacian on the isometrically isomorphic Hilbert space $L^2(\mathbb{R})$, thus, membership in $\mathcal{W}^k$ mirrors the existence and integrability of higher-order derivatives that characterise classical Sobolev spaces. Similar to their classical counterparts, QSSs are Banach spaces and, as demonstrated in [GMR24], are also compactly embedded into one another. More precisely, for $k' > k$, it holds that $\mathcal{W}^{k'} \subset \mathcal{W}^k$, with $\|\cdot\|_{\mathcal{W}^k} \le \|\cdot\|_{\mathcal{W}^{k'}}$ on $\mathcal{W}^{k'}$, while bounded sets in $(\mathcal{W}^{k'}, \|\cdot\|_{\mathcal{W}^{k'}})$ are precompact (that is, their closure is compact) in $(\mathcal{W}^k, \|\cdot\|_{\mathcal{W}^k})$. For $k, k' \in \mathbb{R}_+$, we can define the set of bounded linear operators $\mathcal{O} \in \mathcal{B}(\mathcal{W}^k, \mathcal{W}^{k'})$ mapping from $\mathcal{W}^k$ to $\mathcal{W}^{k'}$, which are bounded with respect to the norm

$$\|\mathcal{O}\|_{\mathcal{W}^{k'}\to\mathcal{W}^k} \equiv \sup_{X\in\mathcal{W}^{k'}\setminus\{0\}} \frac{\|\mathcal{O}(X)\|_{\mathcal{W}^k}}{\|X\|_{\mathcal{W}^{k'}}}\,. \tag{1.98}$$

By standard Banach space theory, $(\mathcal{B}(\mathcal{W}^{k'}, \mathcal{W}^k), \|\cdot\|_{\mathcal{W}^{k'}\to\mathcal{W}^k})$ is itself a Banach space for which we can prove a quantum analogue of the Stein-Weiss interpolation theorem for classical weighted L^p spaces:

Theorem 1.4.2 (Quantum Stein-Weiss theorem [GMR24]) *Let $k, k' \in \mathbb{R}_+$ with $k' > k$.*

1. *Let further $\mathcal{O} \in \mathcal{B}(\mathcal{W}^k)$ such that $\mathcal{O}|_{\mathcal{W}^{k'}} \in \mathcal{B}(\mathcal{W}^{k'})$ with bounds $c = \|\mathcal{O}\|_{\mathcal{W}^k\to\mathcal{W}^k}, c' = \|\mathcal{O}\|_{\mathcal{W}^{k'}\to\mathcal{W}^{k'}}$ respectively. Then $\mathcal{O}|_{\mathcal{W}^{k''}} \in \mathcal{B}(\mathcal{W}^{k''})$ for $k'' \in [k, k']$ and*

$$\|\mathcal{O}\|_{\mathcal{W}^{k''}\to\mathcal{W}^{k''}} \le c^{\frac{k'-k''}{k'-k}}(c')^{\frac{k''-k}{k'-k}}\,.$$

2. *Let further $\mathcal{O} \in \mathcal{B}(\mathcal{T}(\mathcal{F}), \mathcal{W}^k)$ such that $\mathcal{O} \in \mathcal{B}(\mathcal{T}(\mathcal{F}), \mathcal{W}^{k'})$ with bounds $c = \|\mathcal{O}\|_{1\to\mathcal{W}^k}, c' = \|\mathcal{O}\|_{1\to\mathcal{W}^{k'}}$ respectively. Then $\mathcal{O} \in \mathcal{B}(\mathcal{T}(\mathcal{F}), \mathcal{W}^{k''})$ for $k'' \in [k, k']$ and*

$$\|\mathcal{O}\|_{1\to\mathcal{W}^{k''}} \le c^{\frac{k'-k''}{k'-k}}(c')^{\frac{k''-k}{k'-k}}\,.$$

Note that [GMR24] proves only the first part of theorem 1.4.2; the second part follows from a largely analogous and arguably simpler line of reasoning. In principle, this method could be generalised to variation of the domain QMSs as well, but such an extension is unnecessary for our current purposes and is therefore omitted.
Finally, we can define the concept of a Sobolev-preserving semigroup, where theorem 1.4.2 naturally establishes a hierarchy.

Definition 1.4.3 (Sobolev-preserving semigroup [GMR24]) We call a C_0-semigroup semigroup $(\mathcal{P}_t)_{t\in\mathbb{R}_+}$ on $(\mathcal{T}(\mathcal{F}), \|\cdot\|_1)$ a Sobolev-preserving semigroup of order $k \in \mathbb{R}_+$, if $(\mathcal{P}_t\,|_{\mathcal{W}^k})_{t\in\mathbb{R}_+}$ is a C_0-semigroup on $(\mathcal{W}^k, \|\cdot\|_{\mathcal{W}^k})$. We furthermore call $(\mathcal{P}_t)_{t\in\mathbb{R}_+}$ Sobolev-preserving if it is Sobolev-preserving for all $k \in \mathbb{R}_+$.

Note that, by virtue of theorem 1.4.2 and the compact embedding of QSSs, a Sobolev-preserving semigroup of order k' is immediately Sobolev-preserving for all $k < k'$ employing [Nag00, Proposition 5.3] for the only non-trivial property of strong continuity.
Returning to our initial objective—to formalise the implication from eq. (1.95) to eq. (1.96)—we can now establish that if $(e^{t\mathcal{L}[a^2-\alpha^2]})_{t\in\mathbb{R}_+}$ was a Sobolev-preserving semigroup, with $(\mathcal{L}[a^2-\alpha^2], \mathcal{T}_f(\mathcal{F}))$ as a core for its generator on all $(\mathcal{W}^k, \|\cdot\|_{\mathcal{W}^k})$ for $k \in \mathbb{R}_+$, then eq. (1.95) would indeed rigorously imply eq. (1.96).
The thorough reader might have noticed that the definitions and results above are slightly more general than the ones presented in [GMR24]. This generality will persist throughout this thesis, so let us shed some light on this in the following remark.

Remark 2. In [GMR24], the results and definitions for QSSs were originally formulated for the real Banach space of self-adjoint trace-class operators. However, it was later recognised by the co-authors and myself—unfortunately too late for inclusion in the published paper—that all proofs remain valid in the more general setting of the complex Banach space of trace-class operators $\mathcal{T}(\mathcal{F})$, along with the corresponding QSSs, rather than being restricted to the real Banach space of its self-adjoint subset (and the respective subspaces of the Sobolev spaces). The primary remaining question was whether the bounds on resolvents and semigroups would also hold in this broader context. This has been affirmed, particularly because the semigroups and equivalently also the resolvents under consideration are completely positive. A proof demonstrating the preservation of these bounds, leveraging the complete positivity of the relevant resolvents and semigroups, can be deduced from the proofs of theorem 4.1.1 and Lemma 2.3.5 in [Mö25]. Consequently, the results presented in this thesis, despite their more general formulation, are established by the combined findings of [GMR24] and [Mö25].

The objectives of [GMR24], which clarify the underlying assumptions and define the goals of the work, are presented in section 2.4, while the corresponding results are discussed in section 3.4.

1.4.5 The generalised quantum Stein's lemma for subalgebra resources ([GR24b])

In this section, we briefly address the well-studied field of quantum hypothesis testing [HP91; NO00; OH04; Aud+08; NS09], focusing on the quantum Stein's lemma and its generalised counterpart. While the latter, in particular, has substantial theoretical implications for resource theories, the operational task of hypothesis testing, which forms the core of the results in [GR24b], will be our primary focus here. Although [GR24b] is a special case of more general findings that were subsequently presented in [Lam25; HY24], we hoped that the algebraic proof and distinct insights offered in [GR24b] may still be of value. For these reasons, the following discussion will be kept relatively brief and only discuss the core concepts. For a more comprehensive treatment of recent developments and the surrounding context of resource theories, we refer the interested reader to seminal works such as [BP10a; Ber+23] and the aforementioned [Lam25; HY24].
Analogous to the classical Chernoff-Stein lemma framework [Cov06, Chapter 11.6] for hypothesis testing, the quantum setting is the following: For $n \in \mathbb{N}$ we are presented with a state $\mu \in \{\rho^{\otimes n}, \sigma^{\otimes n}\} \subset \mathcal{S}(\mathcal{H}^{\otimes n})$, where here and throughout, we use the shorthand $\rho^{\otimes n} = \bigotimes_{i=1}^{n} \rho$. Without prior knowledge regarding the likelihood of either state, we have to distinguish these two states by measurement. This means we have the freedom to choose a self-adjoint operator, say M (as discussed in section 1.3.3) and to each of its eigenvalues associate either one of the states. Without loss of generality, we may assume M has the spectral decomposition $M = (+1)P + (-1)(\mathbb{1} - P)$, where P is an orthogonal projector. The projector onto the alternative outcome is then necessarily $\mathbb{1} - P$, as eigenprojections must sum to the identity. A measurement

outcome of $+1$ leads to the conclusion that $\mu = \rho^{\otimes n}$, whereas an outcome of -1 leads to the conclusion that $\mu = \sigma^{\otimes n}$. The expectation values for the errors of type I and type II are therefore:

1. Type I error (false negative): Concluding $\sigma^{\otimes n}$ (measuring -1) when $\mu = \rho^{\otimes n}$.
   Probability: $\alpha_n(P;\rho) = \mathrm{Tr}[\rho^{\otimes n}(\mathbb{1} - P)]$.

2. Type II error (false positive): Concluding $\rho^{\otimes n}$ (measuring $+1$) when $\mu = \sigma^{\otimes n}$.
   Probability: $\beta_n(P;\sigma) = \mathrm{Tr}[\sigma^{\otimes n} P]$.

Both error probabilities depend on the choice of the measurement projector P, which, as noted, is our sole degree of freedom. This focus clarifies why a binary-outcome measurement is sufficient. Even if more measurement outcomes (eigenvalues) were permitted, each would still need to be assigned to one of the two hypotheses (that is, identifying the state as $\rho^{\otimes n}$ or $\sigma^{\otimes n}$). Since the eigenprojections determine the measurement probabilities and, consequently, the hypothesis testing errors, regrouping these assignments would effectively result in the same binary decision structure, defined by two orthogonal projectors. Each of these would be the sum of the original eigenprojections corresponding to outcomes assigned to the same hypothesis.

Given this setup, one objective might be to choose P to minimise both error probabilities simultaneously, for instance, by minimising their sum. This approach leads to the quantum Neyman-Pearson test and its associated asymptotic error rate, which examines the scaling of $\min_{0 \le P \le \mathbb{1}}\{\alpha_n(P) + \beta_n(P)\}$ in n [Aud+07; NS09]. Quantum Stein's lemma, however, addresses a different scenario: it investigates the asymptotic scaling in n of the minimal probability of a type II error, $\beta_n(P)$, subject to the constraint that the probability of a type I error, $\alpha_n(P)$, does not exceed a fixed threshold $\varepsilon \in (0,1)$. Formally, we examine the behaviour of:

$$\beta_n^*(\varepsilon;\rho,\sigma) = \min\{\beta_n(P;\rho) : 0 \le P \le \mathbb{1}, \alpha_n(P;\sigma) \le \varepsilon\}$$

as $n \to \infty$. One intuitively concludes that the error should decay exponentially $\beta_n^*(\varepsilon;\rho,\sigma) \sim e^{-rn}$, leading to the definition of the asymptotic error rate:

$$r(\varepsilon;\rho,\sigma) = \lim_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon;\rho,\sigma) = \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(\rho^{\otimes n}\|\sigma^{\otimes n}), \tag{1.99}$$

where, for better readability, we introduce the hypothesis testing relative entropy for arbitrary states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and $\varepsilon \in (0,1)$ as

$$D_h^\varepsilon(\rho\|\sigma) \equiv \begin{cases} -\log\min\{\mathrm{Tr}[\sigma P] : 0 \le P \le \mathbb{1}, \mathrm{Tr}[\rho(\mathbb{1} - P)] \le \varepsilon\} & \text{if } \ker\sigma \subseteq \ker\rho\,, \\ +\infty & \text{otherwise}\,. \end{cases} \tag{1.100}$$

This quantity satisfies the DPI, can be defined for positive-semidefinite matrices and then is anti-monotonous in the second argument (i.e., eq. (1.15)) [Tom16]. However, unlike the divergences in section 1.3.1, it is not additive but rather subadditive under tensor products:

$$D_h^\varepsilon(\rho\otimes\rho'\|\sigma\otimes\sigma') \le D_h^\varepsilon(\rho\|\sigma) + D_h^\varepsilon(\rho'\|\sigma')\,,$$

as can be shown by straightforward calculation. This subadditivity, via Fekete's lemma [Fek23], ensures a priori (that is, without invoking quantum Stein's lemma itself) that the limit in eq. (1.99) exists, justifying its use over a $\liminf$ or $\limsup$. Quantum Stein's lemma [HP91; NO00] states that this rate is independent of ε:

$$\lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma)\,,$$

while moreover being given by the quantum relative entropy. In this way, the quantum relative entropy acquires an operational interpretation in the context of asymmetric hypothesis testing.

A natural question is how far this framework can be extended, a topic first addressed in [BP10a], where the quantum Stein setting was generalised: n copies of ρ ($\rho^{\otimes n}$) are compared not against a single alternative $\sigma^{\otimes n}$, but against a set of states $\mathcal{S}_n \subseteq \mathcal{S}(\mathcal{H}^{\otimes n})$. That is, the state μ is either $\rho^{\otimes n}$ or some $\sigma \in \mathcal{S}_n$, where σ is chosen from $\mathcal{S}_n$ to maximise the probability of a type II error for a given measurement. The probability of a type I error remains unchanged, but the type II error is now defined as the worst-case probability over the set $\mathcal{S}_n$:

1. Type I error: Concluding $\mu \in \mathcal{S}_n$ when $\mu = \rho^{\otimes n}$. Probability: $\alpha_n(P;\rho) = \mathrm{Tr}[\rho^{\otimes n}(\mathbb{1} - P)]$.
2. Worst case type II error : Concluding $\rho^{\otimes n}$ when $\mu \in \mathcal{S}_n$, where the specific state from $\mathcal{S}_n$ is the one that maximises this error. Probability: $\beta_n(P;\mathcal{S}_n) = \sup\{\mathrm{Tr}[\sigma P] : \sigma \in \mathcal{S}_n\}$.

Analogously to the standard Stein's lemma, we seek the measurement that minimises the worst-case type II error, given a bounded type I error:

$$\beta_n^*(\varepsilon;\rho,\mathcal{S}_n) = \inf\left\{\beta_n(P;\mathcal{S}_n) : 0 \le P \le \mathbb{1}, \alpha_n(P;\rho) \le \varepsilon\right\} . \tag{1.101}$$

Similarly to the original quantum Stein's lemma, the interest lies in the asymptotic scaling of this minimised worst-case error for a fixed $\varepsilon \in (0,1)$, with its tractability depending on the choice of the family of sets $(\mathcal{S}_n)_{n\in\mathbb{N}}$. For instance, setting $\mathcal{S}_n = \{\sigma^{\otimes n}\}$ recovers the standard quantum Stein's lemma, while an unsuitable choice for $(\mathcal{S}_n)_{n\in\mathbb{N}}$ might preclude the interchange of the supremum and infimum in eq. (1.101), or may not yield an exponential decay in the error rate. Motivated by resource theories (particularly the resource theory of entanglement at the time) and to ensure tractability, [BP10a] imposed certain reasonable restrictions on the family $(\mathcal{S}_n)_{n\in\mathbb{N}}$:

1. For all $n \in \mathbb{N}$, $\mathcal{S}_n \subseteq \mathcal{S}(\mathcal{H}^{\otimes n})$ is a convex and compact set.
2. There exists $\sigma_0 \in \mathcal{S}(\mathcal{H})$, positive-definite, such that $\sigma_0^{\otimes n} \in \mathcal{S}_n$ for all $n \in \mathbb{N}$.
3. For all $n, m \in \mathbb{N}$, if $\tau_n \in \mathcal{S}_n$ and $\tau_m \in \mathcal{S}_m$, it holds that $\tau_n \otimes \tau_m \in \mathcal{S}_{n+m}$.
4. For every $n \in \mathbb{N}$ and all permutations π on n elements, with corresponding unitary $U_\pi : \mathcal{H}^{\otimes n} \to \mathcal{H}^{\otimes n}$ implementing the permutation by exchange of the respective Hilbert spaces $\mathcal{H}$, it holds that $U_\pi\, \mathcal{S}_n\, U_\pi^* \subseteq \mathcal{S}_n$ (permutation invariance).

Under these conditions, one still obtains an exponential decay $\sim e^{-rn}$ of the asymptotic error rate given by:

$$\begin{aligned} r(\varepsilon;\rho,(\mathcal{S}_n)_{n\in\mathbb{N}}) &= \lim_{n\to\infty} -\frac{1}{n}\log\beta_n^*(\varepsilon;\rho,\mathcal{S}_n) \\ &= \lim_{n\to\infty}\inf_{\sigma\in\mathcal{S}_n} D_h^\varepsilon(\rho^{\otimes n}\|\sigma) = \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(\rho^{\otimes n}\|\,\mathcal{S}_n) . \end{aligned} \tag{1.102}$$

The interchange of the infimum and supremum in the first line of eq. (1.102) is justified by the constraints on the sequence $(\mathcal{S}_n)_{n\in\mathbb{N}}$, the linearity of the trace functional involved, the application of Sion's minimax theorem [KW20, Theorem 2.18], and the monotonicity of the logarithm. Using the properties of the family of sets again, we further conclude that $s_n = D_h^\varepsilon(\rho^{\otimes n}\|\,\mathcal{S}_n)$ is a subadditive sequence, i.e., $s_{n+m} \le s_n + s_m$ which by Fekete's lemma [Fek23] justifies the use of the limit instead of a $\liminf$ or $\limsup$ in eq. (1.102). Analogously to quantum Stein's lemma, the generalised quantum Stein's lemma asserts that, for all $\varepsilon \in (0,1)$ and a family of sets $(\mathcal{S}_n)_{n\in\mathbb{N}}$ satisfying the above conditions, this limit is independent of ε and given by:

$$\lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(\rho^{\otimes n}\|\,\mathcal{S}_n) = \lim_{n\to\infty}\frac{1}{n}D(\rho^{\otimes n}\|\,\mathcal{S}_n) . \tag{1.103}$$

This result was initially claimed to be proven in [BP10a], with implications for resource theories (e.g., the reversibility of the resource theory of entanglement [BP10b; BG15]) that were subsequently built upon it. More than a decade after its publication, an error was discovered in the proof [Ber+23], placing the result and its consequences in jeopardy. The details of this error and its impact on resource theories are discussed in [Ber+23]. After another incorrect attempt [YK24], Gao and Rahaman [GR24a] successfully proved the generalised quantum Stein's lemma, for a specific class of families $(\mathcal{S}_n)_{n\in\mathbb{N}}$, namely for

$$\mathcal{S}_n = \{E^{\otimes n}(\sigma) : \sigma \in \mathcal{S}(\mathcal{H}^{\otimes n})\} ,$$

where $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ is a HS-symmetric conditional expectation, i.e., $E = E^\dagger$ (see section 1.3.2). Their proof involved reducing this case to whether the generalised Stein's lemma holds for

$$\mathcal{S}_n = \left\{\iota_{\mathcal{K}}^{\otimes n} \otimes \sigma_n : \sigma_n \in \mathcal{S}(\mathcal{H}^{\otimes n})\right\} , \tag{1.104}$$

when tested against $\rho^{\otimes n}$ with $\rho \in \mathcal{S}(\mathcal{K} \otimes \mathcal{H})$, a scenario whose solution was known, e.g., from [Tom16, Section 6.4].
While this result covered resource theories such as coherence and the bipartite setting of eq. (1.104), it did not extend to the resource theory of entanglement, nor did it recover the original quantum Stein's lemma (as $E(\cdot) = \mathrm{Tr}[\,\cdot\,]\sigma$ for some fixed σ is not generally a HS-symmetric conditional expectation). With our work, [GR24b] we extend the results of [GR24a] to families of the form

$$\mathcal{S}_n = \{(E^\dagger)^{\otimes n}(\sigma) : \sigma \in \mathcal{S}(\mathcal{H}^{\otimes n})\}\,,$$

where $E^\dagger$ denotes the HS-adjoint of a conditional expectation E, not necessarily restricted to HS-symmetric ones, but required to satisfy $\pi = E^\dagger(\iota) > 0$ (see the equivalent conditions in theorem 1.3.2). This last constraint is necessary for the family to be viable under the axioms stated above, as otherwise the second constraint would be violated.
Although this generalisation encompassed the quantum Stein's lemma and the other settings named above, it remains again insufficient for the resource theory of entanglement, where $\mathcal{S}_n$ consists of separable states across a bipartition. In the form of the general quantum Stein's lemma (eq. (1.103) with the conditions from [BP10a]) this was only resolved after our derivation in the work [HY24] and, using different techniques, by [Lam25] as we have already mentioned. Although already discussed here extensively, we will summarise the objective in section 2.5 again and then proof the result of [GR24b] in section 3.5.

Chapter 2

Objectives

Building upon the introduction in the previous chapter, this chapter defines the objectives of the respective projects. The focus will be on the central inequalities we aimed to prove and the anticipated consequences of these proofs. While many of these objectives were alluded to in the preliminaries (sections 1.3 and 1.4), this section distils them into high level goals, whilst maintaining brevity and avoiding excessive detail. The discussions in [Blu+24] and [Aud+25] are combined in the first section, while each of the subsequent projects is presented in its own dedicated section. Notably, projects [Aud+25] and [Cap+24] began with specific objectives that evolved during the research process. Although the initial goals were not fully realised, the investigations yielded significant insights, leading to the proof of alternative or refined results. We will detail this course of change and give both the original and the changed goals in the respective section.

2.1 Objectives of [Aud+25] and [Blu+24]

The primary objective for both [Aud+25] and [Blu+24] is the improvement and, where feasible, the generalisation of existing continuity bounds. Consequently, the central inequalities in these projects are precisely these continuity bounds. For [Aud+25], the initial aim was to prove the conjectured bound eq. (1.54). However, this was not achieved in its full generality. Instead, focusing on the special case where the marginals of both states $\rho_{AB}, \rho'_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ agree ($\rho_B = \rho'_B$) led to an interesting generalisation of the sharp von Neumann entropy bound eq. (1.52). For this case of equal marginals, the conjecture simplifies to the inequality:

$$|S(\rho_{AB}) - S(\rho'_{AB})| \overset{?}{\leq} \varepsilon \log\big(d_A^2 - 1\big) + h(\varepsilon)$$

with $\varepsilon = T(\rho_{AB}, \rho'_{AB})$ and h the binary entropy. A naive application of eq. (1.52) would yield a dimension dependence of d_{AB} rather than the desired d_A. We thus reformulate the problem, seeking to establish the following inequality for $\rho, \rho' \in \mathcal{S}(\mathcal{H})$, with $\varepsilon = T(\rho, \rho')$:

$$S(\rho) - S(\rho') \overset{?}{\leq} \varepsilon(S(\mu) - S(\nu)) + h(\varepsilon) \tag{2.1}$$

where μ, ν are the states given by the JH decomposition (see eq. (1.58)). Such a result does not only address eq. (1.54) for the case of equal marginals, but also gives eq. (1.52) as a consequence, while further improving on the continuity bounds for the maps $\rho \mapsto D(\rho\|\sigma)$ and $(\rho, \sigma) \mapsto D(\rho\|\sigma)$ from [Blu+23a; Blu+23b].
The primary aim of [Blu+24] is to improve the bounds for $\alpha \in (1, \infty)$ on the SR-CE established in [MD22], which are based on a suboptimal duality relation. Furthermore, [Blu+24] aims to generalise the bounds from both [MD22] and [BG23] to encompass general SR resource measures. This means maps of the form $\rho \mapsto \widetilde{D}_\alpha(\rho\| \mathcal{C})$, where $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ is a convex, compact set containing at least one positive-definite element. Specifically, for $\rho, \rho' \in \mathcal{S}(\mathcal{H})$ with $\varepsilon = T(\rho, \rho')$, the objective is to improve and generalise [MD22] to the inequality:

$$|\widetilde{D}_\alpha(\rho\| \mathcal{C}) - \widetilde{D}_\alpha(\rho'\| \mathcal{C})| \leq \begin{cases} \log(1+\varepsilon) + \frac{1}{1-\alpha} \log\Big(1 + \varepsilon^\alpha k(\alpha)^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\Big) & \alpha \in [1/2, 1)\,, \\ \log(1+\varepsilon) + \frac{1}{\alpha-1} \log\Big(1 + \varepsilon \kappa(\alpha)^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\Big) & \alpha \in (1, \infty)\,, \end{cases} \tag{2.2}$$

and to improve and generalise [BG23] for $\alpha \in (1, \infty]$ to the inequality:

$$|\widetilde{D}_\alpha(\rho\| \mathcal{C}) - \widetilde{D}_\alpha(\rho'\| \mathcal{C})| \leq \frac{\alpha}{\alpha - 1} \log\Big(1 + \varepsilon\kappa(\alpha)^{\frac{\alpha-1}{\alpha}}\Big) \tag{2.3}$$

with $\log\kappa(\alpha) = \sup\{\widetilde{D}_\alpha(\rho\| \mathcal{C}) : \rho \in \mathcal{S}(\mathcal{H})\}$. The boundedness of the latter constant can be readily demonstrated using the existence of the positive-definite state in $\mathcal{C}$. A key aspect of generalising [BG23] is the necessity to directly prove the triangle inequality of a norm-like functional. Our approach aims to avoid reliance on the proofs in interpolation theory [Jun96; Pis18] used in [BG23], which are only applicable when $\mathcal{C} = \iota_A \otimes \mathcal{S}(\mathcal{H}_B)$. Similarly to the von Neumann entropy bound discussed earlier, subsequent goals include application of the results to derive continuity bounds, particularly for $\rho \mapsto \widetilde{D}_\alpha(\rho\|\sigma)$ and consequently $(\rho, \sigma) \mapsto \widetilde{D}_\alpha(\rho\|\sigma)$.

2.2 Objectives of [Alh+24]

The overarching goal in [Alh+24] is to achieve a MPO approximation (in ε trace-distance (TD)) of Gibbs states and marginals of Gibbs states at every positive temperature for local, translation-invariant interactions on $\mathbb{Z}$ for which the bond dimension required to approximate a $|\Lambda'|$-site marginal scales sub-polynomially with $\frac{|\Lambda'|}{\varepsilon}$. The central inequalities are the estimate of the error in a single-step recovery through the BS-CMI and its boundedness by eq. (3.18).

Complementing this, the project aims to reconstruct these Gibbs states using MPO representations based on measurements of local marginals, and to further apply the developed techniques to estimate other correlation measures of such Gibbs states. Our discussion here will primarily focus on the MPO reconstruction, briefly mentioning the measurement-based reconstruction that builds upon it, and will not delve into the other correlation measures (which are detailed in [Alh+24]).

As alluded to in section 1.4.2, one pivotal missing element for achieving this result is establishing a connection between the BS-CMI and the quantification of the recovery condition in eq. (1.73). More precisely, this involves proving the following:

$$\begin{aligned}\widehat{D}(\sigma_{AB} \otimes \iota_C\|\sigma_{ABC}) - \widehat{D}(\sigma_B \otimes \iota_C\|\sigma_{BC}) &= \widehat{I}(A;C|B)_\sigma \\ &\leq c(\sigma_{ABC}, \sigma_{BC}, \sigma_{AB}, \sigma_B)\big\|\sigma_{ABC}\sigma_{AB}^{-1}\sigma_B\sigma_{BC}^{-1} - \mathbb{1}\big\|_\infty\end{aligned} \tag{2.4}$$

with a controllable $c(\sigma_{ABC}, \sigma_{BC}, \sigma_{AB}, \sigma_B)$ in the context where σ_{ABC} is the (marginal) Gibbs state of a local, translation-invariant interaction on $\mathbb{Z}$. More generally, the objective is to find an upper bound to the BS-DPI of the form:

$$\widehat{D}(X\|Y) - \widehat{D}(E(X)\|E(Y)) \leq c(X, Y, E(X), E(Y))\big\|YX^{-1}E(X)E(Y)^{-1} - \mathbb{1}\big\|_\infty$$

for positive-definite $X, Y \in \mathcal{B}(\mathcal{H})$ and $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ a HS-symmetric conditional expectation. Equally important sub-objectives include finding a recovery map $\widehat{\mathcal{R}}^\sigma_{B\to BC}$ satisfying the inequality:

$$\Big\|\mathrm{id}_A \otimes \widehat{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB}) - \sigma_{ABC}\Big\|_1 \leq \Big\|\sigma_{AB}^{1/2}\sigma_B^{-1/2}(Z^{BC})^{1/2}\sigma_B^{1/2} - (Z^{ABC})^{1/2}\sigma_{BC}^{1/2}\Big\|_2 \tag{2.5}$$

with $Z^{ABC} = \sigma_{AB}^{-1/2}\sigma_{ABC}\sigma_{AB}^{-1/2}$ and $Z^{BC} = \sigma_B^{-1/2}\sigma_{BC}\sigma_B^{-1/2}$ which at the same time fulfils the uniform TD-bound (eq. (1.64)). The constant in that bound should only depend on the maximal size of the interval in the partition, i.e., $\max\{|\Lambda_i'| : i = 1, \ldots, N\}$, and not on the global interval size $|\Lambda'|$. Once these primary inequalities are established, the remaining objectives include extending the inequality from [BCPH22] (eq. (1.73)) to disjoint adjacent partitions of the form $\Lambda \supseteq ABC$, thereby allowing for side systems; estimating the constants appearing in eq. (2.4) and eq. (1.72); and, finally, combining the above results into the MPO reconstruction.

2.3 Objectives of [Cap+24]

The project initially aimed to prove the existence of an $\alpha(\mathcal{L}_\Lambda) = \Omega(1)$ MLSI for the Davies semigroup $\mathcal{L}_\Lambda$ derived from local and commuting interactions (as described in section 1.4.3). This scaling of the MLSI should be solely dependent on a certain decay of a specific and explicit correlation measure in the underlying Gibbs state σ^{Λ} (eq. (1.48)) at the corresponding temperature. The intention was to extend beyond the special case proven in [CLPG18] (where $\Lambda = A \sqcup B \sqcup C$) and to demonstrate generally, for $\Lambda = A \sqcup B \sqcup C \sqcup D$ and $\rho, \sigma \in \mathcal{S}(\mathcal{H}_\Lambda)$, the inequality:

$$(1 - g(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D))D_{ABC}(\rho\|\sigma) \leq D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma) \tag{2.6}$$

where the relabelling from figure 1.2 was used to reformulate also the indices of the conditional relative entropy in eq. (1.89). Although we could show that in a fully classical system (where all involved states and their possible marginals pairwise commute), one obtains:

$$g(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D) = \left\|\sigma_{ADC}\sigma_{AD}^{-1}\sigma_D\sigma_{DC}^{-1} - \mathbb{1}\right\|_\infty,$$

the general case, where no commutation relations are given, and even the scenario where only $\sigma_{ADC}, \sigma_{AD}$, σ_{DC}, and σ_D pairwise commute, remain unsolved.
Since these initial attempts were unsuccessful, the project's focus was redirected toward establishing an additive correction in the form of the following inequality:

$$D_{ABC}(\rho\|\sigma) \leq D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma) + g'(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D)\,. \tag{2.7}$$

The objective then shifted from proving a MLSI under the assumption of a specific form of correlation decay in the Gibbs state, namely the uniform decay of $g(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D)$ across all decompositions $\Lambda = A \sqcup B \sqcup C \sqcup D$ and independently of Λ. Instead, the focus turned to demonstrating that a different form of correlation decay, namely in $g'(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D)$, which is also uniform across such decompositions and independent of Λ, can, when combined with a suitable decomposition of the lattice, improve the mixing time estimates derived from uniform local gap bounds. More specifically, if a function $f : (0, \infty) \to (0, \infty)$ exists, which has $\frac{1}{f}$ homogeneous[1] such that:

$$\min\{\lambda(\mathcal{L}^\dagger_{A\subseteq\Lambda})/f(|A|) : A \subseteq \Lambda\} = \Omega(1)$$

(see eq. (1.79) for a definition of the local Lindbladians $\mathcal{L}_{A\subseteq\Lambda}$), then the naive mixing time estimate is given as follows: The case $A = \Lambda$ yields $\lambda(\mathcal{L}_\Lambda) = \Omega(f(|\Lambda|))$ which combined with eq. (1.34) and the fact that $\log\left\|\sigma^{-1}\right\| = \Theta(|\Lambda|)$ for local commuting Hamiltonians gives:

$$t_{\text{mix}}(\mathcal{L}_\Lambda; 1/2) = O\Big(\frac{\sqrt{|\Lambda|}}{f(|\Lambda|)}\Big)\,,$$

In contrast our objective is to show that a previously mentioned decay of $g'(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D)$, yields the improved:

$$t_{\text{mix}}(\mathcal{L}_\Lambda; 1/2) = O\Big(\frac{(\log|\Lambda|)^{D+1}}{f((\log|\Lambda|)^D)}\Big)\,,$$

which constitutes an exponential improvement of mixing time compared with the naive approach.
It is worth noting that in [Cap+24], f is particularly assumed to be a monomial in $x \mapsto x^{-1}$ (e.g., $x \mapsto x^{-3}$). However, this is not a strict requirement, and thus the objective here, and the subsequent results, are stated more generally. As a secondary goal, we aim to utilise the inequality eq. (2.7), with the same decay and partition assumptions, to derive a transportation cost inequality. This, in turn, would provide mixing time estimates for the normalised Wasserstein distance, as an alternative to the trace distance, an aspect of the project we will not cover here. Lastly, the project aims to demonstrate that the assumed decay of $g'(\sigma_{ADC}, \sigma_{AD}, \sigma_{DC}, \sigma_D)$ holds for marginal commuting local interactions at high temperatures.
We complement the above, with a result from an unpublished note that the author of this thesis, in collaboration with Cambyse Rouzé and Ángela Capel, achieved when attempting to elevate a mixing time estimate to a MLSI. Although this primary goal was not achieved, the author successfully derived a large-time contraction coefficient eq. (1.92) already mentioned in section 1.4.3 and a gap estimate both from the mixing time, which will be presented in the results and discussion section (section 3.3).

[1] A function $g : (0, \infty) \to (0, \infty)$ is called homogenous if there exists $k \in \mathbb{N}$ such that for all $\lambda > 0, x > 0$ $g(\lambda x) = \lambda^k g(x)$.

2.4 Objectives of [GMR24]

The central goal of [GMR24] is to establish a checkable sufficient condition under which the formal generator:

$$\mathcal{L}(X) = -i[H, X] + \sum_{j=1}^{J} \left(L_j X L_j^{\circledast} - \frac{1}{2} L_j^{\circledast} L_j, X \right), \quad X \in \mathcal{D}(\mathcal{L}) = \mathcal{T}_f(\mathcal{F}) \tag{2.8}$$

with $H = H^{\circledast} = p_0(a, a*)$ and $L_j = p_j(a, a^*)$, $p_j \in \mathbb{C}[x, y]$ polynomials in a and a^* serves as a core (on all $\mathcal{W}^k$ for $k \in \mathbb{R}_+$) for the generator of a Sobolev-preserving QMS. This sufficient condition, inspired by [ASR15], is given by the existence of a strictly monotone and divergent sequence $(k_r)_{r\in\mathbb{N}} \subset \mathbb{R}_+$ and corresponding $(\omega_r)_{r\in\mathbb{N}} \subset \mathbb{R}_+$ such that for all $r \in \mathbb{N}$ and quantum states $\rho \in \mathcal{T}_f(\mathcal{F})$

$$\operatorname{Tr}\left[\mathcal{L}(\rho)(\mathbf{N} + \mathbb{1})^{k_r}\right] \leq \omega_r \operatorname{Tr}\left[\rho(\mathbf{N} + \mathbb{1})^{k_r}\right]. \tag{2.9}$$

In this case the semigroup should further satisfy

$$\left\| e^{t\,\mathcal{L}} \right\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq e^{\omega(k)t} \tag{2.10}$$

for all $t \in \mathbb{R}_+$ with $k \mapsto \omega(k)$ derived from $(\omega_r)_{r\in\mathbb{N}}$. If, in addition, there exist $k \in \mathbb{R}_+$ $\nu > 0$, $\mu \geq 0$ and $\delta \geq 0$ such that for all $\rho \in \mathcal{T}_f(\mathcal{F})$

$$\operatorname{Tr}\left[\mathcal{L}(\rho)(\mathbf{N} + \mathbb{1})^{k}\right] \leq -\nu \operatorname{Tr}\left[\rho(\mathbf{N} + \mathbb{1})^{k+\delta}\right] + \mu\,, \tag{2.11}$$

then there should be a bounded function $t \mapsto c(t)$ derived from μ and ν such that for all $t \in \mathbb{R}_+$

$$\left\| e^{t\,\mathcal{L}} \right\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq c(t) \tag{2.12}$$

and further if $\delta > 0$, $t \mapsto f(t)$ derived from μ, ν, δ such that for $t \in (0, \infty)$

$$\left\| e^{t\,\mathcal{L}} \right\|_{1 \to \mathcal{W}^k} \leq f(t)\,. \tag{2.13}$$

Once the above results are established, the objectives extend to the discussion of examples and include a perturbation analysis of the corresponding semigroups. These examples comprise the quantum OU process, Gaussian semigroups, and semigroups implementing one-mode cat code gates. For some of these cases, we present the condition in eq. (2.10), and in certain instances also the stronger version given by eq. (2.11).

2.5 Objectives of [GR24b]

The central objective of this project is to prove the generalised quantum Stein's lemma, stated as the following limiting equality:

$$\lim_{n\to\infty} \frac{1}{n} D_h^{\varepsilon}(\rho^{\otimes n} \| \mathcal{S}_n) = \lim_{n\to\infty} \frac{1}{n} D(\rho^{\otimes n} \| \mathcal{S}_n)\,,$$

for $\rho \in \mathcal{S}(\mathcal{H})$, $\varepsilon \in (0, 1)$ and families of sets given by:

$$\mathcal{S}_n = \{(E^\dagger)^{\otimes n}(\sigma) : \sigma \in \mathcal{S}(\mathcal{H}^{\otimes n})\}\,,$$

where $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ is a conditional expectation (positive projection onto a von Neumann subalgebra, in this case $\mathcal{N}$) which further satisfies $\pi = E^\dagger(\iota) > 0$ (see theorem 1.3.2 for equivalent conditions). The key inequality underpinning this proof is:

$$\Gamma \otimes X \geq VV^*(\Gamma \otimes X)VV^* \tag{2.14}$$

for $X \in \mathcal{B}(\mathcal{H})$ being a positive-semidefinite fixed point of $E^\dagger$, $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$ an isometry, and $\Gamma \in \mathcal{B}(\mathcal{K})$ a positive-definite matrix, such that the action of $E^\dagger$ is given by $E^\dagger(\cdot) = V^*(\Gamma \otimes (\cdot))V$. The primary aim is to prove the existence of such an isometry V and positive-definite matrix $\Gamma > 0$ that define $E^\dagger(\cdot) = V^*(\Gamma \otimes (\cdot))V$, and critically, to establish the inequality eq. (2.14). As in the approach of [GR24a], proving this inequality reduces the question of the validity of our setup of the generalised quantum Stein's lemma to the specific case:

$$\mathcal{S}_n = \left\{ \gamma^{\otimes n} \otimes \sigma_n : \sigma_n \in \mathcal{S}(\mathcal{H}^{\otimes n}) \right\}, \tag{2.15}$$

where $\gamma \in \mathcal{S}(\mathcal{K})$ is a fixed positive-definite state, tested against $\rho^{\otimes n}$ for $\rho \in \mathcal{S}(\mathcal{K} \otimes \mathcal{H})$. This scenario is already known to be solvable [HT16, Chapter 7].

Chapter 3

Results and Discussion

We now turn to the chapter on results and discussion. As with the second section of the preliminaries (section 1.4) and the objectives (chapter 2), this chapter is divided into individual sections corresponding to each project, except for [Aud+25] and [Blu+24], which are again presented jointly.

3.1 Results and discussion of [Aud+25] and [Blu+24]

3.1.1 A novel bound for the von Neumann entropy and its consequences

The proof of the main result, i.e., eq. (2.1), relies on two key inequalities for the von Neumann entropy. These inequalities can be equivalently stated for quantum states or positive-semidefinite operators; we opt for the latter here to maintain consistency with [Aud+25] and to broaden their scope of application without reverting to the state formulation. Let $X, Y \in \mathcal{B}(\mathcal{H})$ be positive-semidefinite operators with $x = \mathrm{Tr}[X]$, $y = \mathrm{Tr}[Y]$, and let $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a CPTP map. Already in Winter's results on the CE [Win16], the following inequalities were used:

$$S(X+Y) \leq S(X) + S(Y)\,, \tag{3.1}$$

$$S(X+Y) \geq S(X) + S(Y) - (x+y)h\Big(\frac{x}{x+y}\Big)\,, \tag{3.2}$$

where h denotes the binary entropy function. Note that in the case when the ranges of X and Y are disjoint, i.e., $XY = YX = 0$, the first inequality becomes equality, which however, is not reflected by the bound in eq. (3.2).

We noticed that this second inequality can be refined along these lines using the monotonicity of the Holevo χ under CPTP maps, given by

$$S(X+Y) - S(X) - S(Y) \geq S(\Psi(X) + \Psi(Y)) - S(\Psi(X)) - S(\Psi(Y))\,.$$

Our improvement accounts for the degree of overlap between X and Y, with the following lemma making this notion precise:

Lemma 3.1.1 ([Aud+25, Lemma 1]) *Let $X, Y \in \mathcal{B}(\mathcal{H})$ be positive-semidefinite, $\mathcal{M} = \{|X\psi\rangle : |\psi\rangle \in \mathcal{H}\}$ the range of X, and $P_{\mathcal{M}}$ the orthogonal projection onto $\mathcal{M}$. Define the restriction $\cdot|_{\mathcal{M}} = P_{\mathcal{M}} \cdot P_{\mathcal{M}}$, and let $x = \mathrm{Tr}[X|_{\mathcal{M}}] = \mathrm{Tr}[X]$, $y' = \mathrm{Tr}[Y|_{\mathcal{M}}]$. Then,*

$$S(X+Y) \geq S(X) + S(Y) - (x+y')h\Big(\frac{x}{x+y'}\Big)\,.$$

Observe that when the ranges of X and Y are disjoint y' becomes zero, causing the correction to vanish. Combining this result with superadditivity eq. (3.1) and applying the JH decomposition, leads to the main result:

Theorem 3.1.2 ([Aud+25, Theorem 3]) *For $\rho, \rho' \in \mathcal{S}(\mathcal{H})$ with $\varepsilon = T(\rho, \rho')$, one has*

$$S(\rho) - S(\rho') \leq \varepsilon(S(\mu) - S(\nu)) + h(\varepsilon)\,,$$

where μ, ν arise from the JH decomposition of the difference of ρ and ρ' as $(\rho - \rho') = \varepsilon(\mu - \nu)$.

In conjunction with [Aud+25, Lemma 2], which states that for $\mu, \nu \in \mathcal{S}(\mathcal{H})$ with disjoint support (i.e., $\mu\nu = \nu\mu = 0$) and $\sigma \in \mathcal{S}(\mathcal{H})$ with $\ker\sigma \subseteq \ker\mu$,

$$D(\mu\|\sigma) - D(\nu\|\sigma) \leq \log(\exp(D_{\max}(\mu\|\sigma)) - 1)\,,$$

we then obtained the following consequences:

Corollary 3.1.3 ([Aud+25, Theorem 4, 5, 6])

1. *Let $\rho, \rho' \in \mathcal{S}(\mathcal{H})$ with $\varepsilon = T(\rho, \rho')$. Then:*

$$S(\rho) - S(\rho') \leq \varepsilon \log(\exp D_{\max}(\nu\|\iota) - 1) + h(\varepsilon) \leq \varepsilon \log(d_{\mathcal{H}} - 1) + h(\varepsilon)\,.$$

Furthermore, for $\sigma \in \mathcal{S}(\mathcal{H})$ with $\ker\sigma \subseteq \ker\rho \cap \ker\rho'$,

$$\begin{aligned} D(\rho\|\sigma) - D(\rho'\|\sigma) &\leq \varepsilon \log(\exp(D_{\max}(\nu\|\sigma)) - 1) + h(\varepsilon) \\ &\leq \varepsilon \log\big(\|\sigma^{-1}\| - 1\big) + h(\varepsilon)\,, \end{aligned} \tag{3.3}$$

where $\rho - \rho' = \varepsilon(\mu - \nu)$ via the JH decomposition and σ^{-1} denotes the Moore-Penrose pseudoinverse. Notably, setting $\rho' = \sigma$ yields an upper bound on the relative entropy in terms of trace distance.

2. *For $\rho_{AB}, \rho'_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ with $\rho_B = \rho'_B$ and $\varepsilon = T(\rho_{AB}, \rho'_{AB})$, one has*

$$|S(A|B)_\rho - S(A|B)_{\rho'}| \leq \varepsilon \log\big(d_A^2 - 1\big) + h(\varepsilon)\,.$$

If we combine the bound in eq. (3.3) with the fact that for $\sigma, \sigma' \in \mathcal{S}(\mathcal{H})$, with $\sigma, \sigma' > 0$ and $\delta = T(\sigma, \sigma')$,

$$D_{\max}(\sigma'\|\sigma) \leq \log\big(1 + \delta\|\sigma^{-1}\|\big)\,, \tag{3.4}$$

as shown for instance in [Aud+25, Corollary 1], we arrive at:

Corollary 3.1.4 ([Aud+25, Corollary 1]) *For $\rho, \rho', \sigma, \sigma' \in \mathcal{S}(\mathcal{H})$ with $\varepsilon = T(\rho, \rho')$, $\delta = T(\sigma, \sigma')$, and $\sigma, \sigma' > 0$, one has*

$$D(\rho\|\sigma) - D(\rho'\|\sigma') \leq \varepsilon \log\big(\|\sigma^{-1}\| - 1\big) + h(\varepsilon) + \log\big(1 + \delta\|\sigma^{-1}\|\big)\,. \tag{3.5}$$

The proof of this result follows exactly the same structure as that for the bound on $(\rho, \sigma) \mapsto \widetilde{D}_\alpha(\rho\|\sigma)$, presented later (see theorem 3.1.9).
In our discussion of the literature, we already introduced some of the previously known best bounds, but we want to summarise and complement this discussion a bit further, however not reaching the depths of the one in [Blu+24]. The continuity bounds presented in theorems 3.1.3 and 3.1.4, within their respective domains of applicability, clearly improve upon previous results, meaning they give tighter estimates. In the case of the CE, they strengthen eq. (1.53) when marginals are equal. For bounds on $\rho \mapsto D(\rho\|\sigma)$ and $(\rho, \sigma) \mapsto D(\rho\|\sigma)$, the results presented here outperform earlier bounds from [Blu+23a; Blu+23b], the latter of which included a correction term of the form $h\big(\frac{\varepsilon}{1+\varepsilon}\big)$ and featured a more complicated expression for the relative entropy bound. Furthermore, the first statement of theorem 3.1.4 yields an improvement over the sharpest known bound on the von Neumann entropy, namely eq. (1.52). All of this is accomplished by relating the various results to a single inequality involving the von Neumann entropy, thereby offering structural insight into the nature of these continuity bounds.
It is worth noting that nearly identical results to those in theorem 3.1.3 were obtained independently and concurrently using alternative techniques—specifically, newly developed integral representations for the relative entropy [HT24]—by Berta et al. in [BLT25]. However, their approach, like ours, does not yet extend to the fully general setting of unequal marginals, leaving the broader validity of eq. (1.54) as an open question.

3.1.2 Continuity bounds for entropy functionals derived from the sandwiched-Rényi divergence

For [Blu+24], we now shift our focus from the von Neumann entropy to the SR divergences, specifically considering their role as distance measures and investigate their continuity properties in the form of continuity bounds. More specifically, we looked at the SR divergence distance to $\mathcal{C}$:

$$\rho \mapsto \widetilde{D}_\alpha(\rho \| \mathcal{C})$$

with $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ being a convex, compact set containing a positive-definite element. Note that the normalisation of $\mathcal{C}$ is not strictly necessary, but we maintain it to remain consistent with the notion of a divergence as a measure between states (see section 1.3.1). Examples include the SR-CE for $\mathcal{C} = \iota_A \otimes \mathcal{S}(\mathcal{H}_B)$ (where the constant d_A cancels in the difference), the map $\rho \mapsto \widetilde{D}_\alpha(\rho\|\sigma)$ for $\mathcal{C} = \{\sigma\}$ with $\sigma > 0$, and, more generally, any resource measure constructed from SR divergences. This is because the conditions imposed on a resource family (see section 1.4.5) go beyond the ones we impose on $\mathcal{C}$.
As we specified earlier in the introduction (section 1.4.1) and the objectives section (section 2.1), obtaining the bounds follows two strategies. The first is an adaptation and generalisation of the strategy in [MD22] using sub/superadditivity and joint concavity/convexity of the functional

$$(X, Y) \mapsto \widetilde{Q}_\alpha(X, Y) \tag{3.6}$$

for $X, Y \in \mathcal{B}(\mathcal{H})$, $X, Y \geq 0$, and $\alpha \in [1/2, 1) \cup (1, \infty]$. The other strategy instead leverages monotonicity, linearity, and a triangle inequality of

$$X \mapsto \|X\|^*_{\mathcal{C},\alpha,1} \equiv \inf_{\sigma \in \mathcal{C}: \sigma > 0} \left\| \sigma^{\frac{1-\alpha}{\alpha}} X \sigma^{\frac{1-\alpha}{\alpha}} \right\|_\alpha , \tag{3.7}$$

for $X \in \mathcal{B}(\mathcal{H})$, $X \geq 0$, and $\alpha \in (1, \infty]$. Through their connection to the SR divergence distance to $\mathcal{C}$, which is given as follows:

$$\widetilde{D}_\alpha(X \| \mathcal{C}) = \frac{1}{\alpha - 1} \log \inf_{\sigma \in \mathcal{C}} \widetilde{Q}_\alpha(X \| \sigma) = \frac{\alpha}{\alpha - 1} \log \|X\|_{\mathcal{C},\alpha,1} ,$$

for $\alpha \in (1, \infty]$, and

$$\widetilde{D}_\alpha(X \| \mathcal{C}) = \frac{1}{\alpha - 1} \log \sup_{\sigma \in \mathcal{C}} \widetilde{Q}_\alpha(X \| \sigma) ,$$

for $\alpha \in [1/2, 1)$ and $X \in \mathcal{B}(\mathcal{H})$, $X \geq 0$ the above properties, together with the JH decomposition, can be used to derive continuity bounds for the map $\rho \mapsto \widetilde{D}_\alpha(\rho \| \mathcal{C})$.
We begin by listing the summary of non-trivial results for $\widetilde{Q}_\alpha$, where we note that the joint convexity and joint concavity were known previously (see e.g., [Tom16, Proposition 4.7, Theorem 4.1]) and are included here for completeness. The subadditivity (the result for $\alpha \in [1/2, 1)$) was proven in [MD22], and we have complemented it in [Blu+24] with the superadditivity (the result for $\alpha \in (1, \infty]$). For this, we used the fact that for $g : \mathbb{R}_+ \to \mathbb{R}_+$ a convex function vanishing at zero, it holds that $|||g(X + Y)||| \geq |||g(X) + g(Y)|||$ for $X, Y \geq 0$ and $|||\cdot|||$ an unitarily invariant norm [BU07, Theorem 1.2].

Lemma 3.1.5 ([Blu+24, Lemma 4.1, Lemma 4.2]) *For $X, X', Y, Y', Z \in \mathcal{B}(\mathcal{H})$ positive-semidefinite and with $\ker Z \subseteq \ker X \cap \ker X'$, $\lambda \in [0, 1]$, $X_\lambda = \lambda X + (1 - \lambda)X'$, and $Y_\lambda = \lambda Y + (1 - \lambda)Y'$, one has for $\alpha \in (1, \infty]$:*

$$\begin{aligned} \widetilde{Q}_\alpha(X_\lambda, Y_\lambda) &\leq \lambda \widetilde{Q}_\alpha(X, Y) + (1 - \lambda) \widetilde{Q}_\alpha(X', Y') , \\ \widetilde{Q}_\alpha(X, Z) + \widetilde{Q}_\alpha(X', Z) &\leq \widetilde{Q}_\alpha(X + X', Z) , \end{aligned} \tag{3.8}$$

and for $\alpha \in [1/2, 1)$:

$$\begin{aligned} \lambda \widetilde{Q}_\alpha(X, Y) + (1 - \lambda) \widetilde{Q}_\alpha(X', Y') &\leq \widetilde{Q}_\alpha(X_\lambda, Y_\lambda) , \\ \widetilde{Q}_\alpha(X + X', Y) &\leq \widetilde{Q}_\alpha(X, Y) + \widetilde{Q}_\alpha(X', Y) . \end{aligned} \tag{3.9}$$

To conclude the necessary properties of the functional in eq. (3.7), we need to make a small detour and begin by introducing a norm that will later turn out to have a dual-like relationship with eq. (3.7). We will state these results in the most general setting, mirroring the related results from interpolation theory [Jun96; Pis18; BG23].

Definition 3.1.6 ([Blu+24, Definition 4.8 and 4.10]) For $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ a convex, compact set (note again that the normalisation is not necessary) with at least one positive-definite element, $1 \leq p \leq q \leq \infty$ with corresponding Hölder conjugates $1 \leq q' \leq p' \leq \infty$ defined through $1 = \frac{1}{p} + \frac{1}{p'}$, $1 = \frac{1}{q} + \frac{1}{q'}$, and invariant $\frac{1}{r} = \frac{1}{q'} - \frac{1}{p'} = \frac{1}{p} - \frac{1}{q}$, we set

$$\|\cdot\|_{\mathcal{C},p,q} : \mathcal{B}(\mathcal{H}) \to [0,\infty), \quad X \mapsto \|X\|_{\mathcal{C},p,q} \equiv \sup_{\sigma\in\mathcal{C}} \left\|\sigma^{\frac{1}{2r}} X \sigma^{\frac{1}{2r}}\right\|_p$$

and

$$\|\cdot\|^*_{\mathcal{C},p',q'} : \mathcal{B}(\mathcal{H}) \to [0,\infty), \quad X \mapsto \|X\|^*_{\mathcal{C},p',q'} \equiv \inf_{\sigma\in\mathcal{C}:\sigma>0} \left\|\sigma^{-\frac{1}{2r}} X \sigma^{-\frac{1}{2r}}\right\|_{p'} .$$

With these definitions in place, one relatively easily shows a Hölder-type inequality (see [Blu+24, Lemma 4.12]) for these maps, hinting at their dual relation, namely: For $X, Y \in \mathcal{B}(\mathcal{H})$, $1 \leq p \leq q \leq \infty$, one has

$$|\operatorname{Tr}[XY]| \leq \|X\|_{\mathcal{C},p,q} \|Y\|^*_{\mathcal{C},p',q'} .$$

Building upon this, we can present the following collection of their properties, all of which are proven in [Blu+24]:

Lemma 3.1.7 ([Blu+24, Lemma 4.9, 4.13, 4.16, Theorem 4.14, Corollary 4.15]) *For $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ a convex, compact set with at least one positive-definite element, $1 \leq p \leq q \leq \infty$, then*

1. *$\|\cdot\|_{\mathcal{C},p,q}$ is a norm on $\mathcal{B}(\mathcal{H})$.*
2. *$\|\cdot\|^*_{\mathcal{C},p',q'}$ is homogeneous and monotonic on positive-semidefinite elements, that is, for $X, X', Z \in \mathcal{B}(\mathcal{H})$ with $X \geq X' \geq 0$, $c \in \mathbb{C}$,*

$$\|cZ\|^*_{\mathcal{C},p',q'} = |c| \|Z\|^*_{\mathcal{C},p',q'} , \quad \|X'\|^*_{\mathcal{C},p',q'} \leq \|X\|^*_{\mathcal{C},p',q'} . \tag{3.10}$$

3. *$\|\cdot\|^*_{\mathcal{C},p,q}$ is the dual to $\|\cdot\|_{\mathcal{C},p',q'}$, that is, for $X \in \mathcal{B}(\mathcal{H})$:*

$$\|X\|_{\mathcal{C},p,q} = \sup\{|\operatorname{Tr}[XY]| : \|Y\|^*_{\mathcal{C},p',q'} \leq 1\} .$$

4. *$\|\cdot\|_{\mathcal{C},p',q'}$ is the dual to $\|\cdot\|^*_{\mathcal{C},p,q}$ on positive-semidefinite elements, that is, for $X \in \mathcal{B}(\mathcal{H})$, $X \geq 0$:*

$$\|X\|^*_{\mathcal{C},p',q'} = \sup\{|\operatorname{Tr}[XY]| : \|Y\|_{\mathcal{C},p,q} \leq 1\} ,$$

   which consequently gives that for $Y \in \mathcal{B}(\mathcal{H})$, $Y \geq 0$:

$$\|X+Y\|^*_{\mathcal{C},p',q'} \leq \|X\|^*_{\mathcal{C},p',q'} + \|Y\|^*_{\mathcal{C},p',q'} . \tag{3.11}$$

The proofs of all but the last relations are relatively straightforward. The identification of $\|\cdot\|_{\mathcal{C},p,q}$ as the dual to $\|\cdot\|^*_{\mathcal{C},p',q'}$ on positive-semidefinite operators, which also allows us to immediately conclude eq. (3.11), however, is more intricate. It required the use of a rewriting of the Schatten norms as a supremum of a trace functional. Together with the already existing infimum over $\mathcal{C}$, the validity of the duality formula now hinges upon the exchange of infimum and supremum. We achieve this by identifying the functional as one covered by Lieb's concavity theorem [Lie73], yielding its convexity in the infimum variable and concavity in the supremum one, allowing for the application of Sion's minimax theorem [KW20, Theorem 2.18], hence the exchange of infimum and supremum.
With both theorem 3.1.5 and theorem 3.1.7 established, the derivation of the continuity bounds follows directly from their combination with the JH decomposition, leading to our main objective:

Theorem 3.1.8 ([Blu+24, Theorem 4.4 and 4.17]) *For $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ a convex, compact set with at least one positive-definite element, and $\rho, \rho' \in \mathcal{S}(\mathcal{H})$ with $\varepsilon = T(\rho, \rho')$, it holds that for $\alpha \in [1/2, 1)$:*

$$|\widetilde{D}_\alpha(\rho\|\, \mathcal{C}) - \widetilde{D}_\alpha(\rho'\|\, \mathcal{C})| \leq \log(1+\varepsilon) + \frac{1}{1-\alpha}\log\left(1 + \varepsilon^\alpha \kappa(\alpha)^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right),$$

and for $\alpha \in (1, \infty]$:

$$|\widetilde{D}_\alpha(\rho\|\, \mathcal{C}) - \widetilde{D}_\alpha(\rho'\|\, \mathcal{C})| \leq \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log\Big(1 + \varepsilon\kappa(\alpha)^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\Big) \\ \frac{\alpha}{\alpha-1}\log\Big(1 + \varepsilon\kappa(\alpha)^{\frac{\alpha-1}{\alpha}}\Big) \end{cases}, \tag{3.12}$$

where $\kappa(\alpha) \equiv \sup\{\widetilde{D}_\alpha(\rho\|\, \mathcal{C}) : \rho \in \mathcal{S}(\mathcal{H})\} < \infty$.

Note that the boundedness of $\kappa(\alpha)$ is a consequence of the existence of a positive-definite state in $\mathcal{C}$ and is detailed in [Blu+24, Remark 4.5]. It is now interesting to investigate the behaviour of these bounds in the respective limits $\alpha \to \infty$ and $\alpha \to 1$, while also comparing their performance against each other. The bound derived from the properties of $\widetilde{Q}_\alpha$ converges to a constant independent of ε for $\alpha \to \infty$ and is therefore no longer a continuity bound. In the limit $\alpha \to 1$, when $\widetilde{D}_\alpha(\rho\|\, \mathcal{C}) \to D(\rho\|\, \mathcal{C})$, this bound reduces to the almost sharp bound by Winter [Win16], that is

$$|D(\rho\|\, \mathcal{C}) - D(\rho'\|\, \mathcal{C})| \leq \varepsilon \log \kappa(1) + h\Big(\frac{\varepsilon}{1+\varepsilon}\Big). \tag{3.13}$$

In contrast, the bound derived using eq. (3.7) for $\alpha \to 1$ is divergent; however, it remains a continuity bound for $\alpha \to \infty$. Numerically comparing the bounds also confirmed what we could already deduce from this asymptotic behaviour, namely that the raised bound in eq. (3.12) performs well for α close to 1, while for medium and large values of α, the lowered bound is superior.
Having derived bounds that exhibit favourable behaviour in one limit from separated strategies, we further combined them to obtain a bound that remains stable in both limits, as $\alpha \to \infty$ and as $\alpha \to 1$. Although this combined bound also reduces to eq. (3.13) for $\alpha \to 1$, it yields a weaker bound than the lowered bound in eq. (3.12) for large values of α, particularly in the limit $\alpha \to \infty$. Despite its worse performance, it still might be of interest if one seeks for a bound continuously parametrised in α. Given the context of theorem 3.1.8, this combined bound for $\alpha \in (1, \infty]$ is given by

$$|\widetilde{D}_\alpha(\rho\|\, \mathcal{C}) - \widetilde{D}_\alpha(\rho'\|\, \mathcal{C})| \leq \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1 + \varepsilon\kappa(\alpha)^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{\frac{2\alpha-1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right).$$

Let us comment on the performance of theorem 3.1.8 in comparison to the previous bounds discussed in section 1.4.1 and mark again the improvements. Our bounds clearly generalise the bounds found in [MD22; BG23], while for $\alpha > 1$ restricted to the setting of SR-CE, they even improve upon the bounds presented in either [MD22] (eq. (1.56)) or [BG23] (eq. (1.57)) that we discussed in section 1.4.1. The improvement over [MD22] is achieved through the use of the new subadditivity for $\widetilde{Q}_\alpha$ for $\alpha > 1$, while the improvement over [BG23] simply stems from the use of the monotonicity in the norm-like functional (eq. (3.10)).
A discussion of the use cases, such as translating the bounds to specific resource theories, along with an application of theorem 3.1.5 to derive continuity bounds for a SR-MI, can be found in [Blu+24], but will not be discussed here. Instead, we at last want to provide a simplified derivation with an improved continuity bound for $(\rho, \sigma) \mapsto \widetilde{D}_\alpha(\rho\|\sigma)$. Using the so-called ALAFF method, this was the central result of [Blu+24, Section 6] where it was employed to analyse α-approximate quantum Markov chains. We only want to give here the stronger continuity bound, leaving out the application, as this discussion, except for replacing the respective bound, remains the same. The new approach uses just a combination of theorem 3.1.8 together with eq. (3.4) to conclude:

Corollary 3.1.9 *Let $\rho, \rho', \sigma, \sigma' \in \mathcal{S}(\mathcal{H})$, with $\sigma, \sigma' > 0$, $\delta = T(\sigma, \sigma')$, and $\varepsilon = T(\rho, \rho')$. Then, for $\alpha \in [1/2, 1)$,*

$$\widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho'\|\sigma') \leq \log(1+\varepsilon) + \frac{1}{1-\alpha}\log\left(1 + \varepsilon^\alpha \kappa^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right) + \log(1+\delta\kappa),$$

and for $\alpha \in (1,\infty)$*,*

$$\widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho'\|\sigma') \leq \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log\Big(1+\varepsilon\kappa^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\Big) + \log(1+\delta\kappa) \\ \frac{\alpha}{\alpha-1}\log\Big(1+\varepsilon\kappa^{\frac{\alpha-1}{\alpha}}\Big) + \log(1+\delta\kappa) \end{cases}$$

with $\kappa = \left\|\sigma^{-1}\right\|_\infty$*.*

Proof. Through the monotonicity of the SR divergences in the second argument (see eq. (1.15)) together with $\sigma' \leq e^{D_{\max}(\sigma'\|\sigma)}\sigma$, we obtain

$$\begin{aligned} \widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho'\|\sigma') &= \widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho'\|\sigma) + \widetilde{D}_\alpha(\rho'\|\sigma) - \widetilde{D}_\alpha(\rho'\|\sigma') \\ &\leq \widetilde{D}_\alpha(\rho\|\sigma) - \widetilde{D}_\alpha(\rho'\|\sigma) + D_{\max}(\sigma'\|\sigma)\,. \end{aligned}$$

Applying theorem 3.1.8 with $\mathcal{C} = \{\sigma\}$ to the difference and eq. (3.4) to the last term yields the claim. □

In addition to the considerably simpler proof and the improved performance, the bound for $\alpha \in [1/2, 1)$ and the raised one of the $\alpha \in (1,\infty]$ bounds reduce, albeit not to the new improved bound in theorem 3.1.4, but to the previously known best bound obtainable from eq. (3.13).

3.1.3 Open questions and future work

Several questions remain unresolved or have been raised by the two projects [Aud+25; Blu+24]. Foremost is the conjectured bound for the CE, that is, eq. (1.54), which still remains an open problem in its most general form. To gain insights into how to establish this bound, one might attempt to refine the strategies employed to derive the bounds for the SR resource measures, aiming to obtain bounds that reduce to the improved bounds for the relative entropy (e.g., theorem 3.1.4) and thereby potentially resolve the conjecture. Alternatively, one might try to further tweak the strategy that lead to theorem 3.1.2. Neither of the two strategies appears to be straightforward to the author, however.
Another interesting avenue of research would be to investigate other quantum generalisations of the classical Rényi divergences, such as the GR or PR divergences. However, adapting the strategies used in [Blu+24] to these families is not straightforward. While these quantities have analogues of $\widetilde{Q}_\alpha$ that are jointly convex or concave depending on the range of α, they both lack the super- or subadditivity of the $\widetilde{Q}_\alpha$ as established in theorem 3.1.5—a fact that we numerically tested in [Blu+24]. Addressing these quantities for continuity bounds, beyond simply using the bounds of the maximal divergence (eq. (3.4)), therefore would require some novel or altered techniques and could lead to new tools and insights.

3.2 Results and discussion of [Alh+24]

Building upon the detailed introduction in section 1.4.2 and the objectives outlined in section 2.2, we now address the remaining technical challenges on the path toward efficient MPO approximations of Gibbs states for local, translation-invariant interactions on spin chains.
Our approach proceeds through several key steps. We begin by establishing both upper and lower additive bounds on the contraction of BS entropy under HS-symmetric conditional expectations. Next, we extend eq. (1.73) to incorporate side systems and leverage this generalisation to demonstrate super-exponential decay of the BS-CMI for Gibbs states of local, translation-invariant interactions on spin chains. This analysis enables us to derive a bound on the reconstruction error incurred in a single step.
Finally, we apply these results to propose a MPO-based reconstruction scheme for marginals of such Gibbs states, achieving sub-polynomial bond dimension scaling in $\frac{|\Lambda'|}{\varepsilon}$, where $|\Lambda'|$ denotes the size of the marginal system and ε represents the reconstruction error in TD. We conclude by discussing the relevance of this reconstruction scheme for learning Gibbs states from local tomography and present the corresponding result from [Alh+24].

3.2.1 Upper and lower bounds on the BS-DPI

As already discussed in sections 1.4.2 and 2.2, a key component in the MPO approximation is the derivation of upper and lower bounds for the difference

$$\widehat{D}(X\|Y) - \widehat{D}(E(X)\|E(Y)), \quad \text{for } X, Y \in \mathcal{B}(\mathcal{H}),\ X > 0\,, Y > 0\,, \tag{3.14}$$

that is, the contractivity (DPI) of the BS entropy under a HS-symmetric conditional expectation $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$.
We begin with the lower bound, obtained by adapting a result from [BC19] (see also eq. (1.72) for the special case of the BS-CMI) to the setting of unnormalised but positive-definite operators $X, Y \in \mathcal{B}(\mathcal{H})$:

$$\left(\frac{\pi}{4}\right)^4 \frac{\left\|X^{1/2}E(X)^{-1/2}Z_E^{1/2}E(X)^{1/2} - Z^{1/2}X^{1/2}\right\|_2^4}{\|X\|_1\|Z\|_\infty^2} \leq \widehat{D}(X\|Y) - \widehat{D}(E(X)\|E(Y))\,, \tag{3.15}$$

where $Z = X^{-1/2}YX^{-1/2}$ and $Z_E = E(X)^{-1/2}E(Y)E(X)^{-1/2}$.
We now invoke the inequality, proved in [CV20, Lemma 2.2], stating that for any $V, W \in \mathcal{B}(\mathcal{H})$ with $\operatorname{Tr}[V^*V] = \operatorname{Tr}[W^*W] = 1$, we have

$$\|V^*V - W^*W\|_1 \leq 2\|V - W\|_2.$$

Applying this inequality to the left-hand side of eq. (3.15), with

$$V = \frac{X^{1/2}E(X)^{-1/2}Z_E^{1/2}E(X)^{1/2}}{\|Y\|_1} \quad \text{and} \quad W = \frac{Z^{1/2}X^{1/2}}{\|Y\|_1},$$

allows us to establish the following result. We state it here in a slightly more general form than in [Alh+24], to place it on equal footing with theorem 3.2.2.

Lemma 3.2.1 ([Alh+24, Lemma 4.1]) *Let $X, Y \in \mathcal{B}(\mathcal{H})$ be positive-definite operators, i.e., $X > 0, Y > 0$, and let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ be a HS-symmetric conditional expectation. Then:*

$$\left(\frac{\pi}{4}\right)^4 \frac{\left\|\widehat{\mathcal{R}}^{X,Y}_{\mathcal{B}(\mathcal{H})\to\mathcal{N}}(X) - Y\right\|_1^4}{\|X\|_1\|Y\|_1\|Z\|_\infty^2} \leq \widehat{D}(X\|Y) - \widehat{D}(E(X)\|E(Y))\,,$$

where $Z = X^{-1/2}YX^{-1/2}$ and

$$\widehat{\mathcal{R}}^{X,Y}_{\mathcal{B}(\mathcal{H})\to\mathcal{N}}(X) \equiv E(X)^{1/2}Z_E^{1/2}E(X)^{-1/2}XE(X)^{-1/2}Z_E^{1/2}E(X)^{1/2}, \tag{3.16}$$

with $Z_E = E(X)^{-1/2}E(Y)E(X)^{-1/2}$.

Although more intricate and structurally distinct from the (rotated) Petz recovery map, and lacking the trace-preserving property, this completely positive map $\widehat{\mathcal{R}}^{X,Y}_{\mathcal{B}(\mathcal{H})\to\mathcal{N}}$ will serve as our fundamental building block for the MPO reconstruction.
Before proceeding, however, we require an upper bound on the contractivity (DPI) of the BS entropy, that is, an upper estimate for the quantity in eq. (3.14), which will later allow us to control the error incurred in a single recovery step.
The proof of this upper bound is comparatively straightforward and relies on the well-known integral representation of the logarithm:

$$\log V = \int_0^\infty \left(\frac{1}{t+1} - \frac{1}{t+V}\right) dt\,,$$

valid for all $V \in \mathcal{B}(\mathcal{H})$ with $V > 0$. This expression, combined with the cyclicity of the trace, the defining properties of conditional expectations (section 1.3.2) and its trace preservation, yields the identity

$$\begin{aligned}&\widehat{D}(X\|Y) - \widehat{D}(E(X)\|E(Y)) \\ &= \int_0^\infty \operatorname{Tr}\left[E(X)^{1/2}\frac{1}{t+Z_E^{-1}}E(X)^{-1/2}\left(XY^{-1} - E(X)E(Y)^{-1}\right)X^{1/2}\frac{1}{t+Z^{-1}}X^{1/2}\right] dt\,,\end{aligned}$$

a result that can be found in [Alh+24, Lemma 3.1]. As before, we denote $Z = X^{-1/2}YX^{-1/2}$ and $Z_E = E(X)^{-1/2}E(Y)E(X)^{-1/2}$.
By applying Hölder's inequality and other estimates to this representation, one arrives at the following upper bound:

Theorem 3.2.2 ([Alh+24, Theorem 3.2]) *Let $X, Y \in \mathcal{B}(\mathcal{H})$ be positive-definite operators and $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ a HS-symmetric conditional expectation. Then:*

$$\begin{aligned}&\widehat{D}(X\|Y) - \widehat{D}(E(X)\|E(Y))\\ &\leq \left\|X^{-1/2}YX^{-1/2}\right\|_\infty \|X\|_1 \|E(X)\|_\infty^{1/2} \left\|E(X)^{-1}\right\|_\infty^{1/2} \left\|Y^{-1}X\right\|_\infty \left\|YX^{-1}E(X)E(Y)^{-1} - \mathbb{1}\right\|_\infty .\end{aligned}$$

We remark that the statement above contains only the part of [Alh+24, Theorem 3.2] that is relevant to our MPO-reconstruction procedure. The original theorem contains an additional inequality that is not pertinent to our present purposes and is therefore omitted here.

3.2.2 Superexponential decay of the BS-CMI on local, translation-invariant spin chains

With the previous section, we now have all necessary tools to derive an explicit upper bound on the recovery error in a single step; that is to determine an explicit ε in

$$T(\sigma_{ABC}, \mathrm{id}_A \otimes \widehat{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB})) \leq \varepsilon . \tag{3.17}$$

Here, the recovery map $\widehat{\mathcal{R}}^\sigma_{B\to BC} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{BC})$ derived from eq. (3.16) is given as

$$\widehat{\mathcal{R}}^\sigma_{B\to BC}(X) \equiv \sigma_B^{1/2}(\sigma_B^{-1/2}\sigma_{BC}\sigma_B^{-1/2})^{1/2}\sigma_B^{-1/2}(X \otimes \mathbb{1}_C)\sigma_B^{-1/2}(\sigma_B^{-1/2}\sigma_{BC}\sigma_B^{-1/2})^{1/2}\sigma_B^{1/2} \tag{3.18}$$

for $X \in \mathcal{B}(\mathcal{H}_B)$, where σ_{ABC} denotes the marginal of $\sigma = \sigma^\Lambda$, the Gibbs state on Λ for a local, translation-invariant interaction (see eq. (1.48)).
Above as well as throughout this section, we fix $A, B, C, \Lambda \Subset \mathbb{Z}$ to be intervals[1], with A, B, and C pairwise disjoint but adjacent, in accordance with their lexicographical ordering. We also require that $ABC \subseteq \Lambda$, where we use the shorthand $ABC \equiv A \sqcup B \sqcup C$.
Before we link the upper (theorem 3.2.2) and lower (theorem 3.2.1) bounds on the contraction of the BS entropy, we first establish the final missing component that allows us to conclude the super-exponential decay of the recovery error in eq. (3.17) as a function of the distance between A and C.
We previously identified this missing piece (section 2.2) as an extension of the result from [BCPH22], presented in eq. (1.73). More precisely, this involved lifting the original statement, which considered only geometries where $ABC = \Lambda$ (i.e., where A, B, and C jointly covered the entire system), to geometries where $ABC \subset \Lambda$, thereby allowing for additional intervals on either side.
We established this extension by reducing the problem to the original result (i.e., eq. (1.73)), introducing only correction factors that depend on the interaction strength J and the interaction range r. This yields the following:

Lemma 3.2.3 ([Alh+24, Corollary 2.7]) *Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}\} \to \mathcal{B}_\mathbb{Z}$ be a local, translation-invariant interaction and $\beta \in \mathbb{R}_+$. Then there exists a constant $c = \Theta(1)$ such that for intervals $A, B, C, \Lambda \Subset \mathbb{Z}$, with A, B, and C pairwise disjoint and adjacent in lexicographical order, and with $ABC \subseteq \Lambda$, the following holds:*

$$\left\|\sigma_{ABC}\sigma_{AB}^{-1}\sigma_B\sigma_{BC}^{-1} - \mathbb{1}\right\|_\infty \leq \frac{c}{(\lfloor\frac{\lfloor|B|/2\rfloor}{r}\rfloor + 1)!} .$$

The involved marginals all stem from $\sigma = \sigma^\Lambda \in \mathcal{B}(\mathcal{H}\,\Lambda)$, the local Gibbs state on Λ (see eq. (1.48)) at inverse temperature β.

[1] Note that when talking about intervals in $\mathbb{Z}$ we mean intervals in $\mathbb{R}$ intersected with $\mathbb{Z}$. Two intervals $A, B \Subset \mathbb{Z}$ are adjacent in this context if there is no integer lying in-between them.

Let us now combine the previous lemma with the estimate from theorem 3.2.2. To this end, we adopt the setting of theorem 3.2.3, choose the conditional expectation as $E(\cdot) = d_A^{-1}\,\mathbb{1}_A \otimes \mathrm{tr}_A[\cdot]$, and set $X = d_C^{-1}\sigma_{AB} \otimes \mathbb{1}_C$ and $Y = \sigma_{ABC}$, cancelling constants where appropriate. This yields:

$$\begin{aligned}\widehat{I}(A;C|B)_\sigma &= \widehat{D}\big(\sigma_{AB} \otimes \iota_C \| \sigma_{ABC}\big) - \widehat{D}\big(\sigma_B \otimes \iota_C \| \sigma_{BC}\big) \\ &\leq \big\|\sigma_{AB}^{-1/2}\sigma_{ABC}\sigma_{AB}^{-1/2}\big\|_\infty \left(\|\sigma_B\|_\infty \|\sigma_B^{-1}\|_\infty\right)^{1/2} \big\|\sigma_{ABC}^{-1}\sigma_{AB}\big\|_\infty \big\|\sigma_{ABC}\sigma_{AB}^{-1}\sigma_B\sigma_{BC}^{-1} - \mathbb{1}\big\|_\infty .\end{aligned}$$

We obtain an upper bound by applying theorem 3.2.3 to the final term on the right-hand side, and by bounding the remaining operator norms of the marginals of the Gibbs state using the results from [Alh+24, Proposition 2.2, Lemma 3.5]. While we omit the detailed derivation of these estimates from [Alh+24], we note that they all rely on Araki's expansionals [Ara69], and follow from similar arguments as in [LRB05; BCPH22].

Combining this with the lower bound of theorem 3.2.1, which in this setting reads

$$T\big(\sigma_{ABC}, \mathrm{id}_A \otimes \widehat{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB})\big) \leq \frac{2}{\pi}\big(d_C \cdot \big\|\sigma_{AB}^{-1/2}\sigma_{ABC}\sigma_{AB}^{-1/2}\big\|_\infty\big)^{1/2}\big(\widehat{I}(A;C|B)_\sigma\big)^{1/4},$$

and estimating the appearing norm as above, we arrive at the following result:

Theorem 3.2.4 ([Alh+24, Lemma 4.1, Theorem 3.6]) *Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}\} \to \mathcal{B}_\mathbb{Z}$ be a local, translation-invariant interaction, and let $\beta \in \mathbb{R}_+$. Then there exist constants $c, c', \alpha, \alpha' = \Theta(1)$ such that for intervals $A, B, C, \Lambda \Subset \mathbb{Z}$, where A, B, and C are disjoint and adjacent in lexicographical order and $ABC \subseteq \Lambda$ with the local Gibbs state $\sigma = \sigma^\Lambda$ at inverse temperature β the following inequalities hold:*

$$\widehat{I}(A;C|B)_\sigma \leq c\, e^{\alpha(|B|+|C|)}\frac{1}{\big(\lfloor\frac{\lfloor |B|/2\rfloor}{r}\rfloor + 1\big)!},$$

and consequently,

$$T(\sigma_{ABC}, \mathrm{id}_A \otimes \widehat{\mathcal{R}}^\sigma_{B\to BC}(\sigma_{AB})) \leq c'\, e^{\alpha'(|B|+|C|)}\frac{1}{\big((\lfloor\frac{\lfloor |B|/2\rfloor}{r}\rfloor + 1)!\big)^{1/4}}.$$

The recovery map is given by

$$\begin{aligned}\widehat{\mathcal{R}}^\sigma_{B\to BC} &: \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_{BC}), \\ &X \mapsto \sigma_B^{1/2}(\sigma_B^{-1/2}\sigma_{BC}\sigma_B^{-1/2})^{1/2}\sigma_B^{-1/2}(X\otimes \mathbb{1}_C)\sigma_B^{-1/2}(\sigma_B^{-1/2}\sigma_{BC}\sigma_B^{-1/2})^{1/2}\sigma_B^{1/2}.\end{aligned}$$

3.2.3 A MPO reconstruction of Gibbs states on local, translation-invariant spin chains

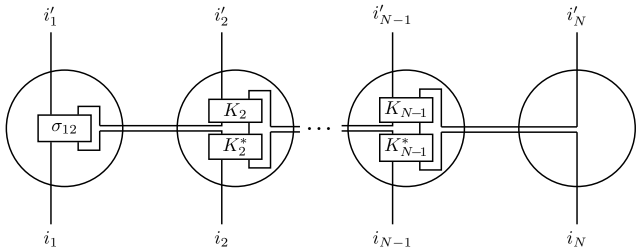
Building on the superexponential decay established for a single reconstruction step (theorem 3.2.4), we now turn to the resolution of the main objective of this project (section 2.2): the efficient MPO reconstruction of (marginals of) Gibbs states for local, translation-invariant interactions on a spin chain. To this end, we follow the approach outlined in section 1.4.2.

Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}\} \to \mathcal{B}_\mathbb{Z}$ be a local, translation-invariant interaction, and let $\beta \in \mathbb{R}_+$ denote the inverse temperature. For any finite region $\Lambda \Subset \mathbb{Z}$, let $\sigma = \sigma^\Lambda$ be the Gibbs state on Λ at inverse temperature β. Consider a subinterval $\Lambda' \subseteq \Lambda$; we may, without loss of generality, partition Λ' into disjoint adjacent intervals of equal size $2rs \in \mathbb{N}$, where $rs \in \mathbb{N}$. That is, we write $\Lambda' = \bigcup_{n=1}^N \Lambda'_n$, with $N = \frac{|\Lambda'|}{2rs}$ and define the reconstruction at step n by

$$\widehat{\mathcal{R}}_n \equiv \mathrm{id}_{\Lambda'_1\dots\Lambda'_{n-1}} \otimes \widehat{\mathcal{R}}^\sigma_{\Lambda'_n \to \Lambda'_n\Lambda'_{n+1}}.$$

The overall MPO reconstruction is then given as a composition of such maps:

$$(\bigcirc_{n=2}^{N}\widehat{\mathcal{R}}_n)(\sigma_{12})_{i_1\dots i_N}^{i'_1\dots i'_N} = \tag{3.19}$$



with $K_i = \sigma_i^{1/2}\big(\sigma_i^{-1/2}\sigma_{i:i+1}\sigma_i^{-1/2}\big)^{1/2}\sigma_i^{-1/2}$, where we use the shorthand $\sigma_i = \sigma_{\Lambda'_i}$ and $\sigma_{i:i+1} = \sigma_{\Lambda'_i\Lambda'_{i+1}}$. As already discussed after eq. (1.67) in section 1.4.2, this expression serves merely as a shorthand for the MPO in the original basis. The bond dimension, while not made explicit in the notation, is upper bounded by $b \le p_r \cdot b_r = (2rs)\cdot(2rs)^2 = (2rs)^3$, i.e., the product of the physical dimension p_r and the bond dimension b_r of the above regrouped representation.

To complete the analysis of the reconstruction error, it remains to establish the norm bound on the chain of recovery maps, i.e., the result in eq. (1.64). This estimate relies solely on the complete positivity of the recovery maps and the fact that they recover the marginal on a subsystem exactly. Hence, the proof below applies in a broader setting beyond the specific composition $\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n$, however, for consistency with [Alh+24], we state it in the specific form:

Lemma 3.2.5 ([Alh+24, Lemma 4.3]) *In the context of section 3.2.3, for $i = 2,\dots,N-1$, the following holds:*

$$T\big((\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n)(X),(\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n)(Y)\big) \le \big\|\sigma_{\Lambda'_i}^{-1}\big\|_\infty T(X,Y) \quad \forall\, X,Y\in\mathcal{B}(\mathcal{H}_{\Lambda'}),\ X,Y>0\,.$$

This generalises the bound from [Alh+24], where it is shown that

$$\Big\|(\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n)(X)\Big\|_1 \le \big\|\sigma_{\Lambda'_i}^{-1}\big\|_\infty\|X\|_1$$

for positive-semidefinite operators $X \ge 0$. When extending this to general self-adjoint operators, a naive application of the triangle inequality introduces an additional factor of 2. However, this can be avoided by decomposing any self-adjoint $X\in\mathcal{B}(\mathcal{H}_{\Lambda'})$ into its positive and non-positive parts $X = (X)_+ - (X)_-$, satisfying $\|X\|_1 = \|(X)_+\|_1 + \|(X)_-\|_1$. Hence,

$$\begin{aligned}\Big\|(\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n)(X)\Big\|_1 &\le \Big\|(\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n)((X)_+)\Big\|_1 + \Big\|(\bigcirc_{n=i}^{N-1}\widehat{\mathcal{R}}_n)((X)_-)\Big\|_1\\ &\le \big\|\sigma_{\Lambda'_i}^{-1}\big\|_\infty\left(\|(X)_+\|_1 + \|(X)_-\|_1\right) \le \big\|\sigma_{\Lambda'_i}^{-1}\big\|_\infty\|X\|_1\,.\end{aligned}$$

Returning to the overall reconstruction error, we may now apply the telescopic sum argument outlined in eq. (1.65), together with theorem 3.2.5 to obtain the final bound:

$$T\big(\sigma,(\bigcirc_{i=2}^{N}\widehat{\mathcal{R}}_i)(\sigma_{12})\big) \le \frac{|\Lambda'|}{s}\frac{\hat{c}\,e^{\hat{\alpha}s}}{\big((s+1)!\big)^{1/4}}\,,$$

where $\hat{c},\hat{\alpha} = \Theta(1)$. In this expression, we have already substituted $N = \frac{|\Lambda'|}{2rs}$, and absorbed the exponential growth in $2rs$ stemming from $\big\|\sigma_{\Lambda'_i}^{-1}\big\|_\infty$ (as derived in [Alh+24, Lemma 3.5]), as well as all constants depending on $2r$, into $\hat{c}$ and $\hat{\alpha}$. By rearranging and optimizing for s, we may now conclude with the following theorem.

Theorem 3.2.6 ([Alh+24, Corollary 4.5]) *For a local, translation-invariant interaction and inverse temperature $\beta\in\mathbb{R}_+$, there exist constants $c_1, c_2 = \Theta(1)$ such that, given intervals $\Lambda,\Lambda'\Subset\mathbb{Z}$ with $\Lambda'\subseteq\Lambda$ and a desired accuracy ε, there exists a MPO reconstruction of the marginal $\sigma_{\Lambda'}$ of the local Gibbs state $\sigma=\sigma^\Lambda$, constructed as above for the partition $\Lambda' = \bigcup_{n=1}^{N}\Lambda'_n$ with*

$$|\Lambda_n| = \left\lceil c_1\frac{\log(|\Lambda'|/\varepsilon)+c_2}{\log\log(|\Lambda'|/\varepsilon)}\right\rceil\,,$$

satisfying

$$T(\sigma_{\Lambda'}, \bigcirc_{n=2}^{N} \widehat{\mathcal{R}}_n(\sigma_{12})) \leq \varepsilon\,,$$

with bond dimension

$$b \leq \exp\left(3 \log d \cdot \left\lceil c_1 \frac{\log(|\Lambda'|/\varepsilon) + c_2}{\log\log(|\Lambda'|/\varepsilon)} \right\rceil\right).$$

In [Alh+24], this result is presented as a corollary, as it follows from a more general error bound for the same MPO reconstruction method applied to arbitrary states, not necessarily Gibbs states. Since explicit error bounds are currently only available for the Gibbs case, we present the result in this specialised form for clarity. The general version can be found in [Alh+24, Theorem 4.4].

To conclude, we summarise how the above MPO reconstruction can be employed to reconstruct the marginal of a Gibbs state from local measurements—one of the key motivations for preferring a reconstruction based on recovery maps over one directly involving the Hamiltonian (see the discussion in section 1.4.2). We do not delve into technical details here, but rather outline the central ideas.

In this scenario, the exact marginals in the recovery maps are replaced by their measured counterparts. This substitution necessitates the use of a suitable measurement algorithm and an analysis of the robustness of the MPO reconstruction to small errors in the marginals. The measurement procedure is described in [Alh+24, Lemma 4.8] and yields a classical representation of the state ρ^m satisfying

$$T(\rho, \rho^m) \leq \delta\,,$$

with probability at least $1-p$, in runtime $\mathrm{poly}(d_{\mathcal{H}}, 1/\delta)\log(1/p)$, using $\mathrm{poly}(d_{\mathcal{H}}, 1/\delta)\log(1/p)$ samples of ρ. Combining this result with the stability estimates for the recovery-based reconstruction from [Alh+24, Section 4.2] yields the following theorem:

Theorem 3.2.7 ([Alh+24, Theorem 4.9]) *For a local, translation-invariant interaction and inverse temperature $\beta \in \mathbb{R}_+$, there exist constants $c_1, c_2 = \Theta(1)$ such that, for intervals $\Lambda, \Lambda' \Subset \mathbb{Z}$ with $\Lambda' \subseteq \Lambda$ and a given accuracy ε, there exists a MPO reconstruction of the marginal $\sigma_{\Lambda'}$ of the local Gibbs state $\sigma = \sigma^{\Lambda}$, constructed as above for the partition $\Lambda' = \bigcup_{n=1}^{N} \Lambda'_n$ with*

$$|\Lambda_n| = \left\lceil c_1 \frac{\log(|\Lambda'|/\varepsilon) + c_2}{\log\log(|\Lambda'|/\varepsilon)} \right\rceil,$$

where the true marginals are replaced by measured ones. With probability at least $1-p$,

$$T(\sigma_{\Lambda'}, \bigcirc_{n=2}^{N} \widehat{\mathcal{R}}_n^m(\sigma_{12}^m)) \leq \varepsilon\,,$$

while the bond dimension of this reconstruction is bound as

$$b \leq \exp\left(3 \log d \cdot \left\lceil c_1 \frac{\log(|\Lambda'|/\varepsilon) + c_2}{\log\log(|\Lambda'|/\varepsilon)} \right\rceil\right),$$

and the number of samples and runtime required to compute this MPO are both $\mathrm{poly}(|\Lambda'|/\varepsilon)\log(1/p)$.

3.2.4 Open questions and future work

We now comment on the remaining open questions and possible future directions of research. It is worth noting that [Alh+24] already extends the main results to exponentially decaying, translation-invariant interactions. However, this extension is restricted to inverse temperatures below a certain threshold β_0, where the decay of the corresponding result to theorem 3.2.3 becomes only exponential in $\mathrm{dist}(A, C)$. Above this threshold, the decay fails entirely. Now below β_0, the exponential decay must overcome the exponential system-size dependence present in the estimates of theorem 3.2.4 and theorem 3.2.5, both of which scale with the inverse temperature. Since there appears to be little prospect for significantly improving these system-dependent factors, the most promising directions for future work become apparent.

For local interactions, a natural next step is to relax the assumption of translation invariance. This could be approached by building on techniques developed in [Ara69; BCPH22; PGPH23], with the goal of proving

a super-exponential decay result analogous to theorem 3.2.3 and subsequently applying the MPO reconstruction described in the preceding section. In contrast, for exponentially decaying interactions, it is more appropriate to focus on the rotated Petz recovery map (see eq. (1.70)) for reconstruction, as this approach avoids exponential dependence on side-systems in both the recovery error estimate and TD bounds. Following the general framework outlined in section 1.4.2 and leveraging the exponential decay bounds of [Kuw24] (cf. eq. (1.69)), one can achieve improved scaling of bond dimension with approximation error for exponentially decaying interactions at all inverse temperatures $\beta \in \mathbb{R}_+$.
To summarise: for strictly exponentially decaying interactions, MPDO reconstruction using the rotated Petz recovery yields results valid at all temperatures while improving the trade-off between bond dimension and estimation error compared to reconstruction based on the recovery map in eq. (3.18). Conversely, for local interactions, provided the translation invariance constraint can be lifted, reconstruction using eq. (3.18) should prove superior.

3.3 Results and discussion of [Cap+24]

In this section, we investigate the mixing times of the Davies semigroup for local commuting interactions, as outlined in section 1.4.3, with the objective of bootstrapping a stronger mixing time from a uniform decay of local spectral gaps and a suitable form of correlation decay in the Gibbs state (section 2.3).
Naively a uniform lower bound on the local spectral gaps of the form

$$\min\left\{\lambda(\mathcal{L}^\dagger_{A\subseteq\Lambda})/f(|A|) : A \subseteq \Lambda\right\} = \Omega(1) \tag{3.20}$$

where $f : (0,\infty) \to (0,\infty)$ is such that $1/f$ is homogeneous, yields the mixing time bound

$$t_{\mathrm{mix}}(\mathcal{L}_\Lambda; 1/2) = O\Big(\frac{\sqrt{|\Lambda|}}{f(|\Lambda|)}\Big)\,,$$

via eq. (1.34). We will now show that, under a sufficiently strong and uniform decay of a certain correlation measure in the corresponding Gibbs state, this can be improved to

$$t_{\mathrm{mix}}(\mathcal{L}_\Lambda; 1/2) = O\left(\frac{(\log|\Lambda|)^{D+1}}{f((\log|\Lambda|)^D)}\right)\,.$$

The next section introduces this measure explicitly as a correction term in the inequality of eq. (2.7). This is followed by a description of the required lattice covering. Both elements are then combined to establish the bootstrapping argument in section 3.3.3. Subsequently, we consider a family of examples—specifically marginally commuting interactions at high temperature—in which the required decay condition is satisfied. We further examine the requirement of the uniform bound on the local gap and discuss the implications of our results in the context of the relevant literature.
Finally, we derive a spectral gap estimate and a discrete-time contraction coefficient, both obtained from the mixing time, and discuss the interplay between mixing time, spectral gap, and MLSI. In light of these findings, we conclude by outlining several open questions and potential directions for future research, some of which were already anticipated in section 2.3.

3.3.1 Additive approximate tensorisation of conditional relative entropy

The proof of the additive approximate tensorisation of the conditional relative entropy (eq. (1.89)) follows directly by rearranging

$$\begin{aligned} D_{ABC}(\rho\|\sigma) - D_{AB}(\rho\|\sigma) - D_{BC}(\rho\|\sigma) \\ &= -D(\rho\|\sigma) - D(\rho_D\|\sigma_D) + D(\rho_{CD}\|\sigma_{CD}) + D(\rho_{AD}\|\sigma_{AD}) \\ &\leq -D(\rho_{ACD}\|\sigma_{ACD}) - D(\rho_D\|\sigma_D) + D(\rho_{CD}\|\sigma_{CD}) + D(\rho_{AD}\|\sigma_{AD})\,, \end{aligned}$$

where $\rho, \sigma \in \mathcal{S}(\mathcal{H}_{ABCD})$, and for the first inequality we employed the DPI of the relative entropy. Then, by applying Hölder's inequality and, once again, the DPI for the relative entropy in the form of SSA for the von Neumann entropy, we obtain the following result:

Lemma 3.3.1 ([Cap+24, Lemma B.1]) *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H}_{ABCD})$ with $\sigma > 0$. Then*

$$D_{ABC}(\rho\|\sigma) \leq D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma) + \|\mathbf{I}(A:C|D)_\sigma\|_\infty ,$$

where

$$\mathbf{I}(A:C|D)_\sigma \equiv \log\sigma_{ACD} + \log\sigma_D - \log\sigma_{AD} - \log\sigma_{CD} \tag{3.21}$$

is the MCMI. Note that we use the same name to refer to the norm $\|\mathbf{I}(A:C|D)_\sigma\|_\infty$ as well.

This inequality forms the central building block that, when combined with the lattice covering introduced in the next section and under a suitable decay condition on the MCMI, enables the bootstrapping of an improved mixing time from uniform decay of local gaps.
We note the clear structural similarity between the MCMI and the CMI; indeed, the former provides an upper bound on the latter via Hölder's inequality, which motivates both the terminology and notation. It is also worth noting that this quantity previously appeared in the (unfortunately flawed) investigation of CMI decay for Gibbs states of local interactions at high temperature in [KKBa20].

3.3.2 The partition of the lattice

This section introduces a partition of the lattice that slightly deviates from the one presented in [Cap+24, Section B.2]. That earlier construction defined all sets relative to a fixed region $\Lambda \Subset \mathbb{Z}^D$, which necessitated special treatment of boundary cases and imposed the constraint that Λ be a box of a certain side length. In contrast, the partition developed here is defined globally on $\mathbb{Z}^D$ and subsequently restricted to the region Λ via intersection. We note that both the construction in [Cap+24] and the present approach are strongly inspired by the framework of [BK18] and fundamentally rely on three parameters each:

1. The length $\ell \in \mathbb{N}$ determines the effective size of each cell in the partition along the canonical directions $e_1, \dots, e_D$ of $\mathbb{Z}^D$.
2. The overlap length $o \in \mathbb{N}$, which later allows exponential decay of the MCMI across overlapping regions, thereby enabling the decomposition of conditional relative entropies.
3. The buffer length $b \in \mathbb{N}$, which separates cells on the same level of the dimensional hierarchy and ensures the factorisation of their corresponding conditional expectations (see eq. (1.81)).

To guarantee that the decomposition yields non-empty sets at each hierarchical level and satisfies the required separation condition, we assume throughout that $\ell > 2(D+1)(b+o)$.
We begin by decomposing $\mathbb{Z}^D$ into disjoint hypercubes of side length ℓ, denoted

$$(Z^{\shortparallel}_{D,i})_{i\in\mathbb{N}} ,$$

where we chose an appropriate indexing and define for each cube two nested subsets. First, we exclude a buffer region of width b:

$$Z^{\shortmid}_{D,i} \equiv \{z \in Z^{\shortparallel}_{D,i} : \operatorname{dist}(z, \mathbb{Z}^D \setminus Z^{\shortparallel}_{D,i}) > b\}, \qquad \forall i \in \mathbb{N} .$$

Then, we further exclude the overlap region of width o:

$$Z_{D,i} \equiv \{z \in Z^{\shortmid}_{D,i} : \operatorname{dist}(z, \mathbb{Z}^D \setminus Z^{\shortmid}_{D,i}) > o\} , \qquad \forall i \in \mathbb{N} .$$

We define the corresponding aggregate sets as

$$Z^{\shortparallel}_D \equiv \bigsqcup_{i\in\mathbb{N}} Z^{\shortparallel}_{D,i} = \mathbb{Z}^D , \qquad Z^{\shortmid}_D \equiv \bigsqcup_{i\in\mathbb{N}} Z^{\shortmid}_{D,i} , \qquad Z_D \equiv \bigsqcup_{i\in\mathbb{N}} Z_{D,i} . \tag{3.22}$$

Finally, we define the residual set of dimensionally reduced structure as $Z^{D-1} \equiv \mathbb{Z}^D \setminus Z_D$. We proceed recursively to define lower-dimensional cells. For each $a = D-1, \dots, 0$, we define a partition of Z^a by identifying subsets that connect neighbouring cells from level $a+1$. Specifically, we define:

$$Z^{\shortparallel}_a \equiv \{z \in Z^a : \exists j \in \mathbb{Z},\, b \in \{1, \dots, D\} \text{ such that } z + je_b \in Z_{a+1}\} .$$

This set is then decomposed into its disjoint connected components $(Z^{\shortparallel}_{a,i})_{i\in\mathbb{N}}$ from whom we define

$$Z^{\shortmid}_{a,i} \equiv \{z \in Z^{\shortparallel}_{a,i} : \mathrm{dist}(z, Z^a \setminus Z^{\shortparallel}_{a,i}) > b\}\,, \quad Z_{a,i} \equiv \{z \in Z^{\shortmid}_{a,i} : \mathrm{dist}(z, Z^a \setminus Z^{\shortmid}_{a,i}) > o\}\,.$$

The aggregated sets are given as

$$Z^{\shortparallel}_a \equiv \bigsqcup_{i\in\mathbb{N}} Z^{\shortparallel}_{a,i}\,, \qquad Z^{\shortmid}_a \equiv \bigsqcup_{i\in\mathbb{N}} Z^{\shortmid}_{a,i}\,, \qquad Z_a \equiv \bigsqcup_{i\in\mathbb{N}} Z_{a,i}\,. \tag{3.23}$$

Note that here $Z^{\shortparallel}_a \subset Z^a$, and we set $Z^{a-1} \equiv Z^a \setminus Z_a$ as before. At the lowest level $a = 0$, we simply define the components $Z^{\shortparallel}_{0,i}$ as the disjoint connected components of Z^0, and set

$$Z_{0,i} \equiv Z^{\shortmid}_{0,i} \equiv Z^{\shortparallel}_{0,i}\,, \qquad \forall i \in \mathbb{N}\,.$$

The full decomposition of $\mathbb{Z}^D$ is denoted

$$\mathcal{Z}^D(\ell, b, o) \equiv \{Z_{a,i}, Z^{\shortmid}_{a,i}, Z^{\shortparallel}_{a,i} : a = 0, \ldots, D,\ i \in \mathbb{N}\}\,,$$

and its restriction to a finite region $\Lambda \Subset \mathbb{Z}^D$ by

$$\mathcal{Z}^D_\Lambda(\ell, b, o) \equiv \{Z \cap \Lambda : Z \in \mathcal{Z}^D(\ell, b, o)\}\,.$$

We also use Z_a, $Z^{\shortmid}_a$, $Z^{\shortparallel}_a$, and Z^a in this restricted context by intersecting them with Λ, but we do not make this explicit in the notation. While this set is finite, we will continue indexing its elements by $\mathbb{N}$ for consistency, with the understanding that only finitely many of the sets are non-empty. An illustration of the decomposition in two dimensions is provided in figure 3.1.
We summarise the key properties in the following lemma and note that, although the above construction differs slightly from that in [Cap+24], the relevant structural properties remain unchanged. As a result, we do not reproduce the proofs here, but instead refer the reader to the proof of [Cap+24, Lemma B.2].

Lemma 3.3.2 ([Cap+24, Lemma B.2]) *Let $\Lambda \subseteq \mathbb{Z}^D$, and let $\ell, b, o \in \mathbb{N}$ satisfy $\ell > 2(D+1)(b+o)$. Then:*

1. *For each $a = 0, \ldots, D$, the sets $Z_a, Z^{\shortmid}_a, Z^{\shortparallel}_a$, not necessarily restricted to Λ, can be decomposed into disjoint unions of cells as in* (3.22) *and* (3.23)*, with each cell satisfying:*

$$|Z_{a,i}| \le |Z^{\shortmid}_{a,i}| \le |Z^{\shortparallel}_{a,i}| \le \ell^D\,, \qquad \forall i \in \mathbb{N}\,.$$

2. *The decomposition into the $Z^{\shortmid}_{a,i} \in \mathcal{Z}^D_\Lambda(\ell, b, o)$ covers Λ, and each site $z \in \Lambda$ lies in at most $D+1$ of such cells:*

$$\Lambda = \bigcup_{Z^{\shortmid} \in \mathcal{Z}^D_\Lambda(\ell,b,o)} Z^{\shortmid}\,, \qquad |\{Z^{\shortmid} \in \mathcal{Z}^D_\Lambda(\ell, b, o) : z \in Z^{\shortmid}\}| \le D+1\,.$$

3. *For $a = 1, \ldots, D$, one has*

$$o < \mathrm{dist}(Z_a, Z^a \setminus Z^{\shortmid}_a)\,,$$

and for all $a = 0, \ldots, D$ and $i \ne j$, it holds that

$$2b < \mathrm{dist}(Z^{\shortmid}_{a,i}, Z^{\shortmid}_{a,j})\,,$$

whenever the sets involved are non-empty. Again this holds true for the respective sets restricted and not restricted to Λ.

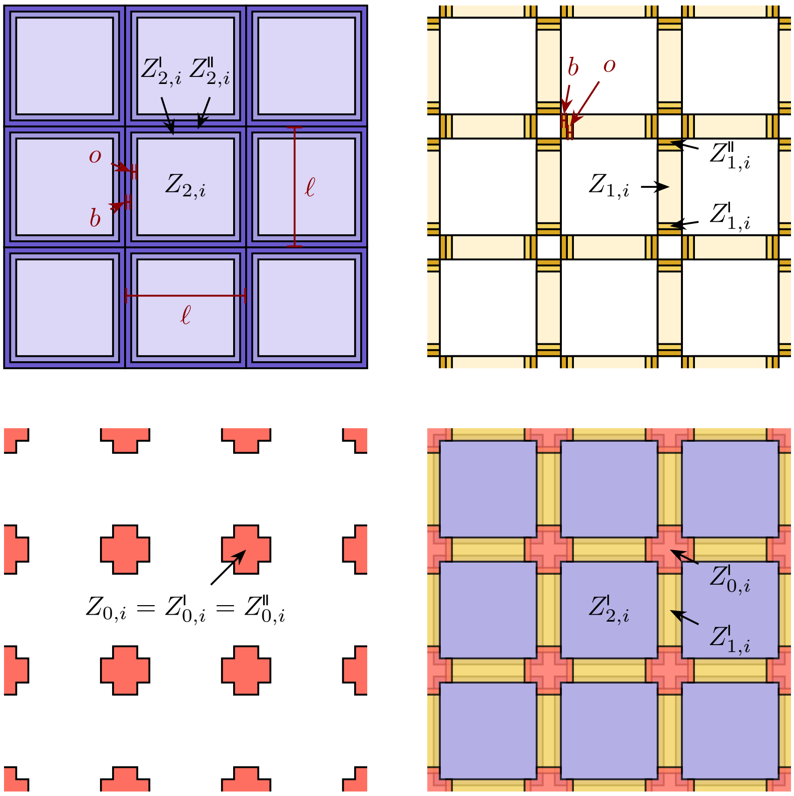


Figure 3.1: A schematic of a subset of the two-dimensional lattice ($\mathbb{Z}^2$) with its decomposition for the three hierarchical levels $a = 2, 1, 0$ and the combined decomposition into $Z^{\mathsf{I}}_2, Z^{\mathsf{I}}_1, Z^{\mathsf{I}}_0$ in the lower left corner, with the overlap visible.

3.3.3 A mixing time improvement from uniform local gap and MCMI decay

We now combine the additive approximate tensorisation with the lattice decomposition introduced in the previous section to derive the claimed improved mixing time. At a high level, the first step is to establish an additive approximate tensorisation of the relative entropy $D(\rho\|\sigma)$ into a sum of terms $D(\rho\|E_{A_i}(\rho))$, where the subsets A_i are sufficiently small to allow the application of the uniform local gap assumption. The resulting additive correction, expressed as a sum of MCMIs in the Gibbs state, can be controlled under an assumption of uniform decay, made precise in eq. (3.26). This leads to a differential inequality of the form of eq. (1.38), now including a small additive error term which, upon integration, results in a corresponding decay of the relative entropy with the same additive correction. Finally, applying Pinsker's inequality yields the desired bound on the mixing time.
We first fix the setting and proceed to prove the additive approximate tensorisation for $D(\rho\|\sigma)$. Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}$ be a local, commuting interaction and let $\Lambda \Subset \mathbb{Z}^D$ be a finite subset. We consider the corresponding Davies generator $\mathcal{L}_\Lambda$ at inverse temperature $\beta \in \mathbb{R}_+$, as defined in eq. (1.74). As discussed in section 1.4.3, the associated dynamics converge to a unique fixed point—the Gibbs state $\sigma = \sigma^\Lambda$ on Λ (eq. (1.48)) at the same inverse temperature. Hence, for any positive-definite $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$, one has: $D(\rho\|E^\dagger_\Lambda(\rho)) = D(\rho\|\sigma) = D_\Lambda(\rho\|\sigma)$, where E_Λ denotes the conditional expectation onto the kernel of $\mathcal{L}^\dagger_\Lambda$. Let $\ell, b, o \in \mathbb{N}$ be such that $b > r$, that is, the buffer exceeds the interaction range, and $\ell > 2(D+1)(b+o)$, ensuring that the hierarchical partition $\mathcal{Z}^D_\Lambda(\ell, b, o)$ satisfies theorem 3.3.2. In this setup define the following subsets (restricted to Λ):

$$A_D \equiv Z_D, \quad B_D \equiv Z^{\mathsf{I}}_D \setminus Z_D, \quad C_D \equiv Z^D \setminus Z^{\mathsf{I}}_D, \quad D_D \equiv \Lambda \setminus Z^D = \emptyset,$$

in accordance with eqs. (3.22), (3.23). Applying theorem 3.3.1, we obtain:

$$\begin{aligned} D_\Lambda(\rho\|\sigma) &= D_{A_D B_D C_D D_D}(\rho\|\sigma) \leq D_{A_D B_D}(\rho\|\sigma) + D_{B_D C_D}(\rho\|\sigma) + \|\mathbf{I}(A_D : C_D|D_D)_\sigma\|_\infty \\ &= D_{Z^!_D}(\rho\|\sigma) + D_{Z^{D-1}}(\rho\|\sigma) + \left\|\mathbf{I}(Z_D : (Z^D \setminus Z^!_D) \mid (\Lambda \setminus Z^D))_\sigma\right\|_\infty . \end{aligned}$$

We proceed by the estimate $D_{Z^!_D}(\rho\|\sigma) \leq D(\rho\|E^\dagger_{Z^!_D}(\rho))$ using eq. (1.88). From theorem 3.3.2, the decomposition

$$Z^!_D = \bigsqcup_{i\in\mathbb{N}} Z^!_{D,i}$$

satisfies $\operatorname{dist}(Z^!_{D,i}, Z^!_{D,j}) > 2b \geq 2r$ for all $i \neq j$, provided the components are non-empty. This implies the boundaries are disjoint, i.e., $Z^!_{D,i}\partial \cap Z^!_{D,j}\partial = \emptyset$, allowing us to apply eq. (1.81), yielding:

$$E^\dagger_{Z^!_D} = \bigcirc_{i\in\mathbb{N}} E^\dagger_{Z^!_{D,i}} ,$$

where we set $E_\emptyset = \mathrm{id}$. The composition is well-defined, as the constituents commute, while also only finitely many of them are non-trivial (i.e., not the identity) due to Λ being finite. Using eq. (1.21), we then obtain:

$$D(\rho\|E^\dagger_{Z^!_D}(\rho)) \leq \sum_{i\in\mathbb{N}} D(\rho\|E^\dagger_{Z^!_{D,i}}(\rho)) .$$

By iterating this argument for $D_{Z^a}(\rho\|\sigma)$, proceeding from $a = D-1$ down to $a = 0$, we arrive at the following result:

Theorem 3.3.3 ([Cap+24, Theorem B.3]) *Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}$ be a local, commuting interaction, and let $\beta \in \mathbb{R}_+$. For a finite region $\Lambda \Subset \mathbb{Z}^D$, let $\mathcal{L}_\Lambda$ be the associated Davies generator (see eq. (1.74)), and $\sigma = \sigma^\Lambda$ the corresponding Gibbs state, both at inverse temperature β. Let $\ell, b, o \in \mathbb{N}$ be such that $b > r$ and $\ell > 2(D+1)(b+o)$, and let the decomposition $\mathcal{Z}^D_\Lambda(\ell, b, o)$ as defined in section 3.3.2. Then for any positive-definite $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$, the following bound holds:*

$$D(\rho\|\sigma) \leq \sum_{a=1}^{D}\sum_{i\in\mathbb{N}} D(\rho\|E^\dagger_{Z^!_{a,i}}(\rho)) + \sum_{a=1}^{D} \left\|\mathbf{I}\left(Z_a : (Z^a \setminus Z^!_a) \mid (\Lambda \setminus Z^a)\right)_\sigma\right\|_\infty , \tag{3.24}$$

where Z_a, Z^a and $Z^!_a$ correspond to $\mathcal{Z}^D_\Lambda(\ell, b, o)$, i.e., are the restricted versions of their counterparts defined in section 3.3.2.

Given the above theorem, we now state our two assumptions. The first was explicitly formulated in section 2.3, while the second was previously only discussed implicitly. For completeness, we present both explicitly here. We assume that, for Φ and β as above, there exist constants $c_1, c_2, \xi = \Theta(1)$ and a function $f : (0,\infty) \to (0,\infty)$ such that $\frac{1}{f}$ is homogeneous of order $k \in \mathbb{N}$, such that for any $\Lambda \Subset \mathbb{Z}^D$, the following hold:

$$\min\{\lambda(\mathcal{L}_{A\subseteq\Lambda})/f(|A|) : A \subseteq \Lambda\} \geq c_1 , \tag{3.25}$$

which is a reformulation of eq. (3.20); and secondly, for all disjoint partitions $\Lambda = A \sqcup B \sqcup C \sqcup D$, where A, B, C, or D may be empty,

$$\|\mathbf{I}(A : C|D)_\sigma\|_\infty \leq c_2|\Lambda|e^{-\operatorname{dist}(A,C)/\xi} . \tag{3.26}$$

We adopt the convection that for any $F \subseteq \mathbb{Z}^D$, $\operatorname{dist}(F, \emptyset) = \infty$, so that the above bound evaluates to zero whenever A or C is empty, corresponding to the behaviour of $\mathbf{I}(A : C|D)_\sigma$ in that case.
Under these two assumptions, and in the setting of theorem 3.3.3, one can simplify eq. (3.24) by setting $o = \log(4c_2|\Lambda|)/\xi$, $b = r+1$, and $c_3 = O(1)$ such that $\ell = 2(D+2)(b+o) \leq c_3 \log|\Lambda|$. By construction, we have $\operatorname{dist}(Z_a, (Z^a \setminus Z^!_a)) > o$, hence:

$$D(\rho\|\sigma) \leq \sum_{a=1,\dots,D}\sum_{i\in\mathbb{N}} D(\rho\|E_{Z^!_{a,i}}(\rho)) + c_2|\Lambda|e^{-o/\xi} \leq \sum_{a=1,\dots,D}\sum_{i\in\mathbb{N}} D(\rho\|E_{Z^!_{a,i}}(\rho)) + \frac{1}{4} .$$

Combining now eq. (1.85) with the fact that the complete Pimsner-Popa index for the local Davies generators satisfies $\max\{2\log 10\, C_c(E_A)/|A| : A \subseteq \Lambda\} = O(1)$ (see [Cap+24, Lemma B.9]), there exists $c_4 = O(1)$ such that

$$D(\rho\|E_{Z^!_{a,i}}(\rho)) \leq \frac{c_4|Z^!_{a,i}|}{\lambda(\mathcal{L}^\dagger_{Z^!_{a,i}})}\operatorname{EP}_{\mathcal{L}_{Z^!_{a,i}}}(\rho) \leq c_1 c_4 \frac{|Z^!_{a,i}|}{f(|Z^!_{a,i}|)}\operatorname{EP}_{\mathcal{L}_{Z^!_{a,i}}}(\rho)\,,$$

where the final inequality follows from eq. (3.25). Since $1/f$ is a homogeneous function of order k, it is in particular monotone, which allows us to apply the size estimate on $Z^!_{a,i}$ from theorem 3.3.2 to obtain

$$D(\rho\|E_{Z^!_{a,i}}(\rho)) \leq c_1 c_4 \frac{\ell^D}{f(\ell^D)}\operatorname{EP}_{\mathcal{L}_{Z^!_{a,i}}}(\rho) \leq c_1 c_3^{D+k} c_4 \frac{(\log|\Lambda|)^D}{f((\log|\Lambda|)^D)}\operatorname{EP}_{\mathcal{L}_{Z^!_{a,i}}}(\rho)\,.$$

Inserting this into the previous inequality, and using that entropy production is non-negative and gauge-invariant under the exchange of fixed points (see theorem 1.3.5), and that both the local (eq. (1.79)) and global (eq. (1.74)) Davies generators are sums of single-site terms, along with the fact that each site is contained in at most $D+1$ of the sets $Z^!_{a,i}$ (see theorem 3.3.2), we find that for some $c = O(1)$,

$$D(\rho\|\sigma) \leq (D+1)c_1 c_3^{D+k} c_4 \frac{(\log|\Lambda|)^D}{f((\log|\Lambda|)^D)}\operatorname{EP}_{\mathcal{L}_\Lambda}(\rho) + \frac{1}{4} = c\frac{(\log|\Lambda|)^D}{f((\log|\Lambda|)^D)}\operatorname{EP}_{\mathcal{L}_\Lambda}(\rho) + \frac{1}{4},.$$

This yields a differential inequality analogous to eq. (1.38), since the only requirement on ρ was positive-definiteness. As such, one may take $\rho_t = e^{t\,\mathcal{L}_\Lambda}(\rho)$, which remains positive-definite for all $t > 0$ (see the discussion following eq. (1.38)). By Grönwall's inequality, this leads to the estimate

$$D(e^{t\,\mathcal{L}_\Lambda}(\rho)\|\sigma) \leq e^{-\alpha t} D(\rho\|\sigma) + \frac{1}{4}\,,$$

with $\alpha = \frac{f((\log|\Lambda|)^D)}{c(\log|\Lambda|)^D}$. Using Pinsker's inequality and the bound $\log\left\|\sigma^{-1}\right\|_\infty = O(|\Lambda|)$ (see, for example, the proof of [Cap+24, Lemma B.6]), one obtains a mixing time estimate as in eq. (1.91). We summarise this result—the main objective of the project—in the following theorem.

Theorem 3.3.4 ([Cap+24, Theorem B.10, Theorem C.2]) *Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}$ be a local, commuting interaction, and let $\beta \in \mathbb{R}_+$ be the inverse temperature. If there exists a function $f : (0,\infty) \to (0,\infty)$ such that $\frac{1}{f}$ is homogeneous, and $\xi \in O(1)$, such that for all $\Lambda \Subset \mathbb{Z}^D$*

$$\min\{\lambda(\mathcal{L}_{A\subseteq\Lambda})/f(|A|) : A \subseteq \Lambda\} = \Omega(1)\,,$$

and

$$\max\{\|\mathbf{I}(A:C|D)_\sigma\|_\infty/(|\Lambda|e^{-\operatorname{dist}(A,C)/\xi}) : A,B,C,D \subseteq \Lambda, A\sqcup B\sqcup C\sqcup D = \Lambda\} = O(1)\,,$$

then

$$t_{\mathrm{mix}}(\mathcal{L}_\Lambda;1/2) = O\Big(\frac{(\log|\Lambda|)^{D+1}}{f((\log|\Lambda|)^D)}\Big)\,,$$

where $\mathcal{L}_\Lambda$ and $\mathcal{L}_{A\subseteq\Lambda}$ denote the global (eq. (1.74)) and local (eq. (1.79)) Davies generators on Λ, respectively, and $\sigma = \sigma^\Lambda$ is the Gibbs state (eq. (1.48)) on Λ, all at the given inverse temperature β.

As previously mentioned, our paper [Cap+24] also presents a related result for a different mixing time using a normalized quantum Wasserstein distance instead of the trace distance. Although the argument follows similar reasoning, the resulting bound is less straightforward to interpret. The paper further applies both mixing time estimates to derive complexity bounds for the Gibbs sampling task. We do not elaborate on these discussions here, but instead focus on the assumptions underlying theorem 3.3.4.

First, note that for classical systems (see [Cap+24, Section A.4.2] for the translation to this setting), the decay of the MCMI is implied by the Dobrushin-Shlosman condition of strong analyticity [Cap+24, Section A.4.2], but not by the weaker condition introduced by Martinelli [Mar99]. Both conditions are commonly used in the analysis of Glauber dynamics, the classical analogue of the Davies semigroup. Strong analyticity

additionally permits a proof of the uniform local gap with $f = 1$ (e.g., following the same arguments as in [KB16], where the classical conditional expectations are given explicitly and one can derive strong clustering eq. (1.83) from strong analyticity). This in turn enables us to recover rapid mixing of Glauber dynamics for classical systems under strong analyticity, albeit with worse system-size dependence (poly-logarithmic rather than logarithmic) then the original classical proofs [Mar99].
Moreover, in our work, we establish decay of the MCMI at sufficiently high temperatures for local, marginal commuting interactions, summarised in the following example:

Example 3 ([Cap+24, Theorem D.2]). Let $\Phi : \{\Lambda : \Lambda \Subset \mathbb{Z}^D\} \to \mathcal{B}_{\mathbb{Z}^D}$ be a local, marginal commuting interaction with range r, strength J, connectivity κ, and maximum overlap g (see section 1.3.4 for an introduction to these constants). Let $\beta \in \mathbb{R}_+$ be an inverse temperature satisfying $\beta < \frac{1}{\kappa g(1+\kappa g)e^2 gJ}$. Then, for every $\Lambda \Subset \mathbb{Z}^D$ and every disjoint partition $\Lambda = A \sqcup B \sqcup C \sqcup D$, one has

$$\|\mathbf{I}(A : C|D)_\sigma\|_\infty \leq 4 \min\{|A|, |C|\} e^{-\mu \operatorname{dist}(A,C)}$$

with $\mu = \frac{1}{r} \log \frac{1}{\kappa J \beta g^2 e^2 (1+\kappa g)}$, and where $\sigma = \sigma^\Lambda$ denotes the local Gibbs state at inverse temperature β on Λ (eq. (1.48)).

However, a corresponding uniform estimate on the local gap—condition eq. (3.20)—is still missing in full generality for local, marginal commuting systems. One known result is the uniform local gap for one-dimensional, local, commuting systems established in [KB16, Theorem 28, Proposition 29] with $f = 1$. Combining this with the above MCMI-decay allows us to recover the rapid mixing result of [Koc+25] for marginal commuting interactions at high temperature in one dimension, albeit again with weaker system-size scaling ($O((\log|\Lambda|)^2)$ versus $O(\log|\Lambda|)$).
As already discussed in section 1.4.3, the aforementioned paper further extends beyond marginal commuting interactions and demonstrates rapid mixing with optimal scaling $O(\log|\Lambda|)$ for general local, commuting interactions at all temperatures in one dimension. Although we expect the MCMI to decay for spin chains at all temperatures in the local commuting case, pursuing such a result would be of limited value, since our framework would still yield inferior mixing time bounds (again $O((\log|\Lambda|)^2)$) compared to [Koc+25].
The only other non-trivial quantum setting known to the author in which a uniform local gap has been established is an unpublished result by Sebastian Stengele and collaborators, previously mentioned in the literature discussion in section 1.4.3. Their work studies the Davies semigroup corresponding to CSS-code interactions, which are local and marginally commuting by construction (see example 1). Their technique not only yields an $\Omega(1)$ uniform local gap estimate under sufficient decay of an explicit correlation measure in the Gibbs state—satisfied in the low inverse temperature regime—but also provides $\Omega(1)$ uniform local and global MLSI constants. They hence have a system-size-independent MLSI for $\mathcal{L}_\Lambda$, leading to an optimal mixing time of $O(\log|\Lambda|)$.
Such performance again outperforms the bound obtained through theorem 3.3.4, which uses their uniform estimate on the local gap with $f = 1$ (derived as a by-product of their analysis in the low inverse temperature regime) combined with the MCMI decay from example 3, resulting in a mixing time of only $O((\log|\Lambda|)^{D+1})$ for low inverse temperature.
Despite its suboptimal scaling, theorem 3.3.4 and similar strategies may still serve as a foundation for future improvements, particularly in the general setting of local, commuting interactions in higher dimensions. In this regime, a decay of the MCMI at high temperature is expected, based on the existence of a strong effective Hamiltonian. However, whether the local gap exhibits uniform decay in this setting remains open.

3.3.4 From mixing time estimates to entropy contraction coefficients and gap

Before concluding our discussion of the Davies semigroup with open problems, we turn to the interrelations between mixing time, spectral gap, and the MLSI. As discussed in section 1.3.3, for KMS-symmetric QMSs, one can derive a mixing time estimate from either the spectral gap or the MLSI. The reverse direction, however, has not yet been addressed. To limit the scope of our analysis, and because it aligns with the assumptions of our results, we henceforth restrict attention to generators $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ of QMSs that are KMS-symmetric with respect to a positive-definite state σ when deriving the spectral gap from mixing time. In contrast, for the result that derives a discrete and large-time entropy contraction from mixing time,

we require only that the semigroup be primitive, meaning that it converges to a positive-definite state σ in the infinite-time limit, without assuming KMS-symmetry.
We first assume only KMS-symmetry and show that under this condition, a mixing time estimate immediately implies a bound on the spectral gap. For this purpose, let us recall the variational definition of the spectral gap, as introduced in section 1.3.3:

$$\lambda(\mathcal{L}^\dagger) = \inf_{\substack{X \in \mathcal{B}(\mathcal{H}),\\ (\mathrm{id}-E)(X)\neq 0}} \frac{\langle X, -\mathcal{L}^\dagger(X)\rangle_{\sigma,1/2}}{\|(\mathrm{id}-E)(X)\|^2_{\sigma,1/2}}$$

for a KMS-symmetric generator and $E : \mathcal{B}(\mathcal{H}) \to \ker \mathcal{L}^\dagger = \mathcal{N}$ the conditional expectation onto the von Neumann algebra that forms its kernel. We note again that E inherits the KMS-symmetry with respect to σ from $\mathcal{L}^\dagger$.
Alternatively, one may define the spectral gap via the spectral properties of the generator or its HS-adjoint:

$$\omega(\mathcal{L}) \equiv \min\{\mathrm{Re}(\mu) : \mu \in \mathrm{Eig}(-\mathcal{L}) \setminus 0\}\,.$$

Since the spectrum of a bounded linear operator and its adjoint are complex conjugates [Bü18, Lemma 5.3.9], we have $\omega(\mathcal{L}) = \omega(\mathcal{L}^\dagger)$, and we shall write simply $\omega(\mathcal{L})$ from now on.
In the case of a KMS-symmetric generator $\mathcal{L}^\dagger$, its spectrum is entirely real, and hence $\mathcal{L}$ also has real spectrum, while both notions of gap agree, as detailed in the following lemma:

Lemma 3.3.5 *Let $\mathcal{L}^\dagger : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be the generator of a QMS in the Heisenberg picture that is KMS-symmetric with respect to a positive-definite $\sigma \in \mathcal{S}(\mathcal{H})$. Then*

$$\lambda(\mathcal{L}^\dagger) = \omega(\mathcal{L})\,.$$

Proof. By KMS-symmetry, $\mathcal{L}^\dagger$ admits an orthonormal eigenbasis $(X_i)_{i=1}^{d^2}$ with corresponding eigenvalues $(-\mu_i)_{i=1}^{d^2}$, all non-positive and ordered increasingly. Then clearly $\mu_k = \omega(\mathcal{L})$. For any X such that $(\mathrm{id}-E)(X) \neq 0$, we may write $X = \sum_{i=1}^{d^2} c_i X_i$ for $c_i \in \mathbb{C}$. Then

$$\frac{\langle X, -\mathcal{L}^\dagger(X)\rangle_{\sigma,1/2}}{\|(\mathrm{id}-E)(X)\|^2_{\sigma,1/2}} = \frac{1}{\sum_{i=1}^k |c_i|^2} \sum_{i=1}^k \mu_i |c_i|^2 = \sum_{i=1}^k p_i \mu_i\,,$$

where we have used that the components in the kernel (those with eigenvalue 0) are cancelled, and defined $(p_i)_{i=1}^k$ which forms a probability distribution. Taking the infimum over all admissible X is equivalent to taking the infimum over all such distributions, yielding the minimal value $\mu_k = \omega(\mathcal{L})$, as claimed. ☐

With this lemma, we are now ready to derive the gap estimate from the mixing time bound:

Theorem 3.3.6 *Let $\mathcal{L}^\dagger : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be the generator of a QMS in the Heisenberg picture that is further KMS-symmetric with respect to a positive-definite $\sigma \in \mathcal{S}(\mathcal{H})$. Then, for any $\varepsilon > 0$, the mixing time $t_{\mathrm{mix}}(\mathcal{L};\varepsilon)$ satisfies:*

$$\lambda(\mathcal{L}^\dagger) \geq \frac{\log(1/\varepsilon)}{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)}\,.$$

Proof. By theorem 3.3.5, we have $\lambda(\mathcal{L}^\dagger) = \omega(\mathcal{L})$. Since $\mathcal{L}$ is hermiticity-preserving and has a real spectrum (by KMS-symmetry), $\omega(\mathcal{L})$ agrees with the corresponding eigenvalue, and we may choose a self-adjoint eigenvector X corresponding to $-\omega(\mathcal{L})$. Then:

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(X) \right\|_1 = \lim_{n\to\infty} \frac{1}{n} \Big(\log(\|X\|_1) - n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\omega(\mathcal{L}) \Big) = -t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\omega(\mathcal{L})\,.$$

Since the mixing time definition applies to quantum states, choose $c \in \mathbb{R}_+$ such that $X + c\sigma > 0$, and define the normalised state $\rho = \frac{1}{\mathrm{Tr}[X]+c}(X + c\sigma) \in \mathcal{S}(\mathcal{H})$. By theorem 1.3.4:

$$\begin{aligned}\frac{1}{2(\mathrm{Tr}[X]+c)} \left\| e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(X) \right\|_1 &= \frac{1}{2(\mathrm{Tr}[X]+c)} \left\| e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(X) - E^\dagger(X) \right\|_1 \\ &= T\big(e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho), E^\dagger(\rho)\big) \leq \varepsilon^n\,,\end{aligned}$$

where we have used $E^\dagger(X) = 0$, which follows from the representation of $E^\dagger$ by the Cesàro mean (see eq. (1.27)). Thus,

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(X) \right\|_1 \leq \lim_{n\to\infty} \frac{1}{n} \log(2(\mathrm{Tr}[X] + c)\varepsilon^n) = \log(\varepsilon)\,.$$

Combining both estimates gives the desired bound:

$$\lambda(\mathcal{L}^\dagger) = \omega(\mathcal{L}) \geq \frac{\log(1/\varepsilon)}{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)}\,.$$

□

Note that such an estimate is only useful if $\varepsilon < 1$; otherwise, the lower bound becomes negative and thus trivial.

Naturally, one would hope for a similar result in the context of MLSI—namely, that a mixing time estimate might imply a MLSI, or even a CMLSI, under suitable conditions. However, to the best of our knowledge, no such result is currently known (despite the indirect route through the gap, i.e., theorem 3.3.6 combined with eq. (1.85) for GNS-symmetric semigroups). The only known implication, we managed to show, is a bound on a discrete-time contraction coefficient, where the step size is given by the mixing time. Even this result holds only in the case of a primitive semigroup. To prove it, we first require the concept of contraction coefficients for both the relative entropy and the trace distance.

To this end, let $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a CPTP map, i.e., a quantum channel. Then the contraction coefficient with respect to the TD is defined as

$$\eta_T(\Psi) \equiv \sup_{\substack{\rho,\sigma\in\mathcal{S}(\mathcal{H})\\ \rho\neq\sigma}} \frac{T(\Psi(\rho), \Psi(\sigma))}{T(\rho,\sigma)}\,,$$

and the one for the relative entropy by

$$\eta_D(\Psi) \equiv \sup_{\substack{\rho,\sigma\in\mathcal{S}(\mathcal{H})\\ \ker\sigma\subseteq\ker\rho}} \frac{D(\Psi(\rho)\|\Psi(\sigma))}{D(\rho\|\sigma)}\,.$$

Clearly, by the DPI for the both quantities, we have $\eta_T(\Psi), \eta_D(\Psi) \in [0,1]$. Moreover, the contraction coefficients were recently shown to satisfy

$$\eta_T(\Psi) \geq \eta_D(\Psi) \tag{3.27}$$

in [HT24]. With this result, establishing a relative entropy contraction coefficient for a primitive semigroup via the mixing time becomes relatively straightforward. The key step is to relate the mixing time to the contraction coefficient in trace distance. We summarise the result in the following theorem.

Theorem 3.3.7 *Let $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be the generator of a primitive QMS, meaning $E^\dagger(\cdot) = \lim_{t\to\infty} e^{t\,\mathcal{L}}(\cdot) =$ $\mathrm{Tr}[\,\cdot\,]\sigma$ for some positive-definite $\sigma \in \mathcal{S}(\mathcal{H})$. Then for all $\varepsilon > 0$,*

$$\eta_D(e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}) \leq \eta_T(e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}) \leq \varepsilon^n\,,$$

and in particular,

$$D(e^{n t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho)\|E^\dagger(\rho)) \leq \varepsilon^n D(\rho\|E^\dagger(\rho))\,,$$

for all $n \in \mathbb{N}_0$.

Proof. Applying the JH decomposition (eq. (1.58)) to $\rho, \rho' \in \mathcal{S}(\mathcal{H})$, with $\rho - \rho' = \delta(\mu - \nu)$ and $\delta = T(\rho, \rho')$, and noting that $E^\dagger(\mu) = E^\dagger(\nu) = \sigma$, we obtain

$$\begin{aligned} T(e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho), e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho')) &= \frac{\delta}{2}\left\| e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\mu) - \sigma + \sigma - e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\nu) \right\|_1 \\ &\leq \delta \max\left\{ T(e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\mu), E^\dagger(\mu)), T(e^{t_{\mathrm{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\nu), E^\dagger(\nu)) \right\} \\ &\leq \varepsilon\, T(\rho, \rho')\,. \end{aligned}$$

Since ρ and ρ' were arbitrary, this can be iterated to yield for all $n \in \mathbb{N}$,

$$T(e^{nt_{\text{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho), e^{nt_{\text{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho')) \leq \varepsilon^n T(\rho,\rho')\,.$$

Taking the supremum over $\rho \neq \rho'$ and dividing gives

$$\eta_T(e^{nt_{\text{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}) \leq \varepsilon^n\,,$$

and by eq. (3.27), it follows that

$$\eta_D(e^{nt_{\text{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}) \leq \varepsilon^n\,,$$

which immediately implies

$$D(e^{nt_{\text{mix}}(\mathcal{L};\varepsilon)\,\mathcal{L}}(\rho)\|E^\dagger(\rho)) \leq \varepsilon^n D(\rho\|E^\dagger(\rho))\,.$$

□

While the structure bears some resemblance to intermediate steps in the proof of theorem 3.3.6, we have so far been unable to extend this result to a MLSI, even under additional assumptions on the generator such as KMS- or GNS-symmetry. This does not, of course, amount to a proof of a fundamental separation between mixing time and MLSI, but it does provide increasing evidence that such a separation may exist in certain settings.

To further support this conjecture we note that in unpublished work by Ángela Capel, Barbara Roos, and Sebastian Stengele, the authors investigate the sensitivity of both mixing time and MLSI to perturbations in the generator and observe significant discrepancies. Continuing along this path, a more in-depth study of such topics—particularly the interplay between different notions of mixing—could yield deeper insights and potentially novel proof techniques for the existence of MLSI and CMLSI beyond GNS-symmetric semigroups. This direction is particularly compelling, as to the best of our knowledge, the only known approach for proving CMLSI from a spectral gap (eq. (1.85)) requires the generator to be GNS-symmetric. While this holds for Davies generators that we have seen in sections 1.4.3, 2.3 and 3.3, the new Gibbs-sampling semigroups proposed in [CKG23; DLL25b] are only KMS-symmetric and therefore currently lack the existence of positive MLSI and CMLSI.

Before summarising these questions alongside the other open problems left by this project [Cap+24], let us combine theorem 3.3.4 with theorems 3.3.6 and 3.3.7 in the following corollary, giving an interesting application of the previous results:

Corollary 3.3.8 *In the context of theorem 3.3.4, we obtain the estimate*

$$\lambda(\mathcal{L}_\Lambda^\dagger) = \Omega\left(\frac{f((\log|\Lambda|)^D)}{(\log|\Lambda|)^{D+1}}\right)\,,$$

from the mixing time bound and theorem 3.3.6, thereby improving upon the global gap estimate derived directly from assumption eq. (3.25). Moreover,

$$D(e^{nt_{\text{mix}}(\mathcal{L};1/2)\,\mathcal{L}}(\rho)\|\sigma) \leq \frac{1}{2^n} D(\rho\|\sigma)\,,$$

for all $n \in \mathbb{N}$, follows from the mixing time estimate and theorem 3.3.7.

3.3.5 Open questions and future work

Let us now turn to the remaining open questions and potential directions for future work, beginning with the most immediate one: the existence of a MLSI (or even a CMLSI) for Davies semigroups with commuting local interactions, under the assumption of a suitable decay of an explicit correlation measure in the Gibbs state. This was, in fact, the original aim of [Cap+24] (see section 2.3). Our result merely enables the lifting of a uniform bound on local gaps to a mixing time estimate under sufficient decay of the MCMI. However, the first assumption always appears in the literature as a by-product of alternative strategies that directly prove mixing times, with our approach yielding strictly weaker results in all known cases. Thus, at least for the Davies semigroup, this strategy appears poorly suited. To achieve the aforementioned goal, one should therefore depart from our method and consider one of the following two strategies. The first would yield even MLSI constants for the local generators $\mathcal{L}_{A\subseteq\Lambda}$, while the second—though only providing a MLSI for the global generator $\mathcal{L}_\Lambda$—may be extendable to the KMS-symmetric generators studied in [CKG23; DLL25b], as already mentioned in section 1.4.3.

1. The first approach builds on work by Sebastian Stengele and collaborators, who obtained explicit expressions for the local Davies conditional expectations E_A in the case of CSS-code interactions, and proved positivity inequalities of the form

$$(1-\varepsilon)E_A E_B \leq E_{A\cup B} \leq (1+\varepsilon)E_A E_B$$

with ε quantifying correlations in the Gibbs state. To generalise this to arbitrary commuting local interactions, one could begin by proving the inequality with $\varepsilon = 0$ in the infinite-temperature setting, assuming sufficient overlap between A and B. From there, one would analyse how the conditional expectations depend on the respective base states $\pi_{A\cup B} = E_{A\cup B}(\mathbb{1}/d)$, $\pi_A = E_A(\mathbb{1}/d)$, and $\pi_B = E_B(\mathbb{1}/d)$, using their structural properties (see theorem 1.3.1) to control the correlation parameter ε. While these base states are defined implicitly for general commuting interactions—as infinite limits of Petz recovery maps applied to $\mathbb{1}/d$ (cf. [Cap+24, Eq. (33)])—in the marginal commuting case, they simplify to $d^{-1}_{\Lambda\setminus A\cup B}\frac{\sigma}{\sigma_{\Lambda\setminus A\cup B}}$, $d^{-1}_{\Lambda\setminus A}\frac{\sigma}{\sigma_{\Lambda\setminus A}}$, and $d^{-1}_{\Lambda\setminus B}\frac{\sigma}{\sigma_{\Lambda\setminus B}}$ respectively, hence should allow for a condition that is easily interpretable.

2. The second strategy involves proving eq. (1.89) or equivalently its relabelled (see figure 1.2) variant eq. (2.6) with an explicit correlation condition on the Gibbs state. While this would only yield a MLSI for $\mathcal{L}_\Lambda$, the result could likely be extended to the generators introduced in [CKG23; DLL25b], since their conditional expectations E^C_A (more precisely, their HS-adjoints) satisfy the same key properties:

$$D_\Lambda(\rho\|\sigma) = D(\rho\|E^{C,\dagger}_\Lambda(\rho)) \quad \text{and} \quad D_A(\rho\|\sigma) \leq D(\rho\|E^{C,\dagger}_A(\rho))\,.$$

The only missing component in that case would be positive CMLSI constants for KMS-symmetric generators.

Of course, either of these strategies leads naturally to further questions should they succeed. For example, one would need to investigate how the resulting correlation measure in the Gibbs state relates to existing ones, particularly to classical correlation conditions. These considerations, however, become relevant only after establishing one of the two proposed results.
Closely tied to the second strategy, though independently interesting, is the question of extending the approach to generators stabilising Gibbs states of non-commutative local interactions, as studied in [CKG23; DLL25b]. For these generators, it is particularly important to examine whether a CMLSI can be inferred from a spectral gap or a mixing time estimate, and more broadly to understand the interplay between these functional inequalities in the setting where the generator is merely KMS-symmetric. In the preceding section, we only briefly touched upon these issues, and a more thorough literature review and conceptual understanding remain to be developed. With that, we conclude our discussion of Davies dynamics. However, we continue our exploration of semigroups in the following section, now in the context of Bosonic systems, where we present the results of [GMR24].

3.4 Results and discussion of [GMR24]

In this section, we pursue the objectives outlined in section 2.4, building on the introduction in section 1.4.4. Specifically, we demonstrate that a formal generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ in GKSL form, with components that are polynomials in a and a^* (see eq. (1.94)), serves as a core for the Sobolev-preserving QMS on all $\mathcal{W}^k$ for $k \in \mathbb{R}_+$, provided eq. (3.37) holds.
Following this discussion of generation theory and an initial bound on the operator norm of the semigroups on $\mathcal{W}^k$ in section 3.4.1, we build upon this result in section 3.4.2. There we discuss how to improve the exponential time dependence of this bound, given the stronger condition of eq. (2.11). We also present the

first proof of a $1 \to \mathcal{W}^k$ estimate for the semigroup when $\delta > 0$. This result extends beyond the findings of [GMR24] and will therefore be proven in detail.
In section 3.4.3, we discuss the examples introduced in section 1.4.4 and provide basic perturbative results. As in the other projects, we conclude with a discussion of future directions and open questions.

3.4.1 Generation of Sobolev-preserving semigroups

To address the first objective—namely, proving the generation result stated in section 2.4 under the assumption of eq. (2.9)—we follow, at least in part, the strategy developed by Davies (see [Dav77], or our summary in section 1.3.3). A key point of divergence from Davies' approach arises immediately. In his work, he selects a domain $\mathcal{D}(\mathcal{L})$ specifically adapted to the operator G, which he assumes to be the generator of a C_0-semigroup on $\mathcal{F}$. Under this assumption, Davies proceeds to show that the induced operator $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$ forms the core of a contractive C_0-semigroup on $\mathcal{T}(\mathcal{F})$. We rewrite eq. (2.8) as

$$\mathcal{L}(X) = G\rho + \rho G^{\circledast} + \sum_{j=1}^{J} L_j \rho L_j^{\circledast} \equiv \mathcal{G}(\rho) + \Sigma(\rho), \qquad X \in \mathcal{T}_f(\mathcal{F}) = \mathcal{D}(\mathcal{L})\,,$$

with

$$G \equiv -iH - \frac{1}{2}\sum_{j=1}^{J} L_j^{\circledast} L_j\,, \qquad \mathcal{D}(G) = \operatorname{span}\{|n\rangle : n \in \mathbb{N}_0\} \subset \mathcal{F} \tag{3.28}$$

and

$$\mathcal{G}(X) \equiv GX + XG^{\circledast}\,, \quad \Sigma(X) \equiv \sum_{j=1}^{J} L_j X L_j^{\circledast}, \qquad X \in \mathcal{T}_f(\mathcal{F}) = \mathcal{D}(\mathcal{G}) = \mathcal{D}(\Sigma)\,. \tag{3.29}$$

Since we do not have a corresponding result at our disposal, we must establish it ourselves. We first show that

$$G_\varepsilon \equiv -\varepsilon(\mathbf{N}+\mathbb{1})^{4d} + G, \quad \mathcal{D}(G_\varepsilon) = \mathcal{D}(G)$$

for all $\varepsilon > 0$, where $d \equiv \max\{\deg(p_j) : j = 0, \ldots, J\}$ is the maximal degree among the polynomials defining eq. (2.8), serves as the core of a contractive C_0-semigroup on $\mathcal{F}$.
The proof of this initial result relies on the simple observation that $-\varepsilon(\mathbf{N}+\mathbb{1})^{4d}$, defined on $\operatorname{span}\{|n\rangle : n \in \mathbb{N}_0\}$, is the core of a closed self-adjoint operator which also relatively bounds G. This directly implies the closability of G_ε and its adjoint, allowing us to apply a variant of the Lumer-Phillips theorem [LP61]: By [Nag00, Corollary 3.17], it is sufficient that both the operator and its adjoint are dissipative, a property that follows directly from the cores $(G_\varepsilon, \mathcal{D}(G_\varepsilon))$ and $(G_\varepsilon^{\circledast}, \mathcal{D}(G_\varepsilon))$, which satisfy this condition.
Following Davies arguments this contractive C_0-semigroup $(e^{t\overline{G}_\varepsilon})_{t\in\mathbb{R}_+}$ can then be lifted to a contractive C_0-semigroup on $\mathcal{T}(\mathcal{F})$ by setting

$$\big(\mathcal{P}_t^\varepsilon(\cdot) \equiv e^{t\overline{G}_\varepsilon} \cdot (e^{t\overline{G}_\varepsilon})^*\big)_{t\in\mathbb{R}_+}\,,$$

which is CP by construction. Disregarding minor technical subtleties, that are dealt with in [GMR24], one can then show that the following operator forms the core of the generator of this semigroup:

Lemma 3.4.1 ([GMR24, Lemma 3.2]) *For $\varepsilon > 0$ and $d \equiv \max\{\deg(p_j) : j = 0, \ldots, J\}$, where the polynomials $p_j \in \mathbb{C}[x, y]$ define the operators in eq. (2.8), the operator*

$$\mathcal{G}_\varepsilon(X) \equiv \mathcal{G}(X) - \varepsilon\{(\mathbf{N}+\mathbb{1})^{4d}, X\}\,, \qquad X \in \mathcal{D}(\mathcal{G}_\varepsilon) \equiv \mathcal{T}_f(\mathcal{F})\,, \tag{3.30}$$

with $\mathcal{G}$ from eq. (3.29), forms the core of a contractive, CP C_0-semigroup on $\mathcal{T}(\mathcal{F})$.

Although interesting in its own right, the primary purpose of this result is to ensure that, for $\lambda > 0$, the resolvent associated with the closure of $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$ is CP. This fact follows from the complete positivity of the semigroup, using the representation of the resolvent as an integral over the semigroup (see [Nag00, Theorem II.1.10 (i)]). This result then plays a crucial role in lifting the semigroup from $\mathcal{T}(\mathcal{F})$ to the QSS, while also ensuring that $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$ remains a core to the generator on these spaces as well. The argument

will build upon the Lumer-Phillips generation theorem (see section 1.3.3) and we will summarise it here in a streamlined and slightly modified form compared to the original [GMR24, Lemma 3.3].
Note first that, due to the compact embedding of QSS and theorem 1.4.2, it suffices to consider only $k = k_r$ for $r \in \mathbb{N}$, and correspondingly set $\omega = \omega_r$. Once the result has been established for all $r \in \mathbb{N}$, one can apply theorem 1.4.2 to interpolate the semigroups and deduce both strong continuity and the core property of $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$, using the compact embedding (see, for example, [Nag00, Proposition 5.3] for an argument of the strong continuity, while the core property naturally follows from the compact embedding).
With the above nomenclature, recall that the Lumer-Phillips theorem requires verification of the following:

1. The operator $(\mathcal{G}_\varepsilon - \omega \,\mathrm{id})$, defined on $\mathcal{D}(\mathcal{G}_\varepsilon)$, is dissipative; that is, for all $X \in \mathcal{D}(\mathcal{G}_\varepsilon)$ and $\lambda > 0$,

$$\lambda \|X\|_{\mathcal{W}^k} \leq \|(\lambda \,\mathrm{id} - (\mathcal{G}_\varepsilon - \omega \,\mathrm{id}))(X)\|_{\mathcal{W}^k} .$$

2. The range of $(\lambda_0 \,\mathrm{id} - \mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$ is dense in $\mathcal{W}^k$ for some $\lambda_0 > \omega$.

The second condition is used in the proof of the first, thus we briefly outline its derivation. The crucial observation is that the operator $(-\varepsilon\{(\mathbf{N} + \mathbb{1})^{4d}, \cdot\}, \mathcal{D}(\mathcal{G}_\varepsilon))$, due to its simple structure, already satisfies the Lumer-Phillips criteria, particularly dissipativity. Given that $\mathcal{G}$ is relatively bounded with respect to this operator on $\mathcal{W}^k$, one can conclude that for sufficiently large $\lambda_0 > \omega$ (depending on ε), the range of $(\lambda_0 \,\mathrm{id} - \mathcal{G}_\varepsilon)$ is dense, and the resolvent is bounded on this range. Hence, $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$ is a core of the closed operator $(\overline{\mathcal{G}}_\varepsilon, \mathcal{D}(\overline{\mathcal{G}}_\varepsilon))$, whose resolvent at λ_0 coincides with the restriction of the semigroup resolvent from theorem 3.4.1. From eq. (2.9), we deduce that for positive-semidefinite $X \in \mathcal{D}(\mathcal{G}_\varepsilon)$,

$$(\lambda_0 - \omega)\|X\|_{\mathcal{W}^k} \leq \|((\lambda_0 - \omega) \,\mathrm{id} - (\mathcal{G}_\varepsilon - \omega \,\mathrm{id}))(X)\|_{\mathcal{W}^k} , \tag{3.31}$$

which by the core property can be extended to all positive-semidefinite $X \in \mathcal{D}(\overline{\mathcal{G}}_\varepsilon)$:

$$(\lambda_0 - \omega)\|X\|_{\mathcal{W}^k} \leq \left\|((\lambda_0 - \omega) \,\mathrm{id} - (\overline{\mathcal{G}}_\varepsilon - \omega \,\mathrm{id}))(X)\right\|_{\mathcal{W}^k} . \tag{3.32}$$

As the resolvent $R(\lambda_0, \overline{\mathcal{G}}_\varepsilon)$, inherits its CP property from the resolvent of the semigroup in theorem 3.4.1, this inequality can be extended first to self-adjoint $X \in \mathcal{D}(\overline{\mathcal{G}}_\varepsilon)$ (see [GMR24, Lemma 3.3]) and then to all $X \in \mathcal{D}(\overline{\mathcal{G}}_\varepsilon)$, following the ideas of [Mö25, Proof of Theorem 4.1.1]. In particular, this yields inequality (3.31) and (3.32) for general $X \in \mathcal{D}(\mathcal{G}_\varepsilon)$ and $X \in \mathcal{D}(\overline{\mathcal{G}}_\varepsilon)$ respectively, and therefore

$$\left\|R(\lambda_0, \overline{\mathcal{G}}_\varepsilon)\right\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq \frac{1}{\lambda_0 - \omega} . \tag{3.33}$$

Since the resolvent set of a closed operator is open and eq. (3.33) holds, we can conclude that the resolvent $R(\lambda, \overline{\mathcal{G}}_\varepsilon)$ exists for all $\lambda \in (\lambda_0 - \delta, \lambda_0 + \delta)$, with $\delta = \frac{1}{\lambda_0 - \omega}$ (see [Nag00, Proposition 1.3 (i)]). Repeating the above argument for these values of λ, we again obtain complete positivity and, using the core property and eq. (2.9), we establish:

$$(\lambda - \omega)\|X\|_{\mathcal{W}^k} \leq \|((\lambda - \omega) \,\mathrm{id} - (\mathcal{G}_\varepsilon - \omega \,\mathrm{id}))(X)\|_{\mathcal{W}^k} \tag{3.34}$$

for all $X \in \mathcal{D}(\mathcal{G}_\varepsilon)$, and

$$\left\|R(\lambda, \overline{\mathcal{G}}_\varepsilon)\right\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq \frac{1}{\lambda - \omega} .$$

By iterating this argument, we find that the resolvent set of $(\overline{\mathcal{G}}_\varepsilon, \mathcal{D}(\overline{\mathcal{G}}_\varepsilon))$ includes the entire interval (ω, ∞), thus yielding (3.34) for all $\lambda \in (\omega, \infty)$. This completes the verification of the missing first and thereby all Lumer-Phillips conditions. As a result, we can complement theorem 3.4.1 and obtain:

Lemma 3.4.2 ([GMR24, Lemma 3.3]) *Let $\varepsilon > 0$, and let $d \equiv \max\{\deg(p_j) : j = 0, \dots, J\}$, where the polynomials $p_j \in \mathbb{C}[x, y]$ define the operators in eq. (2.8). Then $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$ from eq. (3.30) is the core to a generator of a CP, C_0-semigroup semigroup $(e^{t\overline{\mathcal{G}}_\varepsilon})_{t \in \mathbb{R}_+}$ on all QSSs. For $k \in \mathbb{R}_+$ the semigroup satisfies*

$$\left\|e^{t\overline{\mathcal{G}}_\varepsilon}\right\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq e^{\omega(k)t} ,$$

where $\omega(k) = \frac{k_{r'} - k}{k_{r'} - k_r}\omega_r + \frac{k - k_r}{k_{r'} - k_r}\omega_{r'}$, with r and r' chosen such that k_r is the greatest lower bound and $k_{r'}$ the least upper bound of k, or $\omega(k) = \omega_{k_r}$ if $k = k_r$.

The next step is to remove the perturbation and establish the same result—theorem 3.4.2—for $\varepsilon = 0$. While this resembles the minimal semigroup problem (see [Dav77; CF98; Fag18] and our discussion in section 1.3.3), theorem 3.4.2 provides us with a considerably stronger statement. Not only is the semigroup under consideration Sobolev-preserving, but its generator is also bounded as an operator between different QSSs. That is, there exists $d' > 0$ such that for every $k \in \mathbb{R}_+$, there exists $c_k > 0$ satisfying

$$\|\mathcal{G}_\varepsilon\|_{\mathcal{W}^{k+d'} \to \mathcal{W}^k} \leq c_k \,.$$

These two properties together allow us to verify the assumptions of the second Trotter-Kato approximation theorem [Nag00, Theorem III.4.9], which states the following: Let $(e^{t\,\mathcal{O}_n} : \mathcal{X} \to \mathcal{X})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, be a sequence of C_0-semigroups on the Banach space $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ with generators $(\mathcal{O}_n, \mathcal{D}(\mathcal{O}_n))$, and assume there exists a uniform bound

$$\|e^{t\,\mathcal{O}_n}\|_{\mathcal{B}(\mathcal{X})} \leq M e^{\omega t} \,,$$

for all $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. If there exists $\lambda_0 > \omega$ and a densely defined operator $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ such that $\mathcal{O}_n(X) \to \mathcal{O}(X)$ for all $X \in \mathcal{D}(\mathcal{O})$, and the range of $(\lambda_0 - \mathcal{O}, \mathcal{D}(\mathcal{O}))$ is dense in $\mathcal{X}$, then the semigroups $(e^{t\,\mathcal{O}_n} : \mathcal{X} \to \mathcal{X})_{t \in \mathbb{R}_+}$ converge strongly and uniformly on compact time intervals to a strongly continuous semigroup $(e^{t\overline{\mathcal{O}}})_{t \in \mathbb{R}_+}$, where $(\mathcal{O}, \mathcal{D}(\mathcal{O}))$ is a core for the generator $(\overline{\mathcal{O}}, \mathcal{D}(\overline{\mathcal{O}}))$.
Note that, with this, we are already anticipating the resolution of the minimal semigroup problem discussed in [Dav77; CF98; Fag18] and in section 1.3.3. This also explains why we state a more general result that is not specific to $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$, and why we refer to theorem 3.4.3 as the approximation lemma in [GMR24]. The proof strategy follows exactly as outlined above: we use the Sobolev-preserving property along with the boundedness of the generator as a map between QSSs, which then yields:

Lemma 3.4.3 ([GMR24, Lemma E.5]) *Let $J \in \mathbb{N}$, and for each $j = 1, \ldots, J$, let $p_{j,r}, p_{j,l} \in \mathbb{C}[x, y]$ be polynomials and $(c_{j,n})_{n \in \mathbb{N}} \subset \mathbb{C}$ convergent sequences with limits $c_j \in \mathbb{C}$. Assume*

$$\mathcal{O}_n(X) \equiv \sum_{j=1}^{J} c_{j,n}\, p_{j,l}(a, a^*) X p_{j,r}(a, a^*) \,, \qquad X \in \mathcal{D}(\mathcal{O}_n) \equiv \mathcal{T}_f(\mathcal{F})$$

are cores of generators of Sobolev-preserving semigroups (see theorem 1.4.3) $(e^{t\overline{\mathcal{O}}_n})_{t \in \mathbb{R}_+}$ on all QSSs (theorem 1.4.1), uniformly bounded in $n \in \mathbb{N}$. That is, for each $k \in \mathbb{R}_+$, there exist constants $M_k > 0, \omega_k \in \mathbb{R}$ such that

$$\|e^{t\overline{\mathcal{O}}_n}\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq M_k e^{\omega_k t} \,, \tag{3.35}$$

for all $t \in \mathbb{R}_+$ and independently of n. Then the pointwise limit $(\mathcal{O}, \mathcal{D}(\mathcal{O}) = \mathcal{T}_f(\mathcal{F}))$ is a core to the generator of a Sobolev-preserving semigroup $(e^{t\overline{\mathcal{O}}})_{t \in \mathbb{R}_+}$ on all QSSs, while the maps $e^{t\overline{\mathcal{O}}}$, for $t \in \mathbb{R}_+$, are the strong limits—uniformly on compact time intervals—of the $(e^{t\overline{\mathcal{O}}_n})_{n \in \mathbb{N}}$, and thereby inherit the bounds of eq. (3.35).

Combining this lemma with the result of theorem 3.4.2, we can take the limit $\varepsilon \to 0$ for the family $(\mathcal{G}_\varepsilon, \mathcal{D}(\mathcal{G}_\varepsilon))$. This leads to the following result, where complete positivity follows from the fact that the limiting semigroup arises as the strong limit of completely positive maps.

Corollary 3.4.4 ([GMR24, Lemma 3.4]) *The operator $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$ from eq. (3.29) is the core to the generator of a CP, Sobolev-preserving semigroup $(e^{t\overline{\mathcal{G}}})_{t \in \mathbb{R}_+}$ on all QSSs. For $k \in \mathbb{R}_+$ this semigroup satisfies*

$$\|e^{t\overline{\mathcal{G}}}\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq e^{\omega(k) t}$$

where $\omega(k) = \frac{k_{r'} - k}{k_{r'} - k_r}\omega_r + \frac{k - k_r}{k_{r'} - k_r}\omega_{r'}$, choosing r, r' such that k_r is the greatest lower bound and $k_{r'}$ the smallest upper bound to k; in the exact case $k = k_r$, one sets $\omega(k) = \omega_{k_r}$.

With this result in place, we now turn to the operator

$$\mathcal{L}_\delta(X) \equiv \mathcal{G}(X) + \delta\Sigma(X) \,, \qquad X \in \mathcal{D}(\mathcal{L}_\delta) \equiv \mathcal{T}_f(\mathcal{F})$$

with constituents defined in eq. (3.29) and $\delta \in (0, 1)$. Fixing $k = k_r$ and $\omega = \omega_r$ for arbitrary $r \in \mathbb{N}$, we use theorem 3.4.4 to verify the conditions of the Lumer-Phillips theorem for $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$. The general case $k \in \mathbb{R}_+$ then follows by interpolation, via theorem 1.4.2 and the compact embedding of the QSSs as we have already argued before. The two required properties are:

1. The operator $(\mathcal{L}_\delta - \omega\,\mathrm{id}, \mathcal{D}(\mathcal{L}_\delta))$ is dissipative; that is for all $X \in \mathcal{D}(\mathcal{L}_\delta)$, $\lambda > 0$,
$$\lambda \|X\|_{\mathcal{W}^k} \leq \|(\lambda\,\mathrm{id} - (\mathcal{L}_\delta - \omega\,\mathrm{id}))(X)\|_{\mathcal{W}^k}\,.$$
2. The range of $(\lambda_0\,\mathrm{id} - \mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$ is dense for some $\lambda_0 > \omega$.

Following the approach in the proof of theorem 3.4.2, we begin with the second point. The argument mirrors that of Davies (as summarised in section 1.3.3), but applied on $\mathcal{W}^k$. For $\lambda > \omega$, we may rewrite
$$(\lambda - \mathcal{L}_\delta)(X) = (\mathrm{id} - \delta\,\Sigma\, R(\lambda, \overline{\mathcal{G}}))(\lambda\,\mathrm{id} - \mathcal{G})(X)$$
for $X \in \mathcal{D}(\mathcal{L}_\delta)$. From eq. (2.9), we first obtain
$$\left\|\Sigma R(\lambda, \overline{\mathcal{G}})(Y)\right\|_{\mathcal{W}^k} \leq \|Y\|_{\mathcal{W}^k}\,,$$
for positive-semidefinite $Y \in (\lambda\,\mathrm{id} - \mathcal{G})\,\mathcal{D}(\mathcal{L}_\delta)$, then use the CP property of $R(\lambda, \overline{\mathcal{G}})$ and Σ to extend to self-adjoint (as in [GMR24, Lemma 3.3]) and finally arbitrary $Y \in (\lambda\,\mathrm{id} - \mathcal{G})\,\mathcal{D}(\mathcal{L}_\delta)$, analogously to [Mö25, Proof of Theorem 4.1.1]. By the density of $(\lambda\,\mathrm{id} - \mathcal{G})\,\mathcal{D}(\mathcal{L}_\delta)$ in $\mathcal{W}^k$, the operator $\Sigma R(\lambda, \overline{\mathcal{G}})$ extends to a CP, bounded and even contractive operator $\mathcal{A}_\lambda$ on $\mathcal{W}^k$. This then yields the existence of $R(1, \delta\,\mathcal{A}_\lambda) = \sum_{n=0}^{\infty} \delta^n\,\mathcal{A}_\lambda^n$ defined by the von Neumann series, since $\delta < 1$. Thus, the composition $(\lambda - \mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$ has dense range, and is the core for the closed operator $(\overline{\mathcal{L}}_\delta, \mathcal{D}(\overline{\mathcal{L}}_\delta))$, whose CP resolvent exists for all $\lambda > \omega$ and is given by
$$R(\lambda, \overline{\mathcal{L}}_\delta) = R(\lambda, \overline{\mathcal{G}}) \sum_{n=0}^{\infty} \delta^n\,\mathcal{A}_\lambda^n\,,$$
which, in particular, establishes the second requirement of Lumer-Phillips. To conclude dissipativity, we combine that $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$ is a core with eq. (2.9) to deduce that for all positive-semidefinite $X \in \mathcal{D}(\overline{\mathcal{L}}_\delta)$,
$$\lambda \|X\|_{\mathcal{W}^k} \leq \left\|(\lambda\,\mathrm{id} - (\overline{\mathcal{L}}_\delta - \omega\,\mathrm{id}))(X)\right\|_{\mathcal{W}^k}\,.$$
Via the CP property of $R(\lambda, \overline{\mathcal{L}}_\delta)$ one can extend this to all $X \in \mathcal{D}(\overline{\mathcal{L}}_\delta)$, meaning in particular that for $X \in \mathcal{D}(\mathcal{L}_\delta)$
$$\lambda \|X\|_{\mathcal{W}^k} \leq \|(\lambda\,\mathrm{id} - (\mathcal{L}_\delta - \omega\,\mathrm{id}))(X)\|_{\mathcal{W}^k}\,,$$
thereby fulfilling the missing first requirement of the Lumer-Phillips theorem.
In summary, we have established that $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$ is the core of the generator of a CP, Sobolev-preserving semigroup on all QSSs, with bounds exactly as given in theorem 3.4.4, and in particular, independent of δ. Consequently, the approximation lemma (theorem 3.4.3) is applicable, yielding one of the main objectives of the project [GMR24] summarised below, where trace preservation follows from $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ being a core.

Theorem 3.4.5 ([GMR24, Theorem 3.1]) *Let*
$$\mathcal{L}(X) = -i[H, X] + \sum_{j=1}^{J} \Big(L_j X L_j^{\circledast} - \frac{1}{2} L_j^{\circledast} L_j, X\Big),, \qquad X \in \mathcal{D}(\mathcal{L}) = \mathcal{T}_f(\mathcal{F})\,, \tag{3.36}$$
where $p_j \in \mathbb{C}[x, y]$ for $j = 0, \ldots, J$, with $H = p_0(a, a^*) = p_0^{\circledast}(a, a^*) = H^{\circledast}$ and $L_j = p_j(a, a^*)$ for $j = 1, \ldots, J$. Suppose that for a strictly monotone divergent sequence $(k_r)_{r\in\mathbb{N}} \subset \mathbb{R}_+$ with associated $(\omega_r)_{r\in\mathbb{N}} \subset \mathbb{R}_+$, we have for all quantum states $\rho \in \mathcal{T}_f(\mathcal{F})$,
$$\mathrm{Tr}\big[\mathcal{L}(\rho)(\mathbf{N} + \mathbb{1})^{k_r}\big] \leq \omega_r\,\mathrm{Tr}\big[\rho(\mathbf{N} + \mathbb{1})^{k_r}\big]\,. \tag{3.37}$$
Then $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is the core of the generator of a CP, Sobolev-preserving QMS $(e^{t\overline{\mathcal{L}}})_{t\in\mathbb{R}_+}$ on all QSSs, which for all $k, t \in \mathbb{R}_+$ satisfies,
$$\left\|e^{t\overline{\mathcal{L}}}\right\|_{\mathcal{W}^k \to \mathcal{W}^k} \leq e^{\omega(k)t}\,,$$
where $\omega(k) = \frac{k_{r'} - k}{k_{r'} - k_r}\omega_r + \frac{k - k_r}{k_{r'} - k_r}\omega_{r'}$, choosing indices r, r' such that k_r is the largest lower bound and $k_{r'}$ the smallest upper bound to k; in the case $k = k_r$, one sets $\omega(k) = \omega_{k_r}$.

Although this constitutes the primary technical step towards the complete objective from section 2.4, we still lack the refined bounds on the norms $\left\|e^{t\overline{\mathcal{L}}}\right\|_{\mathcal{W}^k \to \mathcal{W}^k}$ and $\left\|e^{t\overline{\mathcal{L}}}\right\|_{1 \to \mathcal{W}^k}$ under the stronger assumption eq. (2.11), which will be addressed next.

3.4.2 Improved estimates on norms of the semigroups

In this section, we show that under assumption eq. (2.11), one obtains an upper bound on $\left\|e^{t\overline{\mathcal{L}}}\right\|_{\mathcal{W}^k\to\mathcal{W}^k}$ that remains uniformly bounded in time. This result has already been established in [GMR24; Mö25]. Moreover, we will derive for the first time an explicit bound on $\left\|e^{t\overline{\mathcal{L}}}\right\|_{1\to\mathcal{W}^k}$ for all $t>0$, provided that eq. (2.11) holds with $\delta>0$, thereby settling an open question of [GMR24]. Both estimates are particularly valuable in perturbative settings, which we will later observe in concrete examples (see section 3.4.3).
To establish the forthcoming results, we begin with two auxiliary lemmas. The first, although known, lacks a readily available reference, while the second was used in [ASR15] but without an explicit proof. For completeness, we provide full derivations of both here. The first lemma is a generalisation of the Grönwall inequality, allowing for a positive power instead of a linear term in the differential inequality.

Lemma 3.4.6 *Let $y(t):\mathbb{R}_+\to\mathbb{R}_+$ be continuous on its domain and differentiable on $(0,\infty)$, and suppose*

$$y'(t)\equiv\frac{dy}{dt}(t)\leq -a(y(t))^p+b$$

for some constants $a>0, b\geq 0$, and $p\geq 1$, and for all $t\in(0,\infty)$. Then:

1. *For $p>1$,*

$$y(t)\leq\frac{1}{(x_0(p)^{1-p}+(p-1)at)^{1/(p-1)}}+\left(\frac{b}{a}\right)^{1/p}\leq\frac{1}{((p-1)at)^{1/(p-1)}}+\left(\frac{b}{a}\right)^{1/p}, \tag{3.38}$$

2. *and for $p=1$,*

$$y(t)\leq x_0(1)e^{-at}+\frac{b}{a}\,. \tag{3.39}$$

These bounds hold for all $t\in[0,\infty)$, where $x_0(p)\equiv\max\left\{y(0)-\left(\frac{b}{a}\right)^{1/p},0\right\}$. Note that in eq. (3.38), we adopt the convention $0^{1-p}=\infty$ and $1/\infty=0$; hence, if $x_0(p)=0$, the first bound reduces to $y(t)\leq\left(\frac{b}{a}\right)^{1/p}$.

Proof. We omit the proof of eq. (3.39), as it follows directly from the standard Grönwall inequality. The arguments for the case $p>1$, which we detail below, proceed along similar lines. Let $z(t)$ denote the expression appearing between the left- and right-hand sides of the inequality in (3.38), and define:

$$t_c\equiv\inf\left\{t:t\in(0,\infty),y(t)\leq\left(\frac{b}{a}\right)^{1/p}\right\}.$$

For $t>t_c$, we then have $y(t)\leq\left(\frac{b}{a}\right)^{1/p}$, and in particular $y(t)\leq z(t)$. We prove this by contradiction. Suppose there exists $t_v>t_c$ such that $y(t_v)>\left(\frac{b}{a}\right)^{1/p}$. By continuity of y one can apply the intermediate value theorem, to conclude existence of $t_0,t_1\in\mathbb{R}_+$ with $t_0<t_1$ such that

$$y(t)\geq\left(\frac{b}{a}\right)^{1/p}\quad\forall t\in[t_0,t_1]\,,\quad\text{and}\quad y(t_0)=\left(\frac{b}{a}\right)^{1/p},\quad y(t_1)>\left(\frac{b}{a}\right)^{1/p}.$$

By the mean value theorem, there exists $t_m\in(t_0,t_1)$ such that

$$y'(t_m)=\frac{y(t_1)-y(t_0)}{t_1-t_0}>0\,.$$

However, since $y(t_m)\geq\left(\frac{b}{a}\right)^{1/p}$, the differential inequality implies

$$y'(t_m)\leq -ay(t_m)^p+b\leq -a\frac{b}{a}+b=0\,,$$

yielding a contradiction.
Now suppose $t_c>0$. Define the function $g:(0,t_c)\to(0,\infty)$ by

$$g(t)\equiv y(t)-\left(\frac{b}{a}\right)^{1/p}.$$

This function is continuous on $[0, t_c]$ and differentiable on $(0, t_c)$, with $g'(t) = y'(t)$. Using the superadditivity of $x \mapsto x^p$ for $x \geq 0$, the differential inequality becomes,

$$g'(t) \leq -a\Big(g(t) + \Big(\frac{b}{a}\Big)^{1/p}\Big)^p + b \leq -ag(t)^p - a\frac{b}{a} + a = -ag(t)^p\,,$$

for $t \in (0, t_c)$. Separating variables and integrating from 0 to $t < t_c$ yields

$$\frac{g(t)^{1-p}}{1-p} - \frac{g(0)^{1-p}}{1-p} = \int_0^t \frac{g'(s)}{g(s)^p}\,ds \leq -at\,.$$

Rearranging gives

$$g(t) \leq \big(g(0)^{1-p} + (p-1)at\big)^{-1/(p-1)}\,,$$

which after resubstitution of $y(t)$ and identification of $g(0) = x_0(p)$ yields, for $t \in (0, t_c)$,

$$y(t) \leq \big(x_0(p)^{1-p} + (p-1)at\big)^{-1/(p-1)} + \left(\frac{b}{a}\right)^{1/p} = z(t)\,.$$

The edge cases $t = 0$ and $t = t_c$ follows by continuity. Combined with the bound for $t > t_c$, this completes the proof, except for the final inequality in eq. (3.38), which follows from the fact that $x_0(p)^{1-p} \geq 0$ and the monotonicity of $x \mapsto x^{-1/(p-1)}$ on $(0, \infty)$. □

The second result is a reformulation of a classical statistical inequality—namely Jensen's inequality—in the context of bosonic systems, applied to the expectation value of moments of the number operator.

Lemma 3.4.7 *Let $\rho \in \mathcal{T}_f(\mathcal{F})$ be a quantum state and $p \geq q > 0$. Then*

$$(\mathrm{Tr}[\rho(\mathbf{N}+\mathbb{1})^q])^{\frac{p}{q}} \leq \mathrm{Tr}[\rho(\mathbf{N}+\mathbb{1})^p]\,.$$

Proof. We compute

$$\begin{aligned}\mathrm{Tr}[\rho(\mathbf{N}+\mathbb{1})^p] &= \sum_{n=1}^{N} \langle n, \rho n\rangle\,(n+1)^p = \sum_{n=1}^{N} \langle n, \rho n\rangle\,(n+1)^{q\cdot\frac{p}{q}} \\ &\geq \Big(\sum_{n=1}^{N} \langle n, \rho n\rangle\,(n+1)^q\Big)^{\frac{p}{q}} = (\mathrm{Tr}[\rho(\mathbf{N}+\mathbb{1})^q])^{\frac{p}{q}}\,,\end{aligned}$$

where we have used that ρ has finite rank $N \in \mathbb{N}$ in the Fock basis, $(p_n \equiv \langle n, \rho n\rangle)_{n=1}^N$ is a probability distribution, and the function $x \mapsto x^{\frac{p}{q}}$ is convex. □

Combining the results above, we now derive the following extension of theorem 3.4.5, addressing our objective of obtaining stronger bounds from stronger inequalities (see section 2.4).

Theorem 3.4.8 *Let $(e^{t\overline{\mathcal{L}}})_{t\in\mathbb{R}_+}$ be a Sobolev-preserving QMS. Let $k \in \mathbb{R}_+$, and suppose $(\mathcal{L}, \mathcal{D}(\mathcal{L}) \equiv \mathcal{T}_f(\mathcal{F}))$ is a core for its generator on $\mathcal{W}^k$, with $\mathcal{L}(\mathcal{D}(\mathcal{L})) \subset \mathcal{T}_f(\mathcal{F})$. If there exist $\nu > 0$, $\mu \geq 0$, and $\delta \geq 0$ such that*

$$\mathrm{Tr}\big[\mathcal{L}(\rho)(\mathbf{N}+\mathbb{1})^k\big] \leq -\nu\,\mathrm{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^{k+\delta}\big] + \mu\,, \tag{3.40}$$

for all $\rho \in \mathcal{D}(\mathcal{L})$ quantum states, then

$$\big\|e^{t\overline{\mathcal{L}}}\big\|_{\mathcal{W}^k\to\mathcal{W}^k} \leq e^{-\nu t} + \frac{\mu}{\nu}\,,$$

and if $\delta > 0$, we also have

$$\big\|e^{t\overline{\mathcal{L}}}\big\|_{1\to\mathcal{W}^k} \leq \left(\frac{k}{\delta\nu t}\right)^{k/\delta} + \left(\frac{\mu}{\nu}\right)^{k/(k+\delta)}\,.$$

Proof. We treat the case $\delta > 0$; the case $\delta = 0$ follows as a special case. Since $\mathcal{T}_f(\mathcal{F})$ is dense in $\mathcal{W}^k$ and the semigroup is CP, it suffices to show that

$$\left\|e^{t\overline{\mathcal{L}}}(X)\right\|_{\mathcal{W}^k} \leq \left(e^{-\nu t} + \frac{\mu}{\nu}\right)\|X\|_{\mathcal{W}^k}, \quad \text{and} \quad \left\|e^{t\overline{\mathcal{L}}}(X)\right\|_{\mathcal{W}^k} \leq \left(\left(\frac{k}{\delta\nu t}\right)^{k/\delta} + \left(\frac{\mu}{\nu}\right)^{k/(k+\delta)}\right)\|X\|_1$$

for positive-semidefinite $X \in \mathcal{T}_f(\mathcal{F})$ (cf. [GMR24, Lemma 3.3], [Mö25, Theorem 4.1.1]). As furthermore $\mathcal{W}^k$ is compactly embedded into $\mathcal{T}(\mathcal{F})$, we have $\|X\|_1 \leq \|X\|_{\mathcal{W}^k}$, so it suffices to prove

$$\left\|e^{t\overline{\mathcal{L}}}(\rho)\right\|_{\mathcal{W}^k} \leq e^{-\nu t}\|\rho\|_{\mathcal{W}^k} + \frac{\mu}{\nu}, \quad \text{and} \quad \|e^{t\overline{\mathcal{L}}}(\rho)\|_{\mathcal{W}^k} \leq \left(\frac{k}{\delta\nu t}\right)^{k/\delta} + \left(\frac{\mu}{\nu}\right)^{k/(k+\delta)}$$

for quantum states $\rho \in \mathcal{T}_f(\mathcal{F})$.
As the semigroup is CP and Sobolev-preserving, $(e^{t\overline{\mathcal{L}}}(\rho))_{t\in\mathbb{R}_+}$ lies in $\mathcal{W}^{k'}$ for all $k' \geq 0$ if $\rho \in \mathcal{T}_f(\mathcal{F})$, thus

$$y(t) \equiv \left\|e^{t\overline{\mathcal{L}}}(\rho)\right\|_{\mathcal{W}^k} = \mathrm{Tr}\Big[e^{t\overline{\mathcal{L}}}(\rho)(\mathbf{N}+\mathbb{1})^k\Big]$$

is continuously differentiable for $t > 0$. The assumption on $\mathcal{L}$, i.e., eq. (3.40), by the Sobolev-preservation extends to states in $\bigcup_{t\geq 0} e^{t\overline{\mathcal{L}}}(\mathcal{T}_f(\mathcal{F}))$, hence by $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ being a core:

$$y'(t) = \mathrm{Tr}\Big[\mathcal{L}(e^{t\overline{\mathcal{L}}}(\rho))(\mathbf{N}+\mathbb{1})^k\Big] \leq -\nu\,\mathrm{Tr}\Big[e^{t\overline{\mathcal{L}}}(\rho)(\mathbf{N}+\mathbb{1})^{k+\delta}\Big] + \mu\,.$$

By theorem 3.4.7, we then have

$$y'(t) \leq -\nu y(t)^{(k+\delta)/k} + \mu \leq -\nu y(t) + \mu\,,$$

where the final inequality uses that $y(t) = \mathrm{Tr}\left[e^{t\overline{\mathcal{L}}}(\rho)(\mathbf{N}+\mathbb{1})^k\right] \geq \mathrm{Tr}\left[e^{t\overline{\mathcal{L}}}(\rho)\right] = 1$, as the semigroup is trace-preserving. Applying theorem 3.4.6, we obtain

$$\left\|e^{t\overline{\mathcal{L}}}(\rho)\right\|_{\mathcal{W}^k} = y(t) \leq \begin{cases} e^{-\nu t}\|\rho\|_{\mathcal{W}^k} + \dfrac{\mu}{\nu} & \text{if } \delta \geq 0\,, \\ \left(\dfrac{k}{\delta\nu t}\right)^{k/\delta} + \left(\dfrac{\mu}{\nu}\right)^{k/(k+\delta)} & \text{if } \delta > 0\,, \end{cases}$$

which completes the proof. □

So far the discussion has been rather abstract, so let us now justify the previous definitions and results by providing concrete examples in the next section that fit into this framework.

3.4.3 Examples and discussion of perturbation theory

In this section, we revisit the examples first introduced in section 1.4.4 and further employ the framework of Sobolev-preserving semigroups, along with the stronger bounds established in theorem 3.4.8, to derive perturbation results for them. These case studies and their perturbative analysis are intended as proofs of concept, illustrating the practical use of the Sobolev-preserving semigroup framework in describing physically relevant systems. Accordingly, we do not aim for generality here but focus on simple, illustrative cases.
Although we remain close to concrete examples, we briefly want to outline the general strategy for proving the sufficient condition eq. (3.37), without detailing its technical aspects. For a generator $(\mathcal{L}, \mathcal{T}_f(\mathcal{F}))$ in GKSL-form, built from polynomials in the annihilation and creation operators a and a^*, we must show that for example for $k \in \mathbb{N}$ and a sequence $(\omega_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}_+$, the following inequality holds for all quantum states $\rho \in \mathcal{T}_f(\mathcal{F})$:

$$\mathrm{Tr}\big[\mathcal{L}(\rho)(\mathbf{N}+\mathbb{1})^k\big] \leq \omega_k\,\mathrm{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^k\big].$$

Since ρ has finite rank, we may treat the unbounded operators a, a^*, and $\mathbf{N}$ analogously to bounded ones. In particular, we can write

$$\mathrm{Tr}\big[\mathcal{L}(\rho)(\mathbf{N}+\mathbb{1})^k\big] = \mathrm{Tr}\Big[\rho\,\mathcal{L}^\dagger((\mathbf{N}+\mathbb{1})^k)\Big]$$

by using the HS-adjoint and the cyclicity of the trace. Using the canonical commutation relations $[a, a^*] = \mathbb{1}$, one rewrites the polynomial $\mathcal{L}^\dagger((\mathbf{N}+\mathbb{1})^k)$ in a and a^* so that each monomial involves only the number operator $\mathbf{N}$ and either powers of only a or a^* respectively. Monomials that beside the number operator involve such powers can be paired together with their formal adjoint and estimated by polynomials in $\mathbf{N}$, using spectral estimates as in [GMR24, Lemma 4.6]. Ultimately, one arrives at

$$\mathrm{Tr}\big[\mathcal{L}(\rho)(\mathbf{N}+\mathbb{1})^k\big] \leq \mathrm{Tr}[\rho\, p_k(\mathbf{N})], \tag{3.41}$$

where $p_k \in \mathbb{R}[x]$. Let the leading term of p_k be $cx^{\deg(p_k)}$ with $c \neq 0$. Define

$$\mu_k \equiv \begin{cases} \sup\{p_k(x) - \frac{c}{2}(x+1)^{\deg(p_k)} : x \in \mathbb{R}_+\} & \text{if } c < 0\,, \\ \sup\{p_k(x) - 2c(x+1)^{\deg(p_k)} : x \in \mathbb{R}_+\} & \text{if } c > 0\,, \end{cases}$$

which is finite by construction. It follows that for all $x \in \mathbb{R}_+$,

$$p_k(x) \leq \nu_k(x+1)^{\deg(p_k)} + \mu_k\,, \quad \text{where } \nu_k \equiv \begin{cases} \frac{c}{2} & \text{if } c < 0\,, \\ 2c & \text{if } c > 0\,. \end{cases}$$

Since $\mathbf{N}$ has non-negative spectrum and ρ is a quantum state, we can apply the above estimate directly to the right-hand-side of eq. (3.41) and obtain

$$\mathrm{Tr}\big[\mathcal{L}(\rho)(\mathbf{N}+\mathbb{1})^k\big] \leq \nu_k\,\mathrm{Tr}\Big[\rho(\mathbf{N}+\mathbb{1})^{\deg(p_k)}\Big] + \mu_k\,.$$

If this inequality holds for all $k \in \mathbb{N}$ and $\deg(p_k) \leq k$ whenever $\nu_k > 0$, then eq. (3.37) and for $\nu_k < 0$ even eq. (3.40) follow. In particular, the sufficient condition of theorem 3.4.5 is satisfied. Moreover, in the aforementioned case where $v_k < 0$, one obtains the stronger bounds stated in theorem 3.4.8. Although the associated estimates are tedious, this approach is, in principle, applicable to any operator of the form given in eq. (3.36).

The requirement of $\deg(p_k) \leq k$ if $\nu_k > 0$ of course raises the converse question: namely, whether a generator that fails to satisfy eq. (3.37) cannot generate a QMS. Before exploring such future directions, however, we now present the concrete examples.

We begin with the quantum OU semigroups (see section 1.4.4), for which we now introduce the shorthand notation:

$$\mathcal{L}_{qOU}[\mu,\nu](X) \equiv \mu^2\,\mathcal{L}[a^*](X) + \nu^2\,\mathcal{L}[a](X), \quad X \in \mathcal{D}(\mathcal{L}_{qOU}) \equiv \mathcal{T}_f(\mathcal{F})\,, \tag{3.42}$$

for $\mu, \nu \in \mathbb{R}_+$, where we use the abbreviation $\mathcal{L}[L] \equiv LXL^{\circledast} - \frac{1}{2}\{L^{\circledast}L, X\}$, as introduced earlier. For these semigroups, we state both eq. (3.37) and eq. (3.40) explicitly, with proofs given in [GMR24, Lemma 4.2]. We also include an example of general single-mode Gaussian semigroups, written more compactly than in their earlier introduction:

$$\mathcal{L}_G[\alpha,\beta,\eta,\gamma,\chi](X) \equiv -i[\alpha a + \overline{\alpha} a^* + \beta a^2 + \overline{\beta}(a^*)^2 + \chi\,\mathbf{N}, X] + \mathcal{L}[\eta a^* + \gamma a](X)\,, \quad X \in \mathcal{D}(\mathcal{L}_G) \equiv \mathcal{T}_f(\mathcal{F})\,, \tag{3.43}$$

where $\alpha, \beta, \eta, \gamma \in \mathbb{C}$, $\chi \in \mathbb{R}$. While we do not give explicit parameter estimates here, they can be obtained using the same techniques as in [GMR24, Lemma 4.2, 4.6, Section 4.2].

Heuristically, one may regard $\mathcal{L}_G$ as a quantum OU semigroup with $\mu = |\eta|$ and $\nu = |\gamma|$, perturbed by a quadratic Hamiltonian and additional terms. Applying techniques from [GMR24, Lemma 4.6] one can show that these perturbations also contribute terms of degree at most k in the polynomial (3.41) corresponding to $\mathcal{L}_G$, so the overall leading order does not exceed k.

Example 4 ([GMR24, Lemma 4.2, Section 4.2, Proposition 5.1]).

1. Let $k \in \mathbb{N}$, $\mu, \nu \in \mathbb{R}_+$, and $\rho \in \mathcal{T}_f(\mathcal{F})$ a quantum state. Then

$$\mathrm{Tr}\big[\mathcal{L}_{qOU}[\mu,\nu](\rho)(\mathbf{N}+\mathbb{1})^k\big] \leq \begin{cases} \frac{k}{2}(\mu^2-\nu^2)\,\mathrm{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^k\big] + \Big(2\frac{\mu^2+\nu^2+2k}{\nu^2-\mu^2}\Big)^k & \nu > \mu\,, \\ k(2\mu^2+k)\,\mathrm{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^k\big] & \nu \leq \mu\,. \end{cases}$$

Thus, $(\mathcal{L}_{qOU}[\mu,\nu], \mathcal{D}(\mathcal{L}_{qOU}))$ defines a Sobolev-preserving QMS. When $\nu > \mu$, the stronger bounds of theorem 3.4.8 apply, and interpolation (theorem 1.4.2) extends them to all $k \in \mathbb{R}_+$.

2. Similarly, for each $k \in \mathbb{N}$, we expect there exists $\omega_k \in \mathbb{R}_+$, depending on $\alpha, \beta, \eta, \gamma \in \mathbb{C}$, $\chi \in \mathbb{R}$, and k, such that for any quantum state $\rho \in \mathcal{T}_f(\mathcal{F})$,
$$\operatorname{Tr}\big[\mathcal{L}_G[\alpha,\beta,\eta,\gamma,\chi](\rho)(\mathbf{N}+\mathbb{1})^k\big] \leq \omega_k \operatorname{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^k\big]\,.$$
Hence, $(\mathcal{L}_G[\alpha,\beta,\eta,\gamma,\chi], \mathcal{D}(\mathcal{L}_G))$ also defines a Sobolev-preserving QMS. A detailed parameter analysis, which we omit here, should reveal when the stronger bounds of theorem 3.4.8 are available, similarly to the quantum OU case.

Although generation properties for general Gaussian semigroups were already established in [AFP21]—as discussed in section 1.4.4—the Sobolev preservation result is novel and enables a straightforward perturbative analysis.
Consider, for instance, the formal generator $(\mathcal{L}[a], \mathcal{T}_f(\mathcal{F}))$ and its perturbation $(\mathcal{L}[a] + \varepsilon\,\mathcal{L}[a^*], \mathcal{T}_f(\mathcal{F}))$ for $\varepsilon \in [0,1)$. By the preceding example, both are cores to generators of Sobolev-preserving QMSs, which in particular implies $\big\|e^{t(\overline{\mathcal{L}[a]+\varepsilon\,\mathcal{L}[a^*]})}\big\|_{1\to1} \leq 1$. One further has constants $c_1, c_2 \in \mathbb{R}_+$ such that
$$\big\|e^{t\overline{\mathcal{L}[a]}}\big\|_{\mathcal{W}^2\to\mathcal{W}^2} \leq c_1\,, \quad \forall t \in \mathbb{R} \quad \text{and} \quad \big\|\overline{\mathcal{L}[a^*]}\big\|_{\mathcal{W}^2\to1} \leq c_2\,,$$
with the first inequality following by the improved bound from theorem 3.4.8. For $X \in \mathcal{T}_f(\mathcal{F})$ and $t \in \mathbb{R}_+$, one may apply Duhamel's formula:
$$\big(e^{t\overline{\mathcal{L}[a]}} - e^{t(\overline{\mathcal{L}[a]+\varepsilon\,\mathcal{L}[a^*]})}\big)(X) = \varepsilon\int_0^t e^{s(\overline{\mathcal{L}[a]+\varepsilon\,\mathcal{L}[a^*]})}\,\mathcal{L}[a^*]e^{(t-s)\overline{\mathcal{L}[a]}}(X)\,ds\,,$$
with the right-hand side well-defined by Sobolev preservation and the fact that $X \in \mathcal{T}_f(\mathcal{F})$. Applying the above estimates yields
$$\Big\|\big(e^{t\overline{\mathcal{L}[a]}} - e^{(t-s)(\overline{\mathcal{L}[a]+\varepsilon\,\mathcal{L}[a^*]})}\big)(X)\Big\|_1 \leq \varepsilon t c_1 c_2 \|X\|_{\mathcal{W}^2}\,,$$
which extends to the operator norm bound
$$\Big\|\big(e^{t\overline{\mathcal{L}[a]}} - e^{t(\overline{\mathcal{L}[a]+\varepsilon\,\mathcal{L}[a^*]})}\big)\Big\|_{\mathcal{W}^2\to1} \leq \varepsilon t c_1 c_2\,,$$
by the density of $\mathcal{T}_f(\mathcal{F})$ in $\mathcal{W}^2$.
This example naturally generalises to all QMS with generators of the form in eq. (3.36) that satisfy the sufficient condition for Sobolev preservation in eq. (3.37) (though the time-dependence of bounds scales exponentially in the worst case).
As discussed in [GMR24], such results can be viewed as canonical extensions of perturbation bounds from the finite-dimensional setting. However, in the infinite-dimensional case, one must account for the unbounded operator in the Duhamel formula, which naturally leads to estimates in $\mathcal{W}^k \to 1$ norms rather than the standard $1 \to 1$ norm familiar from finite dimensions.
To improve upon the $\mathcal{W}^k \to 1$ bounds, one is led to consider formal generators whose dissipative parts are of higher than linear order in the annihilation operator.
Before delving into this direction, let us briefly comment on the converse situation—namely, when higher-order terms involve the creation operator a^*. This introduces a qualitative change in behaviour. For instance, in the case of a formal generator such as $(\mathcal{L}[(a^*)^2], \mathcal{T}_f(\mathcal{F}))$, one can no longer establish eq. (3.37) which coincides with the breakdown of trace preservation by the minimal semigroup (see [Dav77, Example 3.3]). Although there exists a workaround to the latter issue by modifying the generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ of the minimal semigroup through a reset mechanism, such a construction appears unphysical—at least naively—thus suggesting the naturalness of requiring eq. (3.37). We call it unphysical, as it requires selecting an arbitrary quantum state $\sigma \in \mathcal{D}(\mathcal{L})$, and subsequently defines the modified operator
$$\mathcal{L}'(X) \equiv \mathcal{L}(X) - \sigma\operatorname{Tr}[\mathcal{L}(X)]\,, \quad X \in \mathcal{D}(\mathcal{L}') \equiv \mathcal{D}(\mathcal{L})\,,$$
thereby not only altering the original dynamics but doing so ambiguously. Only the resulting $(\mathcal{L}', \mathcal{D}(\mathcal{L}'))$ serves as the generator of a QMS (see [Dav77, Theorem 3.4]).

One could further argue that divergence of the expectation of a physical observable (the number operator and its moments) in finite time, i.e., failure of eq. (3.37) is unphysical too—further underpinning the naturalness of our assumption at least from a physics perspective.
This is, of course, only a superficial assessment, made under lack of knowledge of any physical system that exhibits dynamics corresponding to the above, where maybe such reset mechanism is natural to ask for and the divergence of the moments of the number operator have a reasonable explanation. Clearly, a more detailed investigation into the paired occurrence of the failure of Sobolev preservation and the lack of trace preservation in the minimal semigroup is in order. Such analysis may even yield a proof of the necessity of eq. (3.37) for eq. (3.36) to define a core for a QMS.
Whereas higher-order creation operators introduce instability, higher-order dissipations have a stabilising effect. This can for example be observed in the formal generators of the cat codes.
Recall their definition from section 1.4.4, where we drop the explicit $\alpha \in \mathbb{R}$ dependence in their short forms:

$$\begin{aligned} \mathcal{L}_{\mathrm{id}}(X) &\equiv \mathcal{L}[a^2-\alpha^2]\,, & \mathcal{L}_{Z(\theta)}(X) &\equiv \mathcal{L}[a^2-\alpha^2] - i[a+a^*,X]\,,\\ \mathcal{L}_{\mathrm{id},bl}(X) &\equiv \mathcal{L}[a^2-\alpha^2]+\mathcal{L}[a]\,, & \mathcal{L}_{Z(\theta),bl}(X) &\equiv \mathcal{L}[a^2-\alpha^2] - i[a+a^*,X]+\mathcal{L}[a]\,, \end{aligned}$$

for $X \in \mathcal{T}_f(\mathcal{F})$, which forms the domain of all these generators. Although the cat codes are the motivation for the project, we want to complement them with the formal-generators

$$\mathcal{L}\{l\} \equiv \mathcal{L}[a^l], \qquad X \in \mathcal{T}_f(\mathcal{F})\,,$$

that can be regarded as higher order identity-gates with α set to zero. With them, we will be able to see more clearly the codependence between the regularisation effects of the semigroup and the order of a in the dissipation. Summarising the results from [GMR24, Corollaries 4.4 and 4.6], we obtain the following bounds:

Example 5 ([GMR24, Corollaries 4.4 and 4.6]).

1. For $k \in \mathbb{N}$ and $\rho \in \mathcal{T}_f(\mathcal{F})$:

$$\begin{aligned} \operatorname{Tr}\big[\mathcal{L}_{\mathrm{id}}(\rho)(\mathbf{N}+\mathbb{1})^k\big] &\le -\operatorname{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^{k+1}\big] + (6+2^{k+3}|\alpha|^2)^{k+1}\,,\\ \operatorname{Tr}\big[\mathcal{L}_{Z(\theta)}(\rho)(\mathbf{N}+\mathbb{1})^k\big] &\le -\operatorname{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^{k+1}\big] + (6+4k+2^{k+3}|\alpha|^2)^{k+1}\,,\\ \operatorname{Tr}\big[\mathcal{L}_{\mathrm{id},bl}(\rho)(\mathbf{N}+\mathbb{1})^k\big] &\le -\operatorname{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^{k+1}\big] + (6+2^{k+3}|\alpha|^2)^{k+1} + (2+4k)^k\,,\\ \operatorname{Tr}\big[\mathcal{L}_{Z(\theta),bl}(\rho)(\mathbf{N}+\mathbb{1})^k\big] &\le -\operatorname{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^{k+1}\big] + (6+4k+2^{k+3}|\alpha|^2)^{k+1} + (2+4k)^k\,. \end{aligned}$$

   By theorem 3.4.5, all these generators define cores to Sobolev-preserving QMS, which by theorem 3.4.8 admit bounded-time estimates on $\|\cdot\|_{\mathcal{W}^k\to\mathcal{W}^k}$ for all $k \in \mathbb{R}_+$, and also on $\|\cdot\|_{1\to\mathcal{W}^k}$ for $k \in \mathbb{R}_+$ and $t > 0$, with intermediate values bridged via theorem 1.4.2.

2. For $k, l \in \mathbb{N}$ with $l > 1$ and $\rho \in \mathcal{T}_f(\mathcal{F})$:

$$\operatorname{Tr}\big[\mathcal{L}\{l\}(\rho)(\mathbf{N}+\mathbb{1})^k\big] \le -\frac{l}{2}\operatorname{Tr}\big[\rho(\mathbf{N}+\mathbb{1})^{k+l-1}\big] + \frac{l}{2}(l(l+1))^{k+l-1}.$$

   Hence, $(\mathcal{L}\{l\}, \mathcal{T}_f(\mathcal{F}))$ is a core for a Sobolev-preserving QMS, and the corresponding semigroup again admits the estimates mentioned above.

While it was already known that $(\mathcal{L}[a^l-\alpha^l], \mathcal{T}_f(\mathcal{F}))$ generates a QMS [ASR16], the generation theory for other cat code generators was established only recently in the work underlying these sections, namely [GMR24]. That work further complemented the result by proving that the corresponding semigroups preserve regularity, which, together with newly established $1 \to \mathcal{W}^k$ estimates, now enables the application of perturbation theory as in previous cases, but yielding stronger bounds.
As an example, consider $(\mathcal{L}\{3\}, \mathcal{T}_f(\mathcal{F}))$ and its perturbation $(\mathcal{L}\{3\} + i[a+a^*,\cdot], \mathcal{T}_f(\mathcal{F}))$. Both are cores to generators of Sobolev-preserving QMS by the above examples. We have

$$\big\|e^{t(\overline{\mathcal{L}\{3\}+i[a+a^*,\cdot]})}\big\|_{1\to 1} \le 1, \quad \text{and} \quad \big\|e^{t\overline{\mathcal{L}\{3\}}}\big\|_{1\to\mathcal{W}^1} \le \Big(\frac{1}{3t}\Big)^{1/2} + 12$$

for all $t > 0$, and $\|i[a + a^*, \cdot]\|_{\mathcal{W}^1\to 1} \leq 4$. Duhamel's formula yields

$$(e^{t\overline{\mathcal{L}\{3\}}} - e^{t(\overline{\mathcal{L}\{3\}+i[a+a^*,\cdot]})})(X) = \int_0^t e^{s(\overline{\mathcal{L}\{3\}+i[a+a^*,\cdot]})} i[a + a^*, \cdot] e^{(t-s)\overline{\mathcal{L}\{3\}}}(X)\, ds,$$

which is again well-defined for $X \in \mathcal{T}_f(\mathcal{F})$ due to the Sobolev-preservation of the involved semigroups. Taking the trace norm on both sides, we obtain

$$\begin{aligned}\left\|(e^{t\overline{\mathcal{L}\{3\}}} - e^{t(\overline{\mathcal{L}\{3\}+i[a+a^*,\cdot]})})(X)\right\|_1 &\leq 4\Big(\int_0^t \frac{1}{\sqrt{3}\sqrt{(t-s)}}\, ds + 12t\Big)\|X\|_1 \\ &= \left(\frac{8}{\sqrt{3}}\sqrt{t} + 48t\right)\|X\|_1 ,\end{aligned}$$

which lifts to

$$\left\|(e^{t\overline{\mathcal{L}\{3\}}} - e^{t(\overline{\mathcal{L}\{3\}+i[a+a^*,\cdot]})})\right\|_{1\to 1} \leq \frac{8}{\sqrt{3}}\sqrt{t} + 48t\,. \tag{3.44}$$

We hence conclude that if the perturbation is sufficiently weak compared to the dissipative regularisation—e.g., via $(\mathcal{L}\{l\}, \mathcal{T}_f(\mathcal{F}))$—then even $1 \to 1$ norm estimates can be obtained. Note, however, that we have not optimised constants here, and this result should be seen primarily as a proof of principle.
Let us now end this discussion of examples with a brief outlook on potential future research directions.

3.4.4 Open questions and future work

A major open question is whether the condition stated in eq. (3.37) is not only sufficient but also necessary for the generation of a QMS, assuming the formal generator consists entirely of polynomials in creation and annihilation operators. As a first step towards addressing this question, one could examine Davies' counterexample—specifically [Dav77, Example 3.3]—and closely study similar constructions.
On the more practical side, one might aim to employ the norm bounds derived herein—particularly the $1 \to \mathcal{W}^k$ bound for the 'l-legged' identity-gate of the cat codes—in conjunction with eq. (1.97) to estimate the mixing time of the semigroup towards the code space. This could potentially lead to estimates of the spectral gap. Building on this, it would be of interest to investigate whether, under suitable regularity assumptions, some gap or MLSI results established for finite-dimensional systems can be extended to the Sobolev-preserving setting.
In addition, a useful extension of the current toolbox around Sobolev-preserving semigroups would be to obtain $\mathcal{W}^{k'} \to \mathcal{W}^k$ norm estimates for $k' < k$, as an alternative to the $1 \to \mathcal{W}^k$ estimate of theorem 3.4.8. Trading a larger k' for better scaling in powers of $1/t$ could be advantageous. If successful, this approach would enable the extension of the strategy used to derive eq. (3.44) to scenarios in which the degree of the perturbation polynomials is large relative to the order of the dissipation. In such cases, one might not obtain a $1 \to \mathcal{W}^k$ bound, but rather a $1 \to \mathcal{W}^{k'}$ bound with $k' < k$. The path to such bounds clearly lies in interpolation of the $1 \to \mathcal{W}^k$ and the $\mathcal{W}^k \to \mathcal{W}^k$ space using techniques similar to the ones in the proof of theorem 1.4.2.
Finally, we comment on the multimode extension of the above theory, developed in [GMR24]. The norm used there generalises $(\mathbf{N} + \mathbb{1})^k$ to $(\mathbf{N} + \mathbb{1})^{\mathbf{k}} \equiv \bigotimes_{j=1}^J (\mathbf{N}_i + \mathbb{1})^{k_i}$, while the remaining arguments carried over with minimal modifications. Although the compact embedding, the quantum Stein-Weiss theorem (theorem 1.4.2), and the generation theorem (theorem 3.4.5) could be proven in this multimode setting almost analogously, establishing eq. (3.37) for specific examples such as the CNOT gate of the cat code proved to be quite laborious. For the Toffoli gate, the complexity became prohibitive, and the investigation was ultimately abandoned. Tim Möbus and the author believe that a more promising approach would have been to derive a norm from the operator $\sum_{j=1}^J c_j(\mathbf{N}_j + \mathbb{1})^{k_j}$ with $c_j \in \mathbb{C}$ instead, which has already found application in [LRZ23; Mö+25; LRZ25; DLL25a; KVS24]. This alternative norm may facilitate proving an estimate for the Toffoli gate, and thereby help address the well-posedness of all formal generators proposed for the cat code. However, adopting this approach would require re-establishing the aforementioned results within this modified framework, which seems possibly but tedious.
With this, we conclude our discussion of [GMR24] and Sobolev-preserving semigroups more broadly and turn to a brief examination of a special case of the generalised quantum Stein's lemma.

3.5 Results and discussion of [GR24b]

We previously shortly hinted at the strategy to proof the generalised quantum Stein's lemma with a second hypothesis stemming from a conditional expectation onto a von Neumann subalgebra in section 2.5 and now proceed by following that outline in detail.
First, we want to establish the existence of a specific characterisation of the HS-adjoint of a conditional expectation, under the assumption that $E^\dagger(\iota) > 0$.

Lemma 3.5.1 *Let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ be a conditional expectation onto the von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$, with $\pi = E^\dagger(\iota) > 0$ (see theorem 1.3.2 for equivalent conditions). Then there exists a positive-definite operator $\Gamma : \mathcal{K} \to \mathcal{K}$, with $\mathcal{K} = \mathbb{C}^{d_{\mathcal{H}}^2}$, and an isometry $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$, such that*

$$E^\dagger(X) = V^*(\Gamma \otimes X)V ,$$

for all $X \in \mathcal{B}(\mathcal{H})$, where V further satisfies

$$\mathbb{1}_{\mathcal{K}} \otimes \pi^{-p} V \pi^p = \Gamma^p \otimes \mathbb{1}_{\mathcal{H}} V ,$$

for all $p \in \mathbb{R}$.

Proof. Since E commutes with the modular operator $\Delta_\pi(\cdot) = \pi \cdot \pi^{-1}$ by theorem 1.3.1, that is, $[E, \Delta_\pi] = 0$, the proof of [BDR20, Lemma 13] can be adapted directly. This yields a Kraus decomposition of E in terms of eigenvectors of Δ_π: there exist $(K_j)_{j=1}^{d_{\mathcal{H}}^2} \subset \mathcal{B}(\mathcal{H})$ such that $\sum_{j=1}^{d_{\mathcal{H}}^2} K_j K_j^* = \mathbb{1}_{\mathcal{H}}$, $\Delta_\pi(K_j) = \Gamma_j K_j$, with $\Gamma_j > 0$ for all j and

$$E(X) = \sum_{j=1}^{d_{\mathcal{H}}^2} K_j X K_j^* ,$$

for all $X \in \mathcal{B}(\mathcal{H})$. Using the characterisation of $E^\dagger$ from theorem 1.3.1, one obtains

$$E^\dagger(X) = E(X\pi^{-1})\pi = \sum_{j=1}^{d_{\mathcal{H}}^2} K_j X (\Delta_\pi(K_j))^* = \sum_{j=1}^{d_{\mathcal{H}}^2} \Gamma_j K_j X K_j^* .$$

Now defining $V \equiv \sum_{j=1}^{d_{\mathcal{K}}} |j\rangle \otimes K_j^*$ and $\Gamma \equiv \sum_{j=1}^{d_{\mathcal{K}}} \Gamma_j |j\rangle\langle j|$, where $(|j\rangle)_{j=1}^{d\mathcal{K}}$ is an orthonormal basis of $\mathcal{K} = \mathbb{C}^{d_{\mathcal{H}}^2}$ completes the proof of the first claim.
For the second claim, observe that $\Delta_\pi^n(K_j) = \Gamma_j^n K_j$ for all $n \in \mathbb{N}$. By the Stone-Weierstraß theorem, this implies that $f(\Delta_\pi)(K_j) = f(\Gamma_j)K_j$ for any continuous function $f : \mathbb{R}_+ \to \mathbb{R}_+$. Thus,

$$\mathbb{1}_{\mathcal{K}} \otimes \pi^{-p} V \pi^p = \sum_{j=1}^{d_{\mathcal{K}}} |j\rangle \otimes (\Delta_\pi^p(K_j))^* = \sum_{j=1}^{d_{\mathcal{K}}} \Gamma_j^p |j\rangle \otimes K_j^* = \Gamma^p \otimes \mathbb{1}_{\mathcal{H}} V . \tag{3.45}$$

□

Given a characterisation as above (i.e., an isometry V and positive-definite Γ that satisfy the detailed properties) we can proof a slightly modified and generalised versions of Lemma 11 and Proposition 12 from [GR24a], particularly eq. (2.14). Unlike the setting in [GR24a], $E^\dagger$ is not unital, and we cannot apply the Kadison-Choi-Schwarz inequality [Kad52; Cho74] directly. Instead, we invoke its generalisation: the Lieb-Ruskai inequality [LR74] in the form of [Car22, Corollary 1.38] and get the following lemma:

Lemma 3.5.2 *Let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ be a conditional expectation onto the von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$, with $\pi = E^\dagger(\iota) > 0$. Then for any $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$ and $\Gamma : \mathcal{K} \to \mathcal{K}$ as in theorem 3.5.1, and for any $X \in \mathcal{B}(\mathcal{H})$ with $E^\dagger(X) = X$, we have*

$$[\Gamma \otimes X, VV^*] = 0 .$$

If in addition $X \geq 0$, then

$$\Gamma \otimes X \geq VV^*(\Gamma \otimes X)VV^* . \tag{3.46}$$

Proof. By [Car22, Corollary 1.38], if $\Psi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is such that $\mathrm{id}_2 \otimes \Psi$ is positive, then for all $Z, C \in \mathcal{B}(\mathcal{H})$ with $C > 0$, it holds that

$$\Psi(Z^* C^{-1} Z) \geq \Psi(Z)^* \Psi(C)^{-1} \Psi(Z)\,.$$

In our setting, π and X are fixed points of $E^\dagger$, which is hermiticity-preserving. Thus,

$$E^\dagger(X^* \pi^{-1} X) \geq E^\dagger(X)^* \pi^{-1} E^\dagger(X) = X^* \pi^{-1} X\,.$$

Since $E^\dagger$ is also trace-preserving, the inequality must be an equality. Substituting the expression for $E^\dagger$ from theorem 3.5.1 and using eq. (3.45), we obtain:

$$\begin{aligned} V^*(\mathbb{1}_\mathcal{K} \otimes (X^* \pi^{-1} X))V &= E^\dagger(X^* \pi^{-1} X) \\ &= E^\dagger(X^*) \pi^{-1} E^\dagger(X) \\ &= V^*(\Gamma \otimes X^*) V \pi^{-1} V^* (\Gamma \otimes X) V \\ &= V^*(\Gamma^{1/2} \otimes (X')^*) P (\Gamma^{1/2} \otimes X') V\,, \end{aligned}$$

where $P \equiv VV^*$ and $X' \equiv \pi^{-1/2} X$. Applying $V \cdot V^*$, we find:

$$P(\Gamma^{1/2} \otimes X')^* (\Gamma^{1/2} \otimes X') P = P(\Gamma^{1/2} \otimes X')^* P (\Gamma^{1/2} \otimes X') P\,.$$

This implies:

$$((\mathbb{1} - P)(\Gamma^{1/2} \otimes X') P)^* ((\mathbb{1} - P)(\Gamma^{1/2} \otimes X') P) = 0\,,$$

so $(\mathbb{1} - P)(\Gamma^{1/2} \otimes X')P = 0$. Using eq. (3.45) again, we conclude:

$$(\Gamma \otimes X) P = P(\Gamma \otimes X) P\,.$$

Repeating the argument with X replaced by X^* and taking adjoints, we obtain $P(\Gamma \otimes X) = P(\Gamma \otimes X)P$, completing the proof of the commutation relation.
If $X \geq 0$, then so is $\Gamma \otimes X$, hence by the commutativity

$$P(\Gamma \otimes X) P = (\Gamma \otimes X)^{1/2} P^2 (\Gamma \otimes X)^{1/2} \leq \Gamma \otimes X$$

follows from $P^2 = P$ and $P \leq \mathbb{1}$, completing the proof of the claim. ☐

This now enables us to provide alternative expressions for both the relative entropy and the hypothesis testing relative entropy resource measures, expressed in terms of the components of any Stinespring dilation of $E^\dagger$, provided that the isometry and the operator Γ satisfy the conditions in theorem 3.5.1.

Lemma 3.5.3 *Let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ be a conditional expectation onto the von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$, with $\pi = E^\dagger(\iota) > 0$. Then, for any isometry $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}$ and $\Gamma : \mathcal{K} \to \mathcal{K}$ as in theorem 3.5.1, the following holds for all $\rho \in \mathcal{S}(\mathcal{H})$:*

$$D(\rho \| E^\dagger(\rho)) = \inf_{\sigma \in \mathcal{S}(\mathcal{H})} D(\rho \| E^\dagger(\sigma)) = \inf_{\sigma \in \mathcal{S}(\mathcal{H})} D(V\rho V^* \| \Gamma \otimes \sigma) = D(V \rho V^* \| \Gamma \otimes \mathrm{tr}_\mathcal{K}[V\rho V^*])\,, \quad (3.47)$$

and for all $\varepsilon \in (0,1)$,

$$\inf_{\sigma \in \mathcal{S}(\mathcal{H})} D_h^\varepsilon(\rho \| E^\dagger(\sigma)) = \inf_{\sigma \in \mathcal{S}(\mathcal{H})} D_h^\varepsilon(V\rho V^* \| \Gamma \otimes \sigma)\,.$$

Proof. We begin by establishing the outermost equalities, which follow directly from the chain rule (eq. (1.20)). For any $\sigma \in \mathcal{S}(\mathcal{H})$, one has

$$D(\rho \| E^\dagger(\sigma)) = D(\rho \| E^\dagger(\rho)) + D(E^\dagger(\rho) \| E^\dagger(\sigma))\,.$$

Taking now the infimum over σ yields the first equality in eq. (3.47). The rightmost equality follows analogously after normalising both sides by $\mathrm{Tr}[\Gamma]$ and selecting the conditional expectation $E'(\cdot) = \mathrm{tr}_\mathcal{K}[\gamma^{1/2} \otimes \mathbb{1}_\mathcal{H}(\,\cdot\,)\gamma^{1/2} \otimes \mathbb{1}_\mathcal{H}]$, whose HS-adjoint is given by $(E')^\dagger(\cdot) = \gamma \otimes \mathrm{tr}_\mathcal{K}[\,\cdot\,]$, where $\gamma = \Gamma / \mathrm{Tr}[\Gamma]$.

The remaining two equalities may be proven simultaneously by introducing the notation $\mathbb{D} = D, D_h^\varepsilon$. In fact, the proof extends beyond the relative entropy and hypothesis-testing entropy to a broader class of divergences, such as the SR, PR, and GR divergences and even their smooth variants.
By the isometric invariance of the divergence $\mathbb{D}$, and letting $P = VV^*$, we obtain

$$\begin{aligned}\inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(\rho\|E^\dagger(\sigma)) &= \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(\rho\|V^*\Gamma\otimes\sigma V)\\ &= \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|VV^*\Gamma\otimes\sigma VV^*)\\ &= \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|P\Gamma\otimes\sigma P)\,.\end{aligned}$$

Now, since P is an orthogonal projection, we may define the CPTP map

$$\Psi(\cdot) = P(\cdot)P + (\mathbb{1}_{\mathcal{K}\otimes\mathcal{H}} - P)(\cdot)(\mathbb{1}_{\mathcal{K}\otimes\mathcal{H}} - P) = (P(\cdot)P) \oplus ((\mathbb{1}_{\mathcal{K}\otimes\mathcal{H}} - P)(\cdot)(\mathbb{1}_{\mathcal{K}\otimes\mathcal{H}} - P))$$

which projects onto a subspace of $\mathcal{B}(\mathcal{K}\otimes\mathcal{H})$, decomposable into a direct sum. Moreover, since $P = PV = V^*P$, it follows that $\Psi(V\rho V^*) = V\rho V^* = V\rho V^* \oplus 0$. As both divergences satisfies $\mathbb{D}(X\oplus 0\|Y\oplus Y') = \mathbb{D}(X\|Y)$ for arbitrary positive-semidefinite operators X, Y, Y', we find

$$\begin{aligned}\inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|P\Gamma\otimes\sigma P) &= \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}((PV\rho V^*P)\oplus 0\|(P\Gamma\otimes\sigma P)\oplus((\mathbb{1}-P)\Gamma\otimes\sigma(\mathbb{1}-P)))\\ &= \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(\Psi(V\rho V^*)\|\Psi(\Gamma\otimes\sigma))\\ &\le \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|\Gamma\otimes\sigma)\,,\end{aligned}$$

where the inequality follows from the DPI.
We now reverse the inequality by restricting the infimum (1), then applying the fixed point inequality from eq. (3.46) and the antimonotonicity of the divergence in the second argument (2), followed by isometric invariance (3), and finally using the idempotency of $E^\dagger$ (4):

$$\begin{aligned}\inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|\Gamma\otimes\sigma) &\overset{(1)}{\le} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|\Gamma\otimes E^\dagger(\sigma))\\ &\overset{(2)}{\le} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(V\rho V^*\|VV^*\Gamma\otimes E^\dagger(\sigma)VV^*)\\ &\overset{(3)}{=} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(\rho\|V^*\Gamma\otimes E^\dagger(\sigma)V)\\ &= \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(\rho\|(E^\dagger)^2(\sigma))\\ &\overset{(4)}{=} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \mathbb{D}(\rho\|E^\dagger(\sigma))\,.\end{aligned}$$

This completes the proof. □

This result is the cornerstone that enables us to relate our setting to the previously established version of the generalised quantum Stein's lemma involving the second hypotheses

$$\mathcal{S}'_n = \{\gamma^{\otimes n}\otimes\sigma_n : \sigma_n\in\mathcal{S}(\mathcal{H}^{\otimes n})\}\,, \tag{3.48}$$

for a fixed positive-definite $\gamma\in\mathcal{S}(\mathcal{K})$, as presented in [HT16, Corollary 18]. Since the formulation of the generalised Stein's lemma in that reference differs slightly, we now explain how to derive eq. (1.103) from it. Let $U_n : (\mathcal{H}\otimes\mathcal{K})^{\otimes n}\to\mathcal{H}^{\otimes n}\otimes\mathcal{K}^{\otimes n}$ be the sorting unitary and $\rho'\in\mathcal{S}(\mathcal{K}\otimes\mathcal{H})$. Then [HT16, Corollary 18] asserts:

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}D(U_n(\rho')^{\otimes n}U_n^*\|\,\mathcal{S}'_n) = D(\rho'\|\gamma\otimes\operatorname{tr}_{\mathcal{K}}[\rho'])\\ &\quad= \sup\left\{R\in\mathbb{R} : \lim_{n\to\infty}\inf\left\{\alpha_n(U_n^*PU_n;\rho') : 0\le P\le\mathbb{1}, -\frac{1}{n}\log\beta(P;\mathcal{S}'_n)\ge R\right\} = 0\right\}\,,\end{aligned}\tag{3.49}$$

where the notation (i.e., α_n and β_n) follows that of section 1.4.5, and the first equality is a consequence of the chain rule (eq. (1.20)) and the additivity of relative entropy under tensor products.
From this one concludes, that for every $\delta > 0$ and $\varepsilon \in (0,1)$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\inf\left\{\alpha_n(U_n^* P U_n;\rho') : 0 \leq P \leq \mathbb{1}, -\frac{1}{n}\log\beta(P;\mathcal{S}'_n) \geq R^* - \delta\right\} \leq \varepsilon/2\,,$$

where $R^* = \lim_{n\to\infty}\frac{1}{n}D(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n)$. Consequently, for each $n \geq N$ there exists a test operator P with $0 \leq P \leq \mathbb{1}$ satisfying

$$\begin{aligned}\alpha_n(U_n^* P U_n;\rho') &\leq \varepsilon \\ R^* - \delta &\leq -\frac{1}{n}\log\beta_n(U_n^* P U_n;\mathcal{S}'_n)\,.\end{aligned}$$

Together, these imply:

$$R^* - \delta \leq \frac{1}{n}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n)\,,$$

and thus

$$R^* - \delta \leq \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n)\,.$$

Since $\delta > 0$ and $\varepsilon \in (0,1)$ were arbitrary, it follows that

$$R^* = \lim_{n\to\infty}\frac{1}{n}D(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n) \leq \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n)\,.$$

The converse inequality follows from the original quantum Stein's lemma (1), the chain rule, and the additivity of the relative entropy under tensor products (2):

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n) &= \lim_{n\to\infty}\frac{1}{n}\inf_{\sigma_n\in\mathcal{S}(\mathcal{H}^{\otimes n})}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \gamma^{\otimes n}\otimes\sigma_n) \\ &\leq \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \gamma^{\otimes n}\otimes \mathrm{tr}_\mathcal{K}[\rho']^{\otimes n}) \\ &= \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon((\rho')^{\otimes n} \| (\gamma\otimes\mathrm{tr}_\mathcal{K}[\rho'])^{\otimes n}) \\ &\overset{(1)}{=} D(\rho'\|\gamma\otimes\mathrm{tr}_\mathcal{K}[\rho']) \overset{(2)}{=} \lim_{n\to\infty}\frac{1}{n}D(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n)\,.\end{aligned}$$

Hence, in summary, we obtain the identity:

$$\lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n) = \lim_{n\to\infty}\frac{1}{n}D(U_n(\rho')^{\otimes n}U_n^* \| \mathcal{S}'_n) = D(\rho'\|\gamma\otimes\mathrm{tr}_\mathcal{K}[\rho']),, \tag{3.50}$$

placing us in a position to formally state and prove the main result of [GR24b]: the quantum Stein's lemma for subalgebra second hypotheses.

Theorem 3.5.4 *Let $E : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ be a conditional expectation onto the von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$, with $\pi = E^\dagger(\iota) > 0$ (see theorem 1.3.2 for equivalent conditions). For $n \in \mathbb{N}$, define*

$$\mathcal{S}_n = \{(E^\dagger)^{\otimes n}(\sigma) : \sigma \in \mathcal{S}(\mathcal{H}^{\otimes n})\}\,.$$

Then, for all $\varepsilon \in (0,1)$, the following identity holds:

$$D(\rho\|E^\dagger(\rho)) = \lim_{n\to\infty}\frac{1}{n}D(\rho^{\otimes n}\|\mathcal{S}_n) = \lim_{n\to\infty}\frac{1}{n}D_h^\varepsilon(\rho^{\otimes n}\|\mathcal{S}_n)\,.$$

Proof. Let $n \in \mathbb{N}$. Observe that $E^{\otimes n} : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{N}^{\otimes n}$ is a conditional expectation satisfying $(E^\dagger)^{\otimes n}(\iota_{\mathcal{H}^{\otimes n}}) = (E^\dagger(\iota_\mathcal{H}))^{\otimes n} > 0$. We may invoke theorem 3.5.1 to write $E^\dagger(\cdot) = V^*\Gamma\otimes(\cdot)V$, so that

$$(E^\dagger)^{\otimes n}(\cdot) = (V^*)^{\otimes n}\left(U_n^*\left(\Gamma^{\otimes n}\otimes(\cdot)\right)U_n\right)V^{\otimes n} \equiv V_n^*\Gamma_n\otimes(\cdot)V_n\,,$$

where $U_n : (\mathcal{H} \otimes \mathcal{K})^{\otimes n} \to \mathcal{H}^{\otimes n} \otimes \mathcal{K}^{\otimes n}$ denotes the sorting unitaries, as before.
The isometry $V_n : \mathcal{H}^{\otimes n} \to \mathcal{K}^{\otimes n} \otimes \mathcal{H}^{\otimes n}$ and the positive-definite operator $\Gamma_n : \mathcal{K}^{\otimes n} \to \mathcal{K}^{\otimes n}$, satisfy the properties detailed in theorem 3.5.1, inherited from the single-copy isometry V and operator Γ. Hence, by theorem 3.5.3, we obtain:

$$D(\rho^{\otimes n} \| \mathcal{S}_n) = \inf_{\sigma_n \in \mathcal{S}(\mathcal{H}^{\otimes n})} D(V_n \rho^{\otimes n} V_n^* \| \Gamma_n \otimes \sigma_n) = D(U_n (V \rho V^*)^{\otimes n} U_n^* \| \mathcal{S}'_n) + n \log \operatorname{Tr}[\Gamma] \,,$$

and similarly,

$$D_h^\varepsilon(\rho^{\otimes n} \| \mathcal{S}_n) = \inf_{\sigma_n \in \mathcal{S}(\mathcal{H}^{\otimes n})} D_h^\varepsilon(V_n \rho^{\otimes n} V_n^* \| \Gamma_n \otimes \sigma_n) = D_h^\varepsilon(U_n (V \rho V^*)^{\otimes n} U_n^* \| \mathcal{S}'_n) + n \log \operatorname{Tr}[\Gamma] \,,$$

where we define $\gamma = \Gamma / \operatorname{Tr}[\Gamma]$, and $\mathcal{S}'_n$ is as given in eq. (3.48). Applying eq. (3.50) with $\rho' = V \rho V^*$, we may identify the right-hand sides of the above equations after dividing by n and taking the limit $n \to \infty$. We, hence, conclude that

$$\lim_{n\to\infty} \frac{1}{n} D(\rho^{\otimes n} \| \mathcal{S}_n) = \lim_{n\to\infty} \frac{1}{n} D_h^\varepsilon(\rho^{\otimes n} \| \mathcal{S}_n) \,.$$

Finally, by the chain rule (eq. (1.20)) and the additivity of relative entropy under tensor products, we have:

$$\lim_{n\to\infty} \frac{1}{n} D(\rho^{\otimes n} \| \mathcal{S}_n) = \lim_{n\to\infty} \frac{1}{n} D(\rho^{\otimes n} \| (E^\dagger)^{\otimes n}(\rho^{\otimes n})) = D(\rho \| E^\dagger(\rho)) \,.$$

□

Although the above theorem generalises the result of [GR24a]—now encompassing the standard Stein's lemma as a special case for example—we have already discussed in section 1.4.5 that more general versions were subsequently established in [HY24; Lam25]. These works go beyond subalgebra second hypotheses and prove the general quantum Stein's lemma without structural assumptions beyond the ones detailed in section 1.4.5. Nonetheless, the techniques presented here may find utility in broader contexts, motivating the following discussion on open questions and directions for future research.

3.5.1 Open questions and future work

A relatively straightforward extension would be to derive Hoeffding bounds and second-order asymptotics, as carried out in [HT16]. The above embedding of subalgebra resource theories into the framework considered in [HT16] suggests that these results should transfer directly.
Continuing along this path one might try to identify novel applications of subalgebra resource theories that go beyond established examples such as quantum coherence where our techniques as well as formerly mentioned Hoeffding bounds or second-order asymptotics are necessary, hence rendering [HY24; Lam25] insufficient.
Alternatively, one might attempt to extend the above results to more general von Neumann algebraic settings, possibly including infinite-dimensional tracial cases. However, the current proof strategy would require substantial adaptation, as the Stinespring characterisation of $E^\dagger$ provided in theorem 3.5.1 fails in this context—most notably because Γ no longer is tracial nor bounded.

Bibliography

[AF04] R. Alicki and M. Fannes. “Continuity of quantum conditional information”. In: *Journal of Physics A: Mathematical and General* 37.5 (Jan. 2004), p. L55. ISSN: 1361-6447. DOI: 10.1088/0305-4470/37/5/l01.

[AFH09] R Alicki, M Fannes, and M Horodecki. “On thermalization in Kitaev’s 2D model”. In: *Journal of Physics A: Mathematical and Theoretical* 42.6 (Jan. 2009), p. 065303. ISSN: 1751-8121. DOI: 10.1088/1751-8113/42/6/065303.

[AFP21] J. Agredo, F. Fagnola, and D. Poletti. “Gaussian Quantum Markov Semigroups on a One-Mode Fock Space: Irreducibility and Normal Invariant States”. In: *Open Systems & Information Dynamics* 28.01 (Mar. 2021), p. 2150001. ISSN: 1793-7191. DOI: 10.1142/s1230161221500013.

[Alh+24] Álvaro M. Alhambra et al. *Conditional Independence of 1D Gibbs States with Applications to Efficient Learning*. 2024. DOI: 10.48550/ARXIV.2402.18500.

[Ans+21] Anurag Anshu et al. “Sample-efficient learning of interacting quantum systems”. In: *Nature Physics* 17.8 (May 2021), pp. 931–935. ISSN: 1745-2481. DOI: 10.1038/s41567-021-01232-0.

[Ara69] Huzihiro Araki. “Gibbs states of a one dimensional quantum lattice”. In: *Communications in Mathematical Physics* 14.2 (June 1969), pp. 120–157. ISSN: 1432-0916. DOI: 10.1007/bf01645134.

[ASR15] R. Azouit, A. Sarlette, and P. Rouchon. “Convergence and adiabatic elimination for a driven dissipative quantum harmonic oscillator”. In: *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, Dec. 2015, pp. 6447–6453. DOI: 10.1109/cdc.2015.7403235.

[ASR16] Rémi Azouit, Alain Sarlette, and Pierre Rouchon. “Well-posedness and convergence of the Lindblad master equation for a quantum harmonic oscillator with multi-photon drive and damping”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 22.4 (Aug. 2016), pp. 1353–1369. ISSN: 1262-3377. DOI: 10.1051/cocv/2016050.

[Aud+07] K. M. R. Audenaert et al. “Discriminating States: The Quantum Chernoff Bound”. In: *Physical Review Letters* 98.16 (Apr. 2007), p. 160501. ISSN: 1079-7114. DOI: 10.1103/physrevlett.98.160501.

[Aud07] Koenraad M R Audenaert. “A sharp continuity estimate for the von Neumann entropy”. In: *Journal of Physics A: Mathematical and Theoretical* 40.28 (June 2007), pp. 8127–8136. ISSN: 1751-8121. DOI: 10.1088/1751-8113/40/28/s18.

[Aud+08] K. M. R. Audenaert et al. “Asymptotic Error Rates in Quantum Hypothesis Testing”. In: *Communications in Mathematical Physics* 279.1 (Feb. 2008), pp. 251–283. ISSN: 1432-0916. DOI: 10.1007/s00220-008-0417-5.

[Aud+25] Koenraad Audenaert et al. “Continuity bounds for quantum entropies arising from a fundamental entropic inequality”. In: *IEEE Transactions on Information Theory* (2025), pp. 1–1. DOI: 10.1109/TIT.2025.3586478.

[Bak+24] Ainesh Bakshi et al. “Learning Quantum Hamiltonians at Any Temperature in Polynomial Time”. In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. STOC ’24. ACM, June 2024, pp. 1470–1477. DOI: 10.1145/3618260.3649619.

[Bar+24] Ivan Bardet et al. “Entropy Decay for Davies Semigroups of a One Dimensional Quantum Lattice”. In: *Communications in Mathematical Physics* 405.2 (Feb. 2024). ISSN: 1432-0916. DOI: 10.1007/s00220-023-04869-5.

[BC19] Andreas Bluhm and Ángela Capel. "A strengthened data processing inequality for the Belavkin-Staszewski relative entropy". In: *Reviews in Mathematical Physics* 32.02 (Aug. 2019), p. 2050005. ISSN: 1793-6659. DOI: 10.1142/s0129055x20500051.

[BCPH22] Andreas Bluhm, Ángela Capel, and Antonio Pérez-Hernández. "Exponential decay of mutual information for Gibbs states of local Hamiltonians". In: *Quantum* 6 (Feb. 2022), p. 650. ISSN: 2521-327X. DOI: 10.22331/q-2022-02-10-650.

[BCPH24] Andreas Bluhm, Ángela Capel, and Antonio Pérez-Hernández. *Strong decay of correlations for Gibbs states in any dimension.* 2024. DOI: 10.48550/ARXIV.2401.10147.

[BDR20] Salman Beigi, Nilanjana Datta, and Cambyse Rouzé. "Quantum Reverse Hypercontractivity: Its Tensorization and Application to Strong Converses". In: *Communications in Mathematical Physics* 376.2 (May 2020), pp. 753–794. ISSN: 1432-0916. DOI: 10.1007/s00220-020-03750-z.

[Ber+23] Mario Berta et al. "On a gap in the proof of the generalised quantum Steinś lemma and its consequences for the reversibility of quantum resources". In: *Quantum* 7 (Sept. 2023), p. 1103. ISSN: 2521-327X. DOI: 10.22331/q-2023-09-07-1103.

[BG15] Fernando G. S. L. Brandão and Gilad Gour. "Reversible Framework for Quantum Resource Theories". In: *Physical Review Letters* 115.7 (Aug. 2015), p. 070503. ISSN: 1079-7114. DOI: 10.1103/physrevlett.115.070503.

[BG23] Salman Beigi and Milad M. Goodarzi. "Operator-valued Schatten spaces and quantum entropies". In: *Letters in Mathematical Physics* 113.5 (Aug. 2023). ISSN: 1573-0530. DOI: 10.1007/s11005-023-01712-9.

[BK18] Fernando G. S. L. Brandão and Michael J. Kastoryano. "Finite Correlation Length Implies Efficient Preparation of Quantum Thermal States". In: *Communications in Mathematical Physics* 365.1 (May 2018), pp. 1–16. ISSN: 1432-0916. DOI: 10.1007/s00220-018-3150-8.

[BLT25] Mario Berta, Ludovico Lami, and Marco Tomamichel. "Continuity of Entropies via Integral Representations". In: *IEEE Transactions on Information Theory* 71.3 (Mar. 2025), pp. 1896–1908. ISSN: 1557-9654. DOI: 10.1109/tit.2025.3527858.

[Blu+23a] Andreas Bluhm et al. "Continuity of Quantum Entropic Quantities via Almost Convexity". In: *IEEE Transactions on Information Theory* 69.9 (Sept. 2023), pp. 5869–5901. ISSN: 1557-9654. DOI: 10.1109/tit.2023.3277892.

[Blu+23b] Andreas Bluhm et al. "General Continuity Bounds for Quantum Relative Entropies". In: *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, June 2023, pp. 162–167. DOI: 10.1109/isit54713.2023.10206734.

[Blu+24] Andreas Bluhm et al. "Unified Framework for Continuity of Sandwiched Rényi Divergences". In: *Annales Henri Poincaré* (Dec. 2024). ISSN: 1424-0661. DOI: 10.1007/s00023-024-01519-x.

[BLW15] Mario Berta, Marius Lemm, and Mark M. Wilde. "Monotonicity of quantum relative entropy and recoverability". In: *Quantum Information and Computation* 15.15 & 16 (Nov. 2015), pp. 1333–1354. ISSN: 1533-7146. DOI: 10.26421/qic15.15-16-5.

[BP10a] Fernando G. S. L. Brandão and Martin B. Plenio. "A Generalization of Quantum Stein's Lemma". In: *Communications in Mathematical Physics* 295.3 (Feb. 2010), pp. 791–828. ISSN: 1432-0916. DOI: 10.1007/s00220-010-1005-z.

[BP10b] Fernando G. S. L. Brandão and Martin B. Plenio. "A Reversible Theory of Entanglement and its Relation to the Second Law". In: *Communications in Mathematical Physics* 295.3 (Feb. 2010), pp. 829–851. ISSN: 1432-0916. DOI: 10.1007/s00220-010-1003-1.

[BU07] Jean-Christophe Bourin and Mitsuru Uchiyama. "A matrix subadditivity inequality for f(A+B) and f(A)+f(B)". In: *Linear Algebra and its Applications* 423.2-3 (June 2007), pp. 512–518. ISSN: 0024-3795. DOI: 10.1016/j.laa.2007.02.019.

[Bü18] Theo Bühler. *Functional analysis.* Graduate Studies in Mathematics volume 191. Providence, Rhode Island: American Mathematical Society, 2018. 1466 pp. ISBN: 9781470447762.

[Cap+24] Ángela Capel et al. *Quasi-optimal sampling from Gibbs states via non-commutative optimal transport metrics*. 2024. DOI: 10.48550/ARXIV.2412.01732.

[Car22] Eric A. Carlen. *On some convexity and monotonicity inequalities of Elliott Lieb*. 2022. DOI: 10.48550/ARXIV.2202.03591.

[Ces01] Filippo Cesi. "Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields". In: *Probability Theory and Related Fields* 120.4 (Aug. 2001), pp. 569–584. ISSN: 1432-2064. DOI: 10.1007/pl00008792.

[CF98] A.M Chebotarev and F Fagnola. "Sufficient Conditions for Conservativity of Minimal Quantum Dynamical Semigroups". In: *Journal of Functional Analysis* 153.2 (Mar. 1998), pp. 382–404. ISSN: 0022-1236. DOI: 10.1006/jfan.1997.3189.

[Cha+22] Christopher Chamberland et al. "Building a Fault-Tolerant Quantum Computer Using Concatenated Cat Codes". In: *PRX Quantum* 3.1 (Feb. 2022), p. 010329. ISSN: 2691-3399. DOI: 10.1103/prxquantum.3.010329.

[Che03] A. M. Chebotarev. "A Priori Estimates for Quantum Dynamical Semigroups". In: *Theoretical and Mathematical Physics* 134.2 (2003), pp. 160–165. ISSN: 0040-5779. DOI: 10.1023/a:1022219919058.

[Che+21] Song Cheng et al. "Simulating noisy quantum circuits with matrix product density operators". In: *Physical Review Research* 3.2 (Apr. 2021), p. 023005. ISSN: 2643-1564. DOI: 10.1103/physrevresearch.3.023005.

[Che+23] Chi-Fang Chen et al. *Quantum Thermal State Preparation*. 2023. DOI: 10.48550/ARXIV.2303.18224.

[Cho74] Man-Duen Choi. "A schwarz inequality for positive linear maps on C^*-algebras". In: *Illinois Journal of Mathematics* 18.4 (Dec. 1974). ISSN: 0019-2082. DOI: 10.1215/ijm/1256051007.

[CKG23] Chi-Fang Chen, Michael J. Kastoryano, and András Gilyén. *An efficient and exact noncommutative quantum Gibbs sampler*. 2023. DOI: 10.48550/ARXIV.2311.09207.

[CLPG18] Ángela Capel, Angelo Lucia, and David Pérez-García. "Quantum conditional relative entropy and quasi-factorization of the relative entropy". In: *Journal of Physics A: Mathematical and Theoretical* 51.48 (Nov. 2018), p. 484001. ISSN: 1751-8121. DOI: 10.1088/1751-8121/aae4cf.

[CM17] Eric A. Carlen and Jan Maas. "Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance". In: *Journal of Functional Analysis* 273.5 (Sept. 2017), pp. 1810–1869. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2017.05.003.

[Cov06] Thomas M. Cover. *Elements of information theory*. Second edition. Hoboken, N.J: Wiley-Interscience, 2006. 1748 pp. ISBN: 9780471748823.

[CRF20] Ángela Capel, Cambyse Rouzé, and Daniel Stilck França. *The modified logarithmic Sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions*. 2020. DOI: 10.48550/ARXIV.2009.11817.

[CS07] Raffaella Carbone and Emanuela Sasso. "Hypercontractivity for a quantum Ornstein-Uhlenbeck semigroup". In: *Probability Theory and Related Fields* 140.3-4 (May 2007), pp. 505–522. ISSN: 1432-2064. DOI: 10.1007/s00440-007-0073-2.

[CV18] Eric A. Carlen and Anna Vershynina. "Recovery and the Data Processing Inequality for Quasi-Entropies". In: *IEEE Transactions on Information Theory* 64.10 (Oct. 2018), pp. 6929–6938. ISSN: 1557-9654. DOI: 10.1109/tit.2018.2812038.

[CV20] Eric A Carlen and Anna Vershynina. "Recovery map stability for the data processing inequality". In: *Journal of Physics A: Mathematical and Theoretical* 53.3 (Jan. 2020), p. 035204. ISSN: 1751-8121. DOI: 10.1088/1751-8121/ab5ab7.

[Dat09] Nilanjana Datta. "Min- and Max-Relative Entropies and a New Entanglement Monotone". In: *IEEE Transactions on Information Theory* 55.6 (June 2009), pp. 2816–2826. ISSN: 0018-9448. DOI: 10.1109/tit.2009.2018325.

[Dav72] E. B. Davies. “Diffusion for weakly coupled quantum oscillators”. In: *Communications in Mathematical Physics* 27.4 (Dec. 1972), pp. 309–325. ISSN: 1432-0916. DOI: 10.1007/bf01645518.

[Dav73] E. B. Davies. “The harmonic oscillator in a heat bath”. In: *Communications in Mathematical Physics* 33.3 (Sept. 1973), pp. 171–186. ISSN: 1432-0916. DOI: 10.1007/bf01667915.

[Dav74] E. B. Davies. “Markovian master equations”. In: *Communications in Mathematical Physics* 39.2 (June 1974), pp. 91–110. ISSN: 1432-0916. DOI: 10.1007/bf01608389.

[Dav76] E. B. Davies. “Markovian master equations. II”. In: *Mathematische Annalen* 219.2 (June 1976), pp. 147–158. ISSN: 1432-1807. DOI: 10.1007/bf01351898.

[Dav77] E.B. Davies. “Quantum dynamical semigroups and the neutron diffusion equation”. In: *Reports on Mathematical Physics* 11.2 (Apr. 1977), pp. 169–188. ISSN: 0034-4877. DOI: 10.1016/0034-4877(77)90059-3.

[DLL25a] Andreas Deuchert, Jonas Lampart, and Marius Lemm. *Dynamics and equilibrium states of infinite systems of lattice bosons*. 2025. DOI: 10.48550/ARXIV.2505.13170.

[DLL25b] Zhiyan Ding, Bowen Li, and Lin Lin. “Efficient Quantum Gibbs Samplers with Kubo-Martin-Schwinger Detailed Balance Condition”. In: *Communications in Mathematical Physics* 406.3 (Feb. 2025). ISSN: 1432-0916. DOI: 10.1007/s00220-025-05235-3.

[DS85] R. L. Dobrushin and S. B. Shlosman. “Completely Analytical Gibbs Fields”. In: *Statistical Physics and Dynamical Systems*. Birkhäuser Boston, 1985, pp. 371–403. ISBN: 9781489966537. DOI: 10.1007/978-1-4899-6653-7_21.

[DS87] R. L. Dobrushin and S. B. Shlosman. “Completely analytical interactions: Constructive description”. In: *Journal of Statistical Physics* 46.5-6 (Mar. 1987), pp. 983–1014. ISSN: 1572-9613. DOI: 10.1007/bf01011153.

[DW05] Igor Devetak and Andreas Winter. “Distillation of secret key and entanglement from quantum states”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 461.2053 (Jan. 2005), pp. 207–235. ISSN: 1471-2946. DOI: 10.1098/rspa.2004.1372.

[EW07] J. Eisert and M. M. Wolf. “Gaussian Quantum Channels”. In: *Quantum Information with Continuous Variables of Atoms and Light*. Published by Imperial College Press and distributed by World Scientific Publishing Co., Feb. 2007, pp. 23–42. ISBN: 9781860948169. DOI: 10.1142/9781860948169_0002.

[Fag18] Franco Fagnola. “Quantum Markov semigroups”. In: *Proyecciones (Antofagasta)* 18.3 (Apr. 2018), pp. 29–74. ISSN: 0717-6279. DOI: 10.22199/s07160917.1999.0003.00004.

[Fan+23] Marco Fanizza et al. *Learning finitely correlated states: stability of the spectral reconstruction*. 2023. DOI: 10.48550/ARXIV.2312.07516.

[Fan73] M. Fannes. “A continuity property of the entropy density for spin lattice systems”. In: *Communications in Mathematical Physics* 31.4 (Dec. 1973), pp. 291–294. ISSN: 1432-0916. DOI: 10.1007/bf01646490.

[Fek23] M. Fekete. “Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten”. In: *Mathematische Zeitschrift* 17.1 (Dec. 1923), pp. 228–249. ISSN: 1432-1823. DOI: 10.1007/bf01504345.

[FFS24] Hamza Fawzi, Omar Fawzi, and Samuel O. Scalet. “Certified algorithms for equilibrium states of local quantum Hamiltonians”. In: *Nature Communications* 15.1 (Aug. 2024). ISSN: 2041-1723. DOI: 10.1038/s41467-024-51592-3.

[FG99] C.A. Fuchs and J. van de Graaf. “Cryptographic distinguishability measures for quantum-mechanical states”. In: *IEEE Transactions on Information Theory* 45.4 (May 1999), pp. 1216–1227. ISSN: 0018-9448. DOI: 10.1109/18.761271.

[FM19] Franco Fagnola and Carlos M. Mora. “Basic Properties of a Mean Field Laser Equation”. In: *Open Systems & Information Dynamics* 26.03 (Sept. 2019), p. 1950015. ISSN: 1793-7191. DOI: 10.1142/s123016121950015x.

[Gao+22] Li Gao et al. *Relative entropy decay and complete positivity mixing time.* 2022. DOI: 10.48550/ARXIV.2209.11684.

[GGK94] L. Gilles, B. M. Garraway, and P. L. Knight. "Generation of nonclassical light by dissipative two-photon processes". In: *Physical Review A* 49.4 (Apr. 1994), pp. 2785–2799. ISSN: 1094-1622. DOI: 10.1103/physreva.49.2785.

[GKS76] Vittorio Gorini, Andrzej Kossakowski, and E. C. G. Sudarshan. "Completely positive dynamical semigroups of N-level systems". In: *Journal of Mathematical Physics* 17.5 (May 1976), pp. 821–825. ISSN: 1089-7658. DOI: 10.1063/1.522979.

[GM19] Jérémie Guillaud and Mazyar Mirrahimi. "Repetition Cat Qubits for Fault-Tolerant Quantum Computation". In: *Physical Review X* 9.4 (Dec. 2019), p. 041053. ISSN: 2160-3308. DOI: 10.1103/physrevx.9.041053.

[GMR24] Paul Gondolf, Tim Möbus, and Cambyse Rouzé. "Energy preserving evolutions over Bosonic systems". In: *Quantum* 8 (Dec. 2024), p. 1551. ISSN: 2521-327X. DOI: 10.22331/q-2024-12-04-1551.

[Gou24] Gilad Gour. *Resources of the Quantum World.* 2024. DOI: 10.48550/ARXIV.2402.05474.

[GR22] Li Gao and Cambyse Rouzé. "Complete Entropic Inequalities for Quantum Markov Chains". In: *Archive for Rational Mechanics and Analysis* 245.1 (May 2022), pp. 183–238. ISSN: 1432-0673. DOI: 10.1007/s00205-022-01785-1.

[GR24a] Li Gao and Mizanur Rahaman. *Generalized Stein's lemma and asymptotic equipartition property for subalgebra entropies.* 2024. DOI: 10.48550/ARXIV.2401.03090.

[GR24b] Paul Gondolf and Cambyse Rouzé. "The Generalised Quantum Stein Lemma for Subalgebra Entropies". Unpublished notes. 2024.

[GS08] Gilad Gour and Robert W Spekkens. "The resource theory of quantum reference frames: manipulations and monotones". In: *New Journal of Physics* 10.3 (Mar. 2008), p. 033023. ISSN: 1367-2630. DOI: 10.1088/1367-2630/10/3/033023.

[HD22] Eric Patrick Hanson and Nilanjana Datta. "Entropies, majorization flow, and continuity bounds". In: *The Physics and Mathematics of Elliott Lieb.* EMS Press, June 2022, pp. 473–514. ISBN: 9783985475216. DOI: 10.4171/90-1/20.

[HHH96] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. "Separability of mixed states: necessary and sufficient conditions". In: *Physics Letters A* 223.1-2 (Nov. 1996), pp. 1–8. ISSN: 0375-9601. DOI: 10.1016/s0375-9601(96)00706-2.

[HHW10] Teiko Heinosaari, Alexander S. Holevo, and Michael M. Wolf. "The semigroup structure of gaussian channels". In: *Quantum Info. Comput.* 10.7 (July 2010), pp. 619–635. ISSN: 1533-7146. URL: https://arxiv.org/abs/0909.0408.

[HKT24] Jeongwan Haah, Robin Kothari, and Ewin Tang. "Learning quantum Hamiltonians from high-temperature Gibbs states and real-time evolutions". In: *Nature Physics* 20.6 (Mar. 2024), pp. 1027–1031. ISSN: 1745-2481. DOI: 10.1038/s41567-023-02376-x.

[HM17] Fumio Hiai and Milán Mosonyi. "Different quantum f-divergences and the reversibility of quantum operations". In: *Reviews in Mathematical Physics* 29.07 (Aug. 2017), p. 1750023. ISSN: 1793-6659. DOI: 10.1142/s0129055x17500234.

[HP12] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups.* Rev. ed (Online-Ausg.) AMS ebook collection. Providence: American Mathematical Society, 2012. 1808 pp. ISBN: 9780821833957.

[HP91] Fumio Hiai and Dénes Petz. "The proper formula for relative entropy and its asymptotics in quantum probability". In: *Communications in Mathematical Physics* 143.1 (Dec. 1991), pp. 99–114. ISSN: 1432-0916. DOI: 10.1007/bf02100287.

[HT16] Masahito Hayashi and Marco Tomamichel. "Correlation detection and an operational interpretation of the Rényi mutual information". In: *Journal of Mathematical Physics* 57.10 (Oct. 2016). ISSN: 1089-7658. DOI: 10.1063/1.4964755.

[HT24] Christoph Hirche and Marco Tomamichel. "Quantum Rényi and f-Divergences from Integral Representations". In: *Communications in Mathematical Physics* 405.9 (Aug. 2024). ISSN: 1432-0916. DOI: 10.1007/s00220-024-05087-3.

[HV17] Jutho Haegeman and Frank Verstraete. "Diagonalizing Transfer Matrices and Matrix Product Operators: A Medley of Exact and Computational Methods". In: *Annual Review of Condensed Matter Physics* 8.1 (Mar. 2017), pp. 355–406. ISSN: 1947-5462. DOI: 10.1146/annurev-conmatphys-031016-025507.

[HY24] Masahito Hayashi and Hayata Yamasaki. *Generalized Quantum Stein's Lemma and Second Law of Quantum Resource Theories.* 2024. DOI: 10.48550/ARXIV.2408.02722.

[Jun+18] Marius Junge et al. "Universal Recovery Maps and Approximate Sufficiency of Quantum Relative Entropy". In: *Annales Henri Poincaré* 19.10 (Aug. 2018), pp. 2955–2978. ISSN: 1424-0661. DOI: 10.1007/s00023-018-0716-0.

[Jun96] Marius Junge. "Factorization Theory for Spaces of Operators". Habilitation Thesis. Kiel University, 1996.

[KAA21] Tomotaka Kuwahara, Álvaro M. Alhambra, and Anurag Anshu. "Improved Thermal Area Law and Quasilinear Time Algorithm for Quantum Gibbs States". In: *Physical Review X* 11.1 (Mar. 2021), p. 011047. ISSN: 2160-3308. DOI: 10.1103/physrevx.11.011047.

[Kad52] Richard V. Kadison. "A Generalized Schwarz Inequality and Algebraic Invariants for Operator Algebras". In: *The Annals of Mathematics* 56.3 (Nov. 1952), p. 494. ISSN: 0003-486X. DOI: 10.2307/1969657.

[KB16] Michael J. Kastoryano and Fernando G. S. L. Brandão. "Quantum Gibbs Samplers: The Commuting Case". In: *Communications in Mathematical Physics* 344.3 (May 2016), pp. 915–957. ISSN: 1432-0916. DOI: 10.1007/s00220-016-2641-8.

[KB19] Kohtaro Kato and Fernando G. S. L. Brandão. "Quantum Approximate Markov Chains are Thermal". In: *Communications in Mathematical Physics* 370.1 (June 2019), pp. 117–149. ISSN: 1432-0916. DOI: 10.1007/s00220-019-03485-6.

[KGR05] B. Kraus, N. Gisin, and R. Renner. "Lower and Upper Bounds on the Secret-Key Rate for Quantum Key Distribution Protocols Using One-Way Classical Communication". In: *Physical Review Letters* 95.8 (Aug. 2005), p. 080501. ISSN: 1079-7114. DOI: 10.1103/physrevlett.95.080501.

[KKBa20] Tomotaka Kuwahara, Kohtaro Kato, and Fernando G. S. L. Brandão. "Clustering of Conditional Mutual Information for Quantum Gibbs States above a Threshold Temperature". In: *Physical Review Letters* 124.22 (June 2020), p. 220601. ISSN: 1079-7114. DOI: 10.1103/physrevlett.124.220601.

[KL51] S. Kullback and R. A. Leibler. "On Information and Sufficiency". In: *The Annals of Mathematical Statistics* 22.1 (Mar. 1951), pp. 79–86. ISSN: 0003-4851. DOI: 10.1214/aoms/1177729694.

[KLCT16] Anna Kómár, Olivier Landon-Cardinal, and Kristan Temme. "Necessity of an energy barrier for self-correction of Abelian quantum doubles". In: *Physical Review A* 93.5 (May 2016), p. 052337. ISSN: 2469-9934. DOI: 10.1103/physreva.93.052337.

[Koc+25] Jan Kochanowski et al. "Rapid Thermalization of Dissipative Many-Body Dynamics of Commuting Hamiltonians". In: *Communications in Mathematical Physics* 406.8 (July 2025). ISSN: 1432-0916. DOI: 10.1007/s00220-025-05353-y.

[KT13] Michael J. Kastoryano and Kristan Temme. "Quantum logarithmic Sobolev inequalities and rapid mixing". In: *Journal of Mathematical Physics* 54.5 (May 2013). ISSN: 1089-7658. DOI: 10.1063/1.4804995.

[Kuw24] Tomotaka Kuwahara. *Clustering of conditional mutual information and quantum Markov structure at arbitrary temperatures.* 2024. DOI: 10.48550/ARXIV.2407.05835.

[KVS24] Tomotaka Kuwahara, Tan Van Vu, and Keiji Saito. "Effective light cone and digital quantum simulation of interacting bosons". In: *Nature Communications* 15.1 (Mar. 2024). ISSN: 2041-1723. DOI: 10.1038/s41467-024-46501-7.

[KW20] Sumeet Khatri and Mark M. Wilde. *Principles of Quantum Communication Theory: A Modern Approach*. 2020. DOI: 10.48550/ARXIV.2011.04672.

[Lam25] Ludovico Lami. "A Solution of the Generalized Quantum Stein's Lemma". In: *IEEE Transactions on Information Theory* 71.6 (June 2025), pp. 4454–4484. ISSN: 1557-9654. DOI: 10.1109/tit.2025.3543610.

[Leg+15] Z. Leghtas et al. "Confining the state of light to a quantum manifold by engineered two-photon loss". In: *Science* 347.6224 (Feb. 2015), pp. 853–857. ISSN: 1095-9203. DOI: 10.1126/science.aaa2085.

[Lie73] Elliott H Lieb. "Convex trace functions and the Wigner-Yanase-Dyson conjecture". In: *Advances in Mathematics* 11.3 (Dec. 1973), pp. 267–288. ISSN: 0001-8708. DOI: 10.1016/0001-8708(73)90011-x.

[Lin25] Lin Lin. "Dissipative preparation of many-body quantum states: Toward practical quantum advantage". In: *APL Computational Physics* 1.1 (Sept. 2025), p. 010901. ISSN: 3066-0017. DOI: 10.1063/5.0283315.

[Lin76] G. Lindblad. "On the generators of quantum dynamical semigroups". In: *Communications in Mathematical Physics* 48.2 (June 1976), pp. 119–130. ISSN: 1432-0916. DOI: 10.1007/bf01608499.

[LP61] Gunter Lumer and R. S. Phillips. "Dissipative operators in a Banach space". In: *Pacific Journal of Mathematics* 11.2 (June 1961), pp. 679–698. ISSN: 0030-8730. DOI: 10.2140/pjm.1961.11.679.

[LPGPH23] Angelo Lucia, David Pérez-García, and Antonio Pérez-Hernández. "Thermalization in Kitaev's quantum double models via tensor network techniques". In: *Forum of Mathematics, Sigma* 11 (2023). ISSN: 2050-5094. DOI: 10.1017/fms.2023.98.

[LPGPH25] Angelo Lucia, David Pérez-García, and Antonio Pérez-Hernández. *Spectral Gap Bounds for Quantum Markov Semigroups via Correlation Decay*. 2025. DOI: 10.48550/ARXIV.2505.08991.

[LR74] Elliott H. Lieb and Mary Beth Ruskai. "Some operator inequalities of the schwarz type". In: *Advances in Mathematics* 12.2 (Feb. 1974), pp. 269–273. ISSN: 0001-8708. DOI: 10.1016/s0001-8708(74)80004-6.

[LRB05] Marco Lenci and Luc Rey-Bellet. "Large Deviations in Quantum Lattice Systems: One-Phase Region". In: *Journal of Statistical Physics* 119.3-4 (May 2005), pp. 715–746. ISSN: 1572-9613. DOI: 10.1007/s10955-005-3015-3.

[LRZ23] Marius Lemm, Carla Rubiliani, and Jingxuan Zhang. *On the microscopic propagation speed of long-range quantum many-body systems*. 2023. DOI: 10.48550/ARXIV.2310.14896.

[LRZ25] Marius Lemm, Carla Rubiliani, and Jingxuan Zhang. *On the quantum dynamics of long-ranged Bose-Hubbard Hamiltonians*. 2025. DOI: 10.48550/ARXIV.2505.01786.

[Mar99] Fabio Martinelli. "Lectures on Glauber Dynamics for Discrete Spin Models". In: *Lectures on Probability Theory and Statistics*. Springer Berlin Heidelberg, 1999, pp. 93–191. ISBN: 9783540481157. DOI: 10.1007/978-3-540-48115-7_2.

[MD22] Ashutosh Marwah and Frédéric Dupuis. "Uniform continuity bound for sandwiched Rényi conditional entropy". In: *Journal of Mathematical Physics* 63.5 (May 2022). ISSN: 1089-7658. DOI: 10.1063/5.0088507.

[Mir+14] Mazyar Mirrahimi et al. "Dynamically protected cat-qubits: a new paradigm for universal quantum computation". In: *New Journal of Physics* 16.4 (Apr. 2014), p. 045014. ISSN: 1367-2630. DOI: 10.1088/1367-2630/16/4/045014.

[ML+13] Martin Müller-Lennert et al. "On quantum Rényi entropies: A new generalization and some properties". In: *Journal of Mathematical Physics* 54.12 (Dec. 2013). ISSN: 1089-7658. DOI: 10.1063/1.4838856.

[Mö25] Tim Möbus. "Limits of Approximation of Quantum Markov Semigroups in Infinite-Dimensional Systems". en. PhD thesis. Technische Universität München, 2025, p. 308. URL: https://mediatum.ub.tum.de/1764923.

[Mö+25] Tim Möbus et al. *Heisenberg-limited Hamiltonian learning continuous variable systems via engineered dissipation.* 2025. DOI: 10.48550/ARXIV.2506.00606.

[Nag00] Rainer Nagel. *One-parameter semigroups for linear evolution equations.* Graduate texts in mathematics 194. New York: Springer, 2000. 1586 pp. ISBN: 0387226427.

[NC12] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information.* 10th anniversary ed., repr. Cambridge [u.a.]: Cambridge University Press, 2012. 676 pp. ISBN: 9781107002173.

[Neu96] John von Neumann. *Mathematische Grundlagen der Quantenmechanik.* Springer Berlin Heidelberg, 1996. ISBN: 9783642614095. DOI: 10.1007/978-3-642-61409-5.

[NO00] H. Nagaoka and T. Ogawa. "Strong converse and Stein's lemma in quantum hypothesis testing". In: *IEEE Transactions on Information Theory* 46.7 (2000), pp. 2428–2433. ISSN: 0018-9448. DOI: 10.1109/18.887855.

[NS09] Michael Nussbaum and Arleta Szkoła. "The Chernoff lower bound for symmetric quantum hypothesis testing". In: *The Annals of Statistics* 37.2 (Apr. 2009). ISSN: 0090-5364. DOI: 10.1214/08-aos593.

[OH04] T. Ogawa and M. Hayashi. "On Error Exponents in Quantum Hypothesis Testing". In: *IEEE Transactions on Information Theory* 50.6 (June 2004), pp. 1368–1372. ISSN: 0018-9448. DOI: 10.1109/tit.2004.828155.

[Ono+23] Emilio Onorati et al. *Provably Efficient Learning of Phases of Matter via Dissipative Evolutions.* 2023. DOI: 10.48550/ARXIV.2311.07506.

[OP93] Masanori Ohya and Dénes Petz. *Quantum entropy and its use.* Texts and monographs in physics. Berlin: Springer, 1993. 335 pp. ISBN: 0387548815.

[Per96] Asher Peres. "Separability Criterion for Density Matrices". In: *Physical Review Letters* 77.8 (Aug. 1996), pp. 1413–1415. ISSN: 1079-7114. DOI: 10.1103/physrevlett.77.1413.

[Pet88] Dénes Petz. "Sufficiency of Channels over von Neumann Algebras". In: *The Quarterly Journal of Mathematics* 39.1 (1988), pp. 97–108. ISSN: 1464-3847. DOI: 10.1093/qmath/39.1.97.

[PGPH23] David Pérez-García and Antonio Pérez-Hernández. "Locality Estimates for Complex Time Evolution in 1D". In: *Communications in Mathematical Physics* 399.2 (Jan. 2023), pp. 929–970. ISSN: 1432-0916. DOI: 10.1007/s00220-022-04573-w.

[Pis18] Gilles Pisier. "Non-commutative vector valued L_p-spaces and completely p-summing maps". In: *Astérisque* (Nov. 2018). ISSN: 2492-5926. DOI: 10.24033/ast.402.

[Pur+19] Shruti Puri et al. "Stabilized Cat in a Driven Nonlinear Cavity: A Fault-Tolerant Error Syndrome Detector". In: *Physical Review X* 9.4 (Oct. 2019), p. 041009. ISSN: 2160-3308. DOI: 10.1103/physrevx.9.041009.

[Qin+24] Zhen Qin et al. "Quantum State Tomography for Matrix Product Density Operators". In: *IEEE Transactions on Information Theory* 70.7 (July 2024), pp. 5030–5056. ISSN: 1557-9654. DOI: 10.1109/tit.2024.3360951.

[Rén61] Alfréd Rényi. "On measures of entropy and information". In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics.* Vol. 4. University of California Press. 1961, pp. 547–562.

[RFA24] Cambyse Rouzé, Daniel Stilck França, and Álvaro M. Alhambra. *Optimal quantum algorithm for Gibbs state preparation.* 2024. DOI: 10.48550/ARXIV.2411.04885.

[RFA25] Cambyse Rouzé, Daniel Stilck França, and Álvaro M. Alhambra. "Efficient Thermalization and Universal Quantum Computing with Quantum Gibbs Samplers". In: *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. STOC '25. ACM, June 2025, pp. 1488–1495. DOI: 10.1145/3717823.3718268.

[Rou24] Cambyse Rouzé. "Logarithmic Sobolev Inequalities for Finite Dimensional Quantum Markov Chains". In: *Optimal Transport on Quantum Structures*. Springer Nature Switzerland, 2024, pp. 263–321. ISBN: 9783031504662. DOI: 10.1007/978-3-031-50466-2_6.

[SBT16] David Sutter, Mario Berta, and Marco Tomamichel. "Multivariate Trace Inequalities". In: *Communications in Mathematical Physics* 352.1 (Oct. 2016), pp. 37–58. ISSN: 1432-0916. DOI: 10.1007/s00220-016-2778-5.

[Sca+21] Samuel O. Scalet et al. "Computable Rényi mutual information: Area laws and correlations". In: *Quantum* 5 (Sept. 2021), p. 541. ISSN: 2521-327X. DOI: 10.22331/q-2021-09-14-541.

[Sch95] Benjamin Schumacher. "Quantum coding". In: *Physical Review A* 51.4 (Apr. 1995), pp. 2738–2747. ISSN: 1094-1622. DOI: 10.1103/physreva.51.2738.

[Sha48] C. E. Shannon. "A Mathematical Theory of Communication". In: *Bell System Technical Journal* 27.3 (July 1948), pp. 379–423. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[Shi20] M. E. Shirokov. "Advanced Alicki-Fannes-Winter method for energy-constrained quantum systems and its use". In: *Quantum Information Processing* 19.5 (Apr. 2020). ISSN: 1573-1332. DOI: 10.1007/s11128-020-2581-2.

[Shi23] M E Shirokov. "Quantifying continuity of characteristics of composite quantum systems". In: *Physica Scripta* 98.4 (Mar. 2023), p. 042002. ISSN: 1402-4896. DOI: 10.1088/1402-4896/acc1b3.

[SPGC11] Norbert Schuch, David Pérez-García, and Ignacio Cirac. "Classifying quantum phases using matrix product states and projected entangled pair states". In: *Physical Review B* 84.16 (Oct. 2011), p. 165139. ISSN: 1550-235X. DOI: 10.1103/physrevb.84.165139.

[Sto30] M. H. Stone. "Linear Transformations in Hilbert Space: III. Operational Methods and Group Theory". In: *Proceedings of the National Academy of Sciences* 16.2 (Feb. 1930), pp. 172–175. ISSN: 1091-6490. DOI: 10.1073/pnas.16.2.172.

[Sto32] M. H. Stone. "On One-Parameter Unitary Groups in Hilbert Space". In: *The Annals of Mathematics* 33.3 (July 1932), p. 643. ISSN: 0003-486X. DOI: 10.2307/1968538.

[Sut18] David Sutter. "Approximate Quantum Markov Chains". In: *Approximate Quantum Markov Chains*. Springer International Publishing, 2018, pp. 75–100. ISBN: 9783319787329. DOI: 10.1007/978-3-319-78732-9_5.

[SW23] Joseph Schindler and Andreas Winter. "Continuity bounds on observational entropy and measured relative entropies". In: *Journal of Mathematical Physics* 64.9 (Sept. 2023). ISSN: 1089-7658. DOI: 10.1063/5.0147294.

[Tem16] Kristan Temme. "Thermalization Time Bounds for Pauli Stabilizer Hamiltonians". In: *Communications in Mathematical Physics* 350.2 (Sept. 2016), pp. 603–637. ISSN: 1432-0916. DOI: 10.1007/s00220-016-2746-0.

[TK15] Kristan Temme and Michael J. Kastoryano. *How fast do stabilizer Hamiltonians thermalize?* 2015. DOI: 10.48550/ARXIV.1505.07811.

[Tom16] Marco Tomamichel. *Quantum Information Processing with Finite Resources*. Springer International Publishing, 2016. ISBN: 9783319218915. DOI: 10.1007/978-3-319-21891-5.

[TPK14] Kristan Temme, Fernando Pastawski, and Michael J Kastoryano. "Hypercontractivity of quasi-free quantum semigroups". In: *Journal of Physics A: Mathematical and Theoretical* 47.40 (Sept. 2014), p. 405303. ISSN: 1751-8121. DOI: 10.1088/1751-8113/47/40/405303.

[Uhl76] A. Uhlmann. "The "transition probability" in the state space of a *-algebra". In: *Reports on Mathematical Physics* 9.2 (Apr. 1976), pp. 273–279. ISSN: 0034-4877. DOI: 10.1016/0034-4877(76)90060-4.

[Ume62] Hisaharu Umegaki. "Conditional expectation in an operator algebra. IV. Entropy and information". In: *Kodai Mathematical Journal* 14.2 (Jan. 1962). ISSN: 0386-5991. DOI: 10.2996/kmj/1138844604.

[VGRC04] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. "Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems". In: *Physical Review Letters* 93.20 (Nov. 2004), p. 207204. ISSN: 1079-7114. DOI: 10.1103/physrevlett.93.207204.

[Wan+07] X Wang et al. "Quantum information with Gaussian states". In: *Physics Reports* 448.1-4 (Aug. 2007), pp. 1–111. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2007.04.005.

[Wil20] Mark M. Wilde. "Optimal uniform continuity bound for conditional entropy of classical-quantum states". In: *Quantum Information Processing* 19.2 (Jan. 2020). ISSN: 1573-1332. DOI: 10.1007/s11128-019-2563-4.

[Win16] Andreas Winter. "Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints". In: *Communications in Mathematical Physics* 347.1 (Mar. 2016), pp. 291–313. ISSN: 1432-0916. DOI: 10.1007/s00220-016-2609-8.

[Wol12] Michael M. Wolf. *Quantum Channels and Operations - Guided Tour.* 2012. URL: https://mediatum.ub.tum.de/doc/1701036/1701036.pdf.

[YK24] Hayata Yamasaki and Kohdai Kuroiwa. *Generalized Quantum Stein's Lemma: Redeeming Second Law of Resource Theories.* 2024. DOI: 10.48550/ARXIV.2401.01926.

[Yos48] Kosaku Yosida. "On the differentiability and the representation of one-parameter semi-group of linear operators." In: *Journal of the Mathematical Society of Japan* 1.1 (Sept. 1948). ISSN: 0025-5645. DOI: 10.2969/jmsj/00110015.

Appendix A

Published Articles

Continuity bounds for quantum entropies arising from a fundamental entropic inequality

Koenraad Audenaert, Bjarne Bergh, Nilanjana Datta, Michael G Jabbour, Ángela Capel and Paul Gondolf

***Abstract*—We establish a tight upper bound for the difference in von Neumann entropies between two quantum states, ρ_1 and ρ_2. This bound is expressed in terms of the von Neumann entropies of the mutually orthogonal states derived from the Jordan-Hahn decomposition of the difference operator $(\rho_1 - \rho_2)$. This yields a novel entropic inequality that implies the well-known Audenaert-Fannes (AF) inequality. In fact, it also leads to a refinement of the AF inequality. We employ this inequality to obtain a uniform continuity bound for the quantum conditional entropy of two states whose marginals on the conditioning system coincide. We additionally use it to derive a continuity bound for the quantum relative entropy in both variables. Interestingly, the fundamental entropic inequality is also valid in infinite dimensions.**

I. Introduction

Entropies play a crucial role in both classical and quantum information theory since they characterize optimal rates of various information-processing tasks. For example, for a discrete memoryless classical information source, its optimal rate of asymptotically reliable compression (i.e., its data compression limit) is given by its Shannon entropy [1]. For the case of a quantum information source, in an analogous setting, the data compression limit is given by its von Neumann entropy [2]. For a bipartite pure state, the von Neumann entropy of one of its marginals can also be used to quantify entanglement.

There are other entropic quantities corresponding to bipartite systems, e.g. the conditional entropy and the mutual information. The quantum (Umegaki) relative entropy and the Kullback-Leibler divergences act as parent quantities for all these entropies in the quantum and classical setting, respectively. Studying mathematical properties of all these quantities (which are also often referred to as *information measures*) has been the focus of much research. An important property of these quantities, which is of relevance in the study of various information-processing tasks, is that of *continuity*. For any such entropic quantity denoted by f, this property pertains to the following question: *Given two quantum states, ρ_1 and ρ_2, that are close to each other with respect to a chosen distance measure, say t (e.g. the trace distance), how close is $f(\rho_1)$ to $f(\rho_2)$?* In other words, it amounts to finding estimates of

$$\sup\{|f(\rho_1) - f(\rho_2)| \, : \, t(\rho_1, \rho_2) \leq \varepsilon\} \, .$$

A well-known continuity bound for the von Neumann entropy, $S(\rho) := -\operatorname{Tr}(\rho \log \rho)$, of a quantum state ρ, with respect to the trace distance, is referred to as the Audenaert-Fannes (AF) inequality (1) [3], [4], [5]: For two quantum states ρ_1, ρ_2 (i.e. positive semi-definite operators of unit trace) on a finite-dimensional Hilbert space, $\mathcal{H}$, with dimension d that are ε close in trace distance, i.e. $\frac{1}{2}||\rho_1 - \rho_2||_1 = \varepsilon$, for some $\varepsilon \in [0, 1]$, it holds that

$$|S(\rho_1) - S(\rho_2)| \leq \varepsilon \log(d-1) + h(\varepsilon) \, . \tag{1}$$

Similarly, this question was also studied for the conditional entropy, which is given by $S(A|B)_\rho = S(\rho_{AB}) - S(\rho_B)$ for a bipartite state ρ_{AB}, with ρ_B being the marginal on the system B. Alicki and Fannes derived the first continuity bounds for this quantity in [6], with a later improvement by Winter in [7], strengthening it to

$$|S(A|B)_{\rho_1} - S(A|B)_{\rho_2}| \leq 2 \log d_A + (1+\varepsilon) h\Big(\frac{\varepsilon}{1+\varepsilon}\Big) \, , \tag{2}$$

where h denotes the binary entropy. The importance of these results resides not only in their numerous applications but in the generality of the method employed to derive them, which is universal for multiple entropic quantities. This method was coined *AFW method* by Shirokov in [8], [9] after the original authors, and subsequently named *ALAFF method* (for "Almost Locally AFFine") in [10], [11] due to the main property exploited in it. In the past few years, it has been multiply used not only to derive better continuity bounds for quantities derived from the Umegaki relative entropy in finite [12], [13], [14], [15] and infinite dimensions [16], [17], [18], but also for other quantities such as Rényi divergences [19], [20], the Belavkin-Staszewski relative entropy [10], the fidelity [21], and more.

In this paper, we introduce a new upper bound on $S(\rho_1) - S(\rho_2)$ which will turn out to imply the AF inequality (1), and also lead to a uniform continuity bound for the quantum conditional entropy when the marginals on the conditioning system agree, and to a continuity bound on the quantum (Umegaki)

K. Audenaert is in the Institut für Theoretische Physik, Albert-Einstein-Allee 11, Universität Ulm 89069, Germany (e-mail: koenraad.audenaert@uni-ulm.de).

B. Bergh is in the Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, United Kingdom (e-mail: bb536@cam.ac.uk).

N. Datta is in the Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, United Kingdom (e-mail: n.datta@damtp.cam.ac.uk).

M. G. Jabbour is at SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France, and in the Centre for Quantum Information and Communication, École polytechnique de Bruxelles, CP 165/59, Université libre de Bruxelles, 1050 Brussels, Belgium (e-mail: mjabbour@telecom-sudparis.eu).

Á. Capel is in the Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, United Kingdom, and at Fachbereich Mathematik, Universität Tübingen, 72076 Tübingen, Germany (e-mail: ac2722@cam.ac.uk).

P. Gondolf is at Fachbereich Mathematik, Universität Tübingen, 72076 Tübingen, Germany (e-mail: paul.gondolf@uni-tuebingen.de).

relative entropy. Let us introduce the inequality in various equivalent forms. Consider the Jordan-Hahn decomposition of the difference $\rho_1 - \rho_2$:

$$\rho_1 - \rho_2 = \Delta_+ - \Delta_- ,$$

where $\Delta_\pm$ are positive semi-definite operators of orthogonal supports (which we write as $\Delta_+ \perp \Delta_-$). We can express this difference in terms of quantum states $\rho_\pm$, namely, $\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_-$, with the states $\rho_\pm$ being defined through the relation: $\Delta_\pm = \varepsilon\rho_\pm$ if $\varepsilon > 0$, and for $\varepsilon = 0$ we can make an arbitrary choice of states ρ_+ and ρ_- such that $\rho_+ \perp \rho_-$. In this paper, we prove that the difference of the von Neumann entropies of the states ρ_1 and ρ_2 can be expressed in terms of the entropies of the states $\rho_\pm$ via the following inequality:

$$S(\rho_1) - S(\rho_2) \le \varepsilon S(\rho_+) - \varepsilon S(\rho_-) + h(\varepsilon) , \tag{3}$$

where $h(\varepsilon) := -\varepsilon \log \varepsilon - (1-\varepsilon)\log(1-\varepsilon)$ denotes the binary entropy. Moreover, this bound is *tight*, in the sense that, for any $\varepsilon \in [0, 1]$, there exist pairs of states for which the bound is saturated. This inequality can be cast in various equivalent forms. Firstly, note that interchanging ρ_1 and ρ_2 results in interchanging Δ_+ and Δ_-. Hence, we also have

$$S(\rho_2) - S(\rho_1) \le \varepsilon S(\rho_-) - \varepsilon S(\rho_+) + h(\varepsilon) . \tag{4}$$

As an immediate consequence of (3) and (4) one gets

$$|\,(S(\rho_1) - S(\rho_2)) - \varepsilon\,(S(\rho_+) - \varepsilon S(\rho_-))\,| \le h(\varepsilon) . \tag{5}$$

We see that (3) is more fundamental than the AF inequality (1). This is because from (3) it follows that (assuming without loss of generality that $S(\rho_1) \le S(\rho_2)$),

$$\begin{aligned} |S(\rho_1) - S(\rho_2)| &\le \varepsilon S(\rho_+) - \varepsilon S(\rho_-) + h(\varepsilon) \\ &\le \varepsilon S(\rho_+) + h(\varepsilon) \\ &\le \varepsilon \log(d-1) + h(\varepsilon) , \end{aligned} \tag{6}$$

which is the right-hand side of (1). In the second line, we have used the non-negativity of the von Neumann entropy, and the last line follows from the fact that $\mathrm{rank}\,\rho_+ \le d - 1$.

To gain some intuition behind our new entropic inequality (considered in any of its equivalent forms), let us first consider some simple cases where it can be easily seen to hold.

Case 1: The inequality clearly holds when ρ_1 and ρ_2 are states of qubits, i.e. when $\mathcal{H} \simeq \mathbb{C}^2$. This is because in this case, $\rho_\pm$ are pure states and hence $S(\rho_\pm) = 0$. Therefore, (5) reduces to

$$|S(\rho_1) - S(\rho_2)| \le h(\varepsilon) ,$$

which is just the AF inequality (1) for the case of qubits (i.e. $d = 2$).

Case 2: The inequality holds whenever $\rho_1 \ge \varepsilon\rho_+$, which in turn guarantees that $\rho_2 \ge \varepsilon\rho_-$, since via the Jordan-Hahn decomposition we have that $\rho_1 - \varepsilon\rho_+ = \rho_2 - \varepsilon\rho_-$. The inequality can then be proved as follows. Note that $\mathrm{Tr}(\rho_1 - \varepsilon\rho_+) = 1 - \varepsilon = \mathrm{Tr}(\rho_2 - \varepsilon\rho_-)$. Note first that if $\varepsilon = 1$, then $\rho_1 \perp \rho_2$ and so $\rho_+ = \rho_1$, $\rho_- = \rho_2$ and so the inequality holds trivially. If $\varepsilon < 1$, then let us define the quantum state

$$\omega := \frac{\rho_1 - \varepsilon\rho_+}{1-\varepsilon} \equiv \frac{\rho_2 - \varepsilon\rho_-}{1-\varepsilon} . \tag{7}$$

Then, we can write convex decompositions of the states ρ_1 and ρ_2 as follows:

$$\rho_1 = \varepsilon\rho_+ + (1-\varepsilon)\omega , \tag{8}$$

$$\rho_2 = \varepsilon\rho_- + (1-\varepsilon)\omega . \tag{9}$$

The property of "almost convexity" of the von Neumann entropy [22], [23] applied to (8) implies that

$$S(\rho_1) \le \varepsilon S(\rho_+) + (1-\varepsilon)S(\omega) + h(\varepsilon) . \tag{10}$$

Moreover, the concavity of the von Neumann entropy applied to (9) implies that

$$S(\rho_2) \ge \varepsilon S(\rho_-) + (1-\varepsilon)S(\omega) . \tag{11}$$

Then from (10) and (11) we immediately obtain the desired inequality (3).

$$S(\rho_1) - S(\rho_2) \le \varepsilon S(\rho_+) - \varepsilon S(\rho_-) + h(\varepsilon) . \tag{12}$$

Remark 1. It can also be easily seen that the inequality holds when ρ_1 and ρ_2 commute, as this is a special case of *Case 2*. Note first that if ρ_1 and ρ_2 commute then ρ_1 and ρ_2 also commute with $\rho_\pm$. Also, the states ρ_1 and ρ_2 are then simultaneously diagonalizable and hence the operator inequality $\rho_1 \ge \varepsilon\rho_+$ reduces to an inequality between eigenvalues of ρ_1 and ρ_+. Let us fix some ordering of the vectors in this simultaneous eigenbasis, and then write the eigenvalue corresponding to the i^{th} basis vector as $\lambda_i(\sigma)$ for any state σ that is diagonal in this basis. Then the operator inequality $\rho_1 \ge \varepsilon\rho_+$ reduces to

$$\lambda_i(\rho_1) \ge \varepsilon\lambda_i(\rho_+) \quad \forall\ i \in [d], \tag{13}$$

where $[d]$ denotes the index set of d elements. Let $p_i := \lambda_i(\rho_1)$ and $q_i := \lambda_i(\rho_2)$, for all $i \in [d]$, and define the sets

$$\begin{aligned} I &:= \{i \in [d] \,:\, p_i \ge q_i\} , \\ I^c &:= \{i \in [d] \,:\, p_i < q_i\} . \end{aligned} \tag{14}$$

Then $\lambda_i(\rho_1 - \varepsilon\rho_+) = \lambda_i(\rho_1) - \varepsilon\lambda_i(\rho_+)$, where $\lambda_i(\rho_+) = p_i - q_i$ for all $i \in I$ and is equal to zero else. Hence, we have that for all $i \in I$, $\lambda_i(\rho_1 - \varepsilon\rho_+) = (1-\varepsilon)p_i + \varepsilon q_i \ge 0$ and for all $i \in I^c$, $\lambda_i(\rho_1 - \varepsilon\rho_+) = p_i \ge 0$. Thus the required inequality of Case 2, namely, $\rho_1 \ge \varepsilon\rho_+$, holds in this case.

Layout of the paper: Our main result, namely the above-mentioned entropic inequality, is stated in Theorem 1 of Section II, and a sharper version of it is stated in Theorem 3. An extension of this inequality to conditional entropies is given in Theorem 2. In Section III we state and prove a few key lemmas that we employ in the proof of the above theorems and of subsequent results. The proof of Theorem 3 is given in Section IV. In Section V we use our fundamental entropic inequality to state and prove a refined version of the AF inequality (1); see Theorem 4. In Section VI, we apply Theorem 1 to obtain a uniform continuity bound for the conditional entropy whenever the marginals on the second system agree, and a continuity bound on the quantum relative entropy. These are stated in Theorem 5 and Corollary 1, respectively, and their proofs are presented in the same section. We end the paper with an extension of our fundamental entropic inequality from Theorem 1 to infinite dimensions in Section VII.

II. Main Results

For the majority of this paper, with the exception of Section VII at the end where we deal with infinite-dimensional Hilbert spaces, we restrict attention to a finite-dimensional Hilbert space $\mathcal{H}$ of dimension d. Let $\mathcal{B}(\mathcal{H})$ denote the algebra of linear operators acting on $\mathcal{H}$, and $\mathcal{B}_{\rm sa}(\mathcal{H})$ denote the subset of self-adjoint ones. The set of quantum states (density matrices), i.e. positive semi-definite operators of unit trace is denoted by $\mathcal{D}(\mathcal{H}) \subset \mathcal{B}_{\rm sa}(\mathcal{H})$, and its subset of positive definite operators of unit trace by $\mathcal{D}_+(\mathcal{H})$. For $X \in \mathcal{B}(\mathcal{H})$, we denote the kernel of X as $\ker X = \{|\psi\rangle \in \mathcal{H} \; : \; X|\psi\rangle = 0\}$ and its support by $\operatorname{supp} X = (\ker X)^\perp$. Note that when writing $A \le B$ for $A, B \in \mathcal{B}_{\rm sa}(\mathcal{H})$ we refer to the Loewner partial order. The norms on $\mathcal{B}(\mathcal{H})$ that we use are the trace- or one-norm $\|\cdot\|_1$ and the operator- or infinity-norm $\|\cdot\|_\infty$ both of which are members of the wider family of Schatten-p-norms $\|A\|_p = \operatorname{Tr}\big((A^*A)^{p/2}\big)^{1/p}$ for $p \in [1,\infty)$ (where $\|\cdot\|_\infty$ corresponds to the limit $p \to \infty$). The trace distance between two density matrices $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ is given by $\frac{1}{2}\|\rho_1 - \rho_2\|_1$.

For any vector $\underline{p} \in \mathbb{R}^d$ of non-negative entries (not necessarily a normalized vector), we define its Shannon entropy $H(\underline{p})$ as $H(\underline{p}) := \sum_i \eta(p_i) := -\sum_i p_i \log p_i$. Additionally, for any positive semi-definite operator $A \in \mathcal{B}_{\rm sa}(\mathcal{H})$, $A \ge 0$, we define its von Neumann entropy as $S(A) := -\operatorname{Tr}(A \log A)$. The quantum (Umegaki) relative entropy of a state ρ with respect to a positive semi-definite operator A is given by

$$D(\rho\|A) = \begin{cases} \operatorname{Tr}(\rho\log\rho - \rho\log A) & \text{if } \ker A \subseteq \ker\rho, \\ \infty & \text{else.} \end{cases} \tag{15}$$

For a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the conditional entropy of the system A given the system B is given by $S(A|B)_\rho := S(\rho_{AB}) - S(\rho_B)$, where $\rho_B = \operatorname{Tr}_A \rho_{AB}$ denotes the reduced state (i.e. marginal) of the system B. It can be expressed in terms of a relative entropy as follows:

$$\begin{aligned} S(A|B)_\rho &= -D(\rho_{AB}\| \mathbb{1}_A \otimes \rho_B) \\ &= \max_{\nu_B \in \mathcal{D}(\mathcal{H}_B)} \left[-D(\rho_{AB}\| \mathbb{1}_A \otimes \nu_B)\right], \end{aligned} \tag{16}$$

where $\mathbb{1}_A$ denotes the identity operator on the system A. We also employ the max-relative entropy [24] which is defined as follows[1]:

$$D_{\max}(\rho\|\sigma) := \inf\{\lambda > 0 \, : \, \rho \le e^\lambda \sigma\}. \tag{17}$$

Note that throughout this paper, we use log to denote the natural logarithm.

We are now in position to state our main results.

Theorem 1. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$, with $\dim \mathcal{H} = d$, such that $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$, for some $\varepsilon \in [0,1]$. Let $\rho_\pm$ be the normalized Jordan-Hahn decomposition of $(\rho_1 - \rho_2)$ as described above, i.e.*

$$\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_- \,, \tag{18}$$

where $\rho_\pm \in \mathcal{D}(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$ and $\rho_+ \perp \rho_-$. Then

$$S(\rho_1) - S(\rho_2) \le \varepsilon S(\rho_+) - \varepsilon S(\rho_-) + h(\varepsilon)\,. \tag{19}$$

Moreover, the inequality is tight.

Remark 2. To see that the bound (19) is tight, for every $\varepsilon \in [0,1]$ one can simply consider the following commuting states ρ_1 and ρ_2:

$$\rho_1 = (1-\varepsilon)\,|\psi\rangle\langle\psi| + \frac{\varepsilon}{d-1}\,(\mathbb{1} - |\psi\rangle\langle\psi|) \quad \text{and} \quad \rho_2 = |\psi\rangle\langle\psi| \,, \tag{20}$$

where $|\psi\rangle$ is any pure state, while $\mathbb{1}$ denotes the identity operator in $\mathcal{B}(\mathcal{H})$.

We prove this theorem in the next section. The above theorem extends to conditional entropies for bipartite states if the condition given in (22) below holds. This result is stated in Theorem 2.

Theorem 2. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$ and $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$, for some $\varepsilon \in [0,1]$. Let $\rho_\pm$ be the normalized Jordan-Hahn decomposition of $(\rho_1 - \rho_2)$ as described above, i.e.*

$$\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_- \,, \tag{21}$$

where $\rho_\pm \in \mathcal{D}(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$ and $\rho_+ \perp \rho_-$. Further, assume that the operator K defined below is positive semi-definite, i.e.

$$K := \rho_1 - \varepsilon\rho_+ = \rho_2 - \varepsilon\rho_- \ge 0. \tag{22}$$

Then

$$S(A|B)_{\rho_1} - S(A|B)_{\rho_2} \le \varepsilon S(A|B)_{\rho_+} - \varepsilon S(A|B)_{\rho_-} + h(\varepsilon). \tag{23}$$

Proof. Similarly to above, if $\varepsilon = 1$, then $\rho_1 = \rho_+$ and $\rho_2 = \rho_-$ and so the relation holds trivially. For $\varepsilon < 1$, let $\omega \in \mathcal{D}(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$ be defined through the relation $K = (1-\varepsilon)\omega$. Then

$$\rho_1 = \varepsilon\rho_+ + (1-\varepsilon)\omega \tag{24}$$

$$\rho_2 = \varepsilon\rho_- + (1-\varepsilon)\omega. \tag{25}$$

Note that the conditional entropy $S(A|B)_{\rho_1}$ is given by

$$\begin{aligned} S(A|B)_{\rho_1} &= S(\rho_1) - S(\rho_{1,B}) \\ &= S(\rho_1) + \operatorname{Tr}\left(\rho_1 \log(\mathbb{1}_A \otimes \rho_{1,B})\right), \end{aligned} \tag{26}$$

where $\rho_{1,B} = \operatorname{Tr}_A \rho_1$. Further, by (24),

$$\begin{aligned} \operatorname{Tr}\left(\rho_1 \log(\mathbb{1}_A \otimes \rho_{1,B})\right) &= \varepsilon \operatorname{Tr}\left(\rho_+ \log(\mathbb{1}_A \otimes \rho_{1,B})\right) \\ &\quad + (1-\varepsilon)\operatorname{Tr}\left(\omega \log(\mathbb{1}_A \otimes \rho_{1,B})\right) \end{aligned} \tag{27}$$

and

$$S(\rho_1) \le \varepsilon S(\rho_+) + (1-\varepsilon) S(\omega) + h(\varepsilon), \tag{28}$$

[1] The definition in the original paper [24] is $D_{\max}(\rho\|\sigma) := \inf\{\lambda > 0 : \rho \le 2^\lambda \sigma\}$. However, we consider here a slightly modified version since we are using natural logarithms throughout the whole text.

where the last inequality follows from the property of "almost convexity" of the von Neumann entropy. The above inequalities imply that

$$\begin{aligned} S(A|B)_{\rho_1} &= S(\rho_1) + \mathrm{Tr}\left(\rho_1 \log(\mathbb{1}_A \otimes \rho_{1,B})\right) \\ &\le h(\varepsilon) - \varepsilon D(\rho_+ \| \mathbb{1}_A \otimes \rho_{1,B}) \\ &\quad - (1-\varepsilon) D(\omega \| \mathbb{1}_A \otimes \rho_{1,B}) \\ &\le h(\varepsilon) + \varepsilon S(A|B)_{\rho_+} + (1-\varepsilon) S(A|B)_\omega, \end{aligned} \tag{29}$$

where in the last step we also used the variational characterization of the conditional entropy, i.e. (16). On the other hand, by (25) and the concavity of the conditional entropy, we have

$$S(A|B)_{\rho_2} \ge \varepsilon S(A|B)_{\rho_-} + (1-\varepsilon) S(A|B)_\omega. \tag{30}$$

Inequalities (29) and (30) yield the desired inequality (23). □

III. Two key lemmas

The proof of Theorem 1 (and its sharper version, Theorem 3 stated in the next section) will be simplified if we extend the definition of the von Neumann entropy to positive operators that do not necessarily have trace 1. Remember that we defined the functional

$$S(A) := -\mathrm{Tr}(A \log A) \tag{31}$$

for every positive operator A, $A \ge 0$. Clearly, when ρ is a state, $a \ge 0$, and $A = a\rho$, we have

$$S(A) = aS(\rho) - a \log a \tag{32}$$

This identity allows to generalise the inequalities (10) and (11) expressing almost convexity and concavity of the von Neumann entropy, respectively, to this entropy functional. One easily obtains the following, for $A, B \ge 0$ with $a = \mathrm{Tr}\, A$, $b = \mathrm{Tr}\, B$:

$$\begin{aligned} S(A+B) &\le S(A) + S(B) && (33) \\ S(A+B) &\ge S(A) + S(B) - (a+b) h\left(\frac{b}{a+b}\right). && (34) \end{aligned}$$

Thus, almost convexity turns into functional subadditivity (not to be confused with the usual subadditivity of the von Neumann entropy with respect to addition of subsystems), and concavity turns into functional almost-super-additivity.

We now show that the latter inequality can be made sharper when B is not full rank. That this should be possible is already being hinted at by the existence of the identity $S(A+B) = S(A) + S(B)$ when A and B have orthogonal supports. The following lemma extends this fact.

Lemma 1 (Sharpened almost-superadditivity)**.** *Let $A, B \ge 0$ and $\mathcal{M} = \mathrm{supp}\, B$. Denoting the restriction of an operator X to $\mathcal{M}$ by $X|_\mathcal{M}$, and defining $a' = \mathrm{Tr}\, A|_\mathcal{M}$ and $b = \mathrm{Tr}\, B \equiv \mathrm{Tr}\, B|_\mathcal{M}$, we have*

$$\begin{aligned} S(A+B) - S(A) &\ge S((A+B)|_\mathcal{M}) - S(A|_\mathcal{M}) \\ &\ge S(B) - (a'+b) h\left(\frac{b}{a'+b}\right) \end{aligned} \tag{35}$$

This inequality will be an essential ingredient in the proof of Theorem 1.

Proof. Monotonicity of the Holevo chi χ under a CPTP map Φ, applied to a two-element ensemble, explicitly reads as follows:

$$\begin{aligned} &S(p\rho + (1-p)\sigma) - pS(\rho) - (1-p)S(\sigma) \\ &\ge S(p\Phi(\rho) + (1-p)\Phi(\sigma)) - pS(\Phi(\rho)) - (1-p)S(\Phi(\sigma)), \end{aligned} \tag{36}$$

or, rephrased in terms of positive operators A and B,

$$\begin{aligned} S(A+B) - S(A) - S(B) &\ge S(\Phi(A) + \Phi(B)) \\ &\quad - S(\Phi(A)) - S(\Phi(B)). \end{aligned} \tag{37}$$

The third term of each side drops out if Φ leaves B unchanged. Let, in particular, Φ be a pinching to the subspaces $\mathcal{M} = \mathrm{supp}\, B$ and $\mathcal{M}^\perp = \ker B$. Then

$$\begin{aligned} S(A+B) - S(A) &\ge S((A+B)|_\mathcal{M} \oplus (A+B)|_{\mathcal{M}^\perp}) \\ &\quad - S(A|_\mathcal{M} \oplus A|_{\mathcal{M}^\perp}) \\ &= S((A+B)|_\mathcal{M}) - S(A|_\mathcal{M}), \end{aligned}$$

which is the first inequality of the lemma.

The second inequality of the lemma then follows by exploiting almost super-additivity of S given by (34). □

Lemma 2. *For $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ with $\rho \perp \sigma$ (i.e. they have mutually orthogonal supports) and $\omega = t\rho + (1-t)\nu$ with $t \in (0,1)$ and $\nu \in \mathcal{D}(\mathcal{H})$, one has*

$$D(\rho\|\omega) - D(\sigma\|\omega) \le \log\left(\frac{1}{t} - 1\right). \tag{38}$$

Proof. From $\omega = t\rho + (1-t)\nu$ it follows that $\omega \ge t\rho$ which in turn gives

$$D(\rho\|\omega) \le D(\rho\|t\rho) = -\log t\,. \tag{39}$$

Let us define the pinching map Φ which acts on any $\tau \in \mathcal{D}(\mathcal{H})$ as follows:

$$\Phi(\tau) := P_\sigma \tau P_\sigma + P_\sigma^\perp \tau P_\sigma^\perp, \tag{40}$$

where P_σ and $P_\sigma^\perp$ denote orthogonal projections onto the support of σ and its complement, respectively. Then, by the data-processing inequality, we have

$$\begin{aligned} D(\sigma\|\omega) &\ge D(\Phi(\sigma)\|\Phi(\omega)) \\ &= D(\sigma\|t\rho + (1-t)\Phi(\nu)) \\ &= D(\sigma\|(1-t)\Phi(\nu)), \end{aligned} \tag{41}$$

where the last equality holds because $\rho \perp \sigma$. Therefore,

$$D(\sigma\|\omega) \ge -\log(1-t) + D(\sigma\|\Phi(\nu)) \ge -\log(1-t), \tag{42}$$

due to the non-negativity of the relative entropy between two quantum states. The bounds (39) and (42) together yield the statement of the lemma, since

$$D(\rho\|\omega) - D(\sigma\|\omega) \le -\log t + \log(1-t) = \log\left(\frac{1}{t} - 1\right). \tag{43}$$

□

We are now in a position to prove Theorem 1.

IV. Proof of Theorem 1

As mentioned earlier, we actually prove an inequality which is sharper than the one stated in Theorem 1. It not only involves the quantity ε (i.e. the trace distance between the states ρ_1 and ρ_2) but also

$$c := \operatorname{Tr} \rho_2|_{\mathcal{M}}$$

where $\mathcal{M}$ denotes the support of ρ_-. We are grateful to Peter Frenkel for inquiring about a possibility of such kind. The sharper inequality is stated in the following theorem.

Theorem 3. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$, with $\dim \mathcal{H} = d$, such that $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$, for some $\varepsilon \in [0,1]$. Let $\rho_\pm$ be the normalized Jordan-Hahn decomposition of $(\rho_1 - \rho_2)$ as described above i.e.*

$$\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_- , \tag{44}$$

where $\rho_\pm \in \mathcal{D}(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$ and $\rho_+ \perp \rho_-$. Further, let $c := \operatorname{Tr} \rho_2|_\mathcal{M}$, where $\mathcal{M}$ denotes the support of ρ_-. Then

$$S(\rho_1) - S(\rho_2) \leq \varepsilon S(\rho_+) - \varepsilon S(\rho_-) + c\, h\left(\frac{\varepsilon}{c}\right). \tag{45}$$

To see that this inequality is indeed sharper than the one stated in Theorem 1, we need to establish that

$$c\, h\left(\frac{\varepsilon}{c}\right) \leq h(\varepsilon), \tag{46}$$

This follows from concavity of the binary entropy h and the fact that $0 \leq \varepsilon \leq c \leq 1$ (which in turn follows from taking the trace of the restriction to $\mathcal{M} = \operatorname{supp} \rho_-$ of the identity (44)):

$$\begin{aligned} h(\varepsilon) &= h\left(c\frac{\varepsilon}{c} + (1-c)0\right) \\ &\geq c\, h\left(\frac{\varepsilon}{c}\right) + (1-c)h(0) = c\, h\left(\frac{\varepsilon}{c}\right). \end{aligned} \tag{47}$$

We now proceed to prove Theorem 3.

Proof. Let us define the positive operator M by

$$M := \rho_1 + \Delta_- = \rho_2 + \Delta_+.$$

Then we have

$$\begin{aligned} S(\rho_1) - S(\rho_2) &= S(\rho_1) - S(M) + S(M) - S(\rho_2) \quad (48) \\ &= -(S(\rho_1 + \Delta_-) - S(\rho_1)) \\ &\quad + (S(\rho_2 + \Delta_+) - S(\rho_2)). \quad (49) \end{aligned}$$

To find an upper bound on the first term we use the inequality (35) of Lemma 1 (sharpened almost super-additivity), and note that $\operatorname{Tr} \Delta_- = \varepsilon$ and $\operatorname{Tr}(\rho_1 + \Delta_-)|_\mathcal{M} = \operatorname{Tr} \rho_2|_\mathcal{M} = c$. For the second term we use subadditivity (33). This gives

$$\begin{aligned} S(\rho_1) - S(\rho_2) &\leq -\left(S(\Delta_-) - c\, h\left(\frac{\varepsilon}{c}\right)\right) + S(\Delta_+) \\ &= S(\Delta_+) - S(\Delta_-) + c\, h\left(\frac{\varepsilon}{c}\right) \\ &= \varepsilon S(\rho_+) - \varepsilon S(\rho_-) + c\, h\left(\frac{\varepsilon}{c}\right). \end{aligned} \tag{50}$$

□

V. A refined continuity bound for the von Neumann entropy

The fundamental entropic inequality stated in Theorem 1 along with Lemma 2 leads to the refinement of the AF inequality (1) given by Theorem 4 below.

Theorem 4. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$, with $\dim \mathcal{H} = d$, such that $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$, for some $\varepsilon \in [0,1]$. Let $\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_-$ where $\rho_\pm \in \mathcal{D}(\mathcal{H})$, and $\rho_+ \perp \rho_-$. Then*

$$\begin{aligned} &|S(\rho_1) - S(\rho_2)| \\ &\leq \varepsilon \log\left(d \max\{\lambda_{\max}(\rho_-), \lambda_{\max}(\rho_+)\} - 1\right) + h(\varepsilon). \end{aligned} \tag{51}$$

Remark 3. Note that Berta *et al* [25] proved an analogous result but with $\lambda_{\max}(\rho_-)$ replaced by $\lambda_{\max}(\rho_2)$, and for $\varepsilon \leq 1 - (1/(d\lambda_{\max}(\rho_2))$; see Corollary 3 of [25]. In fact, it was their result that inspired us to prove Theorem 4.

Before proving the above theorem, assume without loss of generality that $S(\rho_1) \geq S(\rho_2)$. Recall that equality holds in the AF inequality (1) for the following choice:

$$\rho_1 = (1-\varepsilon)\, |1\rangle\langle 1| + \frac{\varepsilon}{d-1}\sum_{i=2}^{d} |i\rangle\langle i|\,, \quad \text{and} \quad \rho_2 = |1\rangle\langle 1|\,,$$

since $S(\rho_2) = 0$ and $S(\rho_1) = \varepsilon \log(d-1) + h(\varepsilon)$.

Note that for this choice of ρ_1 and ρ_2, we have

$$\rho_+ = \frac{1}{d-1}\sum_{i=2}^{d} |i\rangle\langle i| \quad \text{and} \quad \rho_- = |1\rangle\langle 1|\,.$$

Hence, $\lambda_{\max}(\rho_-) = 1$ so that the RHS of (51) reduces to the RHS of the AF inequality (1), and the inequality is saturated.

A. Proof of Theorem 4

Proof. By Theorem 1,

$$\begin{aligned} S(\rho_1) - S(\rho_2) &\leq \varepsilon(S(\rho_+) - S(\rho_-)) + h(\varepsilon) \\ &= \varepsilon(D(\rho_-\|\tau) - D(\rho_+\|\tau)) + h(\varepsilon), \end{aligned} \tag{52}$$

where $\tau = I/d$, the completely mixed state. Let the spectral decomposition of ρ_- be given by

$$\rho_- = \sum_{i=1}^{d} \lambda_i\, |e_i\rangle\langle e_i| \leq \lambda_{\max}(\rho_-) I = \frac{1}{t}\tau, \tag{53}$$

where $t := \frac{1}{d\lambda_{\max}(\rho_-)}$. Setting $\omega \equiv \tau$, we obtain from (53)

$$\omega \equiv \tau = t\rho_- + (1-t)\nu, \tag{54}$$

where $\nu = \frac{\tau - t\rho_-}{1-t} \in \mathcal{D}(\mathcal{H})$. Since $\rho_+ \perp \rho_-$ and (54) holds, we can apply Lemma 2 to the RHS of (52) to obtain

$$\begin{aligned} \text{RHS of (52)} &\leq \varepsilon \log\left(\frac{1}{t} - 1\right) + h(\varepsilon) \\ &= \varepsilon \log(d\lambda_{\max}(\rho_-) - 1) + h(\varepsilon). \end{aligned} \tag{55}$$

To obtain (51), note when exchanging ρ_1 and ρ_2, ρ_+ will become ρ_- and vice-versa, and so we can obtain the absolute value bound of (51) by combining (55) with its version where ρ_1 and ρ_2 have been exchanged. □

VI. Further Applications

As a straightforward application of our inequality (19), we obtain a uniform continuity bound for the conditional entropy $S(A|B)_\rho$, in the special case in which the marginals of the two states on the system B are identical. The following has been conjectured (also in the case where the two marginals are not identical) [7], [26]: For $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, such that $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$,

$$|S(A|B)_{\rho_1} - S(A|B)_{\rho_2}| \leq \varepsilon \log\bigl(d_A^2 - 1\bigr) + h(\varepsilon) . \quad (56)$$

This continuity bound is tight as the inequality is saturated for certain choices of states $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ [26], [27]. The analogous classical inequality has appeared e.g. in [28] and [29]. Our proof of (56) for the special case of equal marginals is given by Theorem 5.

Using Lemma 2, we can easily prove the following result on the uniform continuity bound for the conditional entropy of bipartite states with equal marginals, using our fundamental inequality (19) stated in Theorem 1.

Theorem 5. *For $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with equal marginals $\rho_{1,B} = \rho_{2,B}$ and $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$ for some $\varepsilon \in [0, 1]$, we find that*

$$|S(A|B)_{\rho_1} - S(A|B)_{\rho_2}| \leq \varepsilon \log\bigl(d_A^2 - 1\bigr) + h(\varepsilon) . \quad (57)$$

Proof. Note that since $\rho_{1,B} = \rho_{2,B}$, we further get that $0 = \mathrm{Tr}_A(\rho_1 - \rho_2) = \mathrm{Tr}_A(\varepsilon(\rho_+ - \rho_-))$. This immediately gives $\rho_{+,B} = \rho_{-,B} \equiv \tilde{\rho}_B$ and hence also $S(\rho_{-,B}) = S(\rho_{+,B})$. Then the proof is a simple application of Theorem 1 and Lemma 2. Without loss of generality, we assume that $S(A|B)_{\rho_1} > S(A|B)_{\rho_2}$. Then,

$$\begin{aligned} S(A|B)_{\rho_1} - S(A|B)_{\rho_2} &= S(\rho_1) - S(\rho_2) \\ &\leq \varepsilon(S(\rho_+) - S(\rho_-)) + h(\varepsilon) \\ &= \varepsilon(S(A|B)_{\rho_+} - S(A|B)_{\rho_-}) \\ &\quad + h(\varepsilon). \end{aligned} \quad (58)$$

To find an upper bound on the difference $\bigl(S(A|B)_{\rho_+} - S(A|B)_{\rho_-}\bigr)$ we first express it as a difference of two relative entropies using (16)

$$\begin{aligned} &S(A|B)_{\rho_+} - S(A|B)_{\rho_-} \\ &\quad = D(\rho_- \| \mathbb{1}_A \otimes \tilde{\rho}_B) - D(\rho_+ \| \mathbb{1}_A \otimes \tilde{\rho}_B) \\ &\quad = D(\rho_- \| \tau_A \otimes \tilde{\rho}_B) - D(\rho_+ \| \tau_A \otimes \tilde{\rho}_B) \end{aligned}$$

where $\tau_A = \mathbb{1}_A / d_A$ (the completely mixed state). In the above, the last line follows from the fact that for any positive constant c, $D(\rho \| c\sigma) = D(\rho\|\sigma) - \log c$.

We can complete the proof of the theorem by applying Lemma 2 with the following choices: $\rho = \rho_-$, $\sigma = \rho_+$, $\omega = \tau_A \otimes \tilde{\rho}_B$ and $t = \frac{1}{d_A^2}$. We can do this because the conditions of the lemma are satisfied for these choices: first, $\rho_+ \perp \rho_-$, and second, $\rho_- \leq d_A^2 \tau_A \otimes \tilde{\rho}_B$, (which follows from e.g. [30, Lemma 5.11]) since $\tilde{\rho}_B = \mathrm{Tr}_A(\rho_-)$, and hence we can write

$$\tau_A \otimes \tilde{\rho}_B = \frac{1}{d_A^2}\rho_- + \Bigl(1 - \frac{1}{d_A^2}\Bigr)\nu ,$$

where $\nu = \frac{d_A^2}{d_A^2 - 1}(\tau_A \otimes \tilde{\rho}_B - \frac{1}{d_A^2}\rho_-)$.

Thus, applying Lemma 2 with the above choices, we obtain

$$S(A|B)_{\rho_1} - S(A|B)_{\rho_2} \leq \varepsilon \log\bigl(d_A^2 - 1\bigr) + h(\varepsilon) . \quad (59)$$

We finalize the proof by swapping the roles of ρ_1 and ρ_2, which gives us the absolute value on the left-hand side of he statement. □

The search for the tight bound (56) conjectured in [7], [26], in the setting beyond the one in which the marginals on the conditioning system are identical, remains open.

As mentioned earlier, our fundamental inequality, (19) of Theorem 1, and Lemma 2 can also be used to obtain a continuity bound on the quantum (Umegaki) relative entropy [31] which improves upon the best-known bounds [10], [11], [32]. We begin with the particular case in which the second states in both relative entropies are identical.

Theorem 6. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$, $\sigma \in \mathcal{D}_+(\mathcal{H})$ with $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$ where $\varepsilon \in [0, 1]$. Then*

$$\begin{aligned} &|D(\rho_1\|\sigma) - D(\rho_2\|\sigma)| \\ &\leq \varepsilon \log\Bigl(e^{\max\{D_{\max}(\rho_+\|\sigma), D_{\max}(\rho_-\|\sigma)\}} - 1\Bigr) + h(\varepsilon) , \end{aligned} \quad (60)$$

where $\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_-$, and $\rho_\pm$ is defined via the Jordan-Hahn decomposition of $(\rho_1 - \rho_2)$ as in (18).

In particular, for $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{D}_+(\mathcal{H})$ with $\frac{1}{2}\|\rho - \sigma\|_1 = \varepsilon$ where $\varepsilon \in [0, 1]$, we obtain

$$D(\rho\|\sigma) \leq \varepsilon \log\Bigl(e^{D_{max}(\omega_+\|\sigma)} - 1\Bigr) + h(\varepsilon) , \quad (61)$$

where ω_+ is defined via the Jordan-Hahn decomposition of $(\rho - \sigma)$ similarly as ρ_+ for $(\rho_1 - \rho_2)$.

Proof. First, note that

$$\begin{aligned} &D(\rho_1\|\sigma) - D(\rho_2\|\sigma) \\ &= - S(\rho_1) + S(\rho_2) - \mathrm{Tr}((\rho_1 - \rho_2)\log\sigma) \\ &= - S(\rho_1) + S(\rho_2) - \varepsilon\,\mathrm{Tr}((\rho_+ - \rho_-)\log\sigma) \\ &\leq \varepsilon(S(\rho_-) - S(\rho_+)) - \varepsilon\,\mathrm{Tr}((\rho_+ - \rho_-)\log\sigma) + h(\varepsilon), \end{aligned} \quad (62)$$

where in the last line we have used (4). Thus, rewriting the RHS of (62) as relative entropies, we obtain

$$D(\rho_1\|\sigma) - D(\rho_2\|\sigma) \leq \varepsilon(D(\rho_+\|\sigma) - D(\rho_-\|\sigma)) + h(\varepsilon) . \quad (63)$$

Now, note that, from the definition (17) of $D_{\max}$, we have

$$\rho_+ \leq e^{D_{\max}(\rho_+\|\sigma)}\sigma , \quad (64)$$

and thus

$$\sigma = e^{-D_{\max}(\rho_+\|\sigma)}\rho_+ + \Bigl(1 - e^{-D_{\max}(\rho_+\|\sigma)}\Bigr)\nu \quad (65)$$

for a certain $\nu \in \mathcal{D}(\mathcal{H})$. Then, as an immediate consequence of Lemma 2, we have

$$D(\rho_+\|\sigma) - D(\rho_-\|\sigma) \leq \log\Bigl(e^{D_{\max}(\rho_+\|\sigma)} - 1\Bigr) , \quad (66)$$

which jointly with (63), and the analogous inequality obtained by swapping the roles of ρ_1 and ρ_2, allows us to conclude (60).

For (61), we consider $\rho = \rho_1$ and $\sigma = \rho_2$ in (60). Noticing the trivial fact that $D(\rho\|\sigma) \geq D(\sigma\|\sigma)$ immediately yields the desired inequality. □

The uniform continuity bound for the relative entropy in the first argument provided in (60) can be compared with the findings of [33, Eq. (43) and (44)] (also based on the previous work [13]), where it was shown that

$$|D(\rho_1\|\sigma) - D(\rho_2\|\sigma)| \leq \max_{i=1,2} \log\left(1 + \frac{\|\rho_1 - \rho_2\|_\infty}{\lambda_{\min}(\rho_i)\lambda_{\min}(\sigma)}\right), \tag{67}$$

whenever $\rho_1, \rho_2 \in \mathcal{D}_+(\mathcal{H})$ and $\min_{i=1,2} \lambda_{\min}(\rho_i) > \|\rho_1 - \rho_2\|_\infty$. Additionally, (60) can be compared with

$$|D(\rho_1\|\sigma) - D(\rho_2\|\sigma)| \leq \varepsilon \log \lambda_{\min}(\sigma)^{-1} + (1+\varepsilon) h\Big(\frac{\varepsilon}{1+\varepsilon}\Big), \tag{68}$$

from [10, Corollary 5.9]. The comparison between (60), (67) and (68) is made explicit in Figure 1, where it is clear that our new bound (60) outperforms the others.

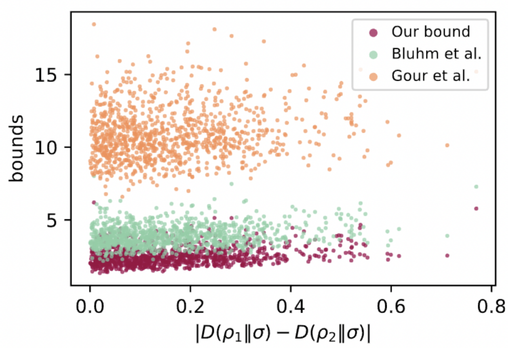


Fig. 1. Comparison between the bounds provided in (60) [our bound], (67) [Gour et al.] and (68) [Bluhm et al.]. Here, we have taken $d = 15$, and we have randomly generated 1000 triples of density matrices ρ_1, ρ_2 and σ, for which we have plotted the corresponding values of the three bounds considered. This shows that (60) is always slightly better than (68), and better than (67).

Next, we can easily extend Theorem 6 to a continuity bound in both arguments of the relative entropy in the following way.

Corollary 1. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ $\sigma_1, \sigma_2 \in \mathcal{D}_+(\mathcal{H})$ with $\frac{1}{2}\|\rho_1 - \rho_2\|_1 = \varepsilon$ and $\frac{1}{2}\|\sigma_1 - \sigma_2\|_1 \leq \delta$ where $\varepsilon, \delta \in (0,1)$. Then*

$$\begin{aligned} &|D(\rho_1\|\sigma_1) - D(\rho_2\|\sigma_2)| \\ &\leq \varepsilon \log\Big(e^{\max\{D_{\max}(\rho_+\|\sigma_1), D_{\max}(\rho_-\|\sigma_2)\}} - 1\Big) \\ &\quad + \log\big(1 + \delta\lambda_{\min}(\sigma)^{-1}\big) + h(\varepsilon)\,, \end{aligned} \tag{69}$$

where $\rho_1 - \rho_2 = \varepsilon\rho_+ - \varepsilon\rho_-$, and $\rho_\pm$ is defined via the Jordan-Hahn decomposition of $(\rho_1 - \rho_2)$ as in (18), and $\lambda_{\min}(\sigma) := \min\{\lambda_{\min}(\sigma_1), \lambda_{\min}(\sigma_2)\}$.

Proof. Let us first begin by rewriting the difference of relative entropies in terms of another difference of relative entropies for which the second states are identical:

$$\begin{aligned} D(\rho_1\|\sigma_1) - D(\rho_2\|\sigma_2) &= D(\rho_1\|\sigma_1) - D(\rho_2\|\sigma_1) \\ &\quad + \operatorname{Tr}(\rho_2(\log\sigma_2 - \log\sigma_1)). \end{aligned} \tag{70}$$

For the first two terms above, we use (63) and (66) from the proof of 6, obtaining the first and last terms of the RHS of (69). For the last term in (70), note that $\sigma_2 \leq e^{D_{\max}(\sigma_2\|\sigma_1)}\sigma_1$. Hence, $\log\sigma_2 \leq D_{\max}(\sigma_2\|\sigma_1)\mathbb{1} + \log\sigma_1$. This in turn implies that

$$\log\sigma_2 - \log\sigma_1 \leq D_{\max}(\sigma_2\|\sigma_1)\mathbb{1} = \mathbb{1}\log\|\sigma_1^{-1/2}\sigma_2\sigma_1^{-1/2}\|_\infty. \tag{71}$$

From this we get

$$\begin{aligned} \operatorname{Tr}(\rho_2(\log\sigma_2 - \log\sigma_1)) &\leq (\operatorname{Tr}\rho_2) D_{\max}(\sigma_2\|\sigma_1) \\ &= \log\|\sigma_1^{-1/2}\sigma_2\sigma_1^{-1/2}\|_\infty\,. \end{aligned} \tag{72}$$

Since we can write σ_2 as

$$\sigma_2 = \sigma_1^{1/2}\left(\sigma_1^{-1/2}(\sigma_2 - \sigma_1)\sigma_1^{-1/2} + \mathbb{1}\right)\sigma_1^{1/2}, \tag{73}$$

we have from (72)

$$\begin{aligned} &\operatorname{Tr}(\rho_2(\log\sigma_2 - \log\sigma_1)) \\ &\quad\leq \log\|\sigma_1^{-1/2}(\sigma_2 - \sigma_1)\sigma_1^{-1/2} + \mathbb{1}\|_\infty \\ &\quad\leq \log\left(1 + \|\sigma_1^{-1}\|_\infty\|\sigma_2 - \sigma_1\|_\infty\right) \\ &\quad\leq \log\left(1 + \delta\lambda_{\min}^{-1}(\sigma)\right)\,, \end{aligned} \tag{74}$$

where we have used the fact that $\|\sigma_2 - \sigma_1\|_\infty \leq (1/2)\|\sigma_2 - \sigma_1\|_1 = \delta$ which can be deduced from the Jordan-Hahn decomposition. ☐

To the best of our knowledge, the only previously existing bound in the form of (69) is that of [10, Theorem 5.13], where it was shown that

$$\begin{aligned} &|D(\rho_1\|\sigma_1) - D(\rho_2\|\sigma_2)| \\ &\leq \left(\varepsilon + \frac{3\delta}{1 - \frac{\lambda_{\min}(\sigma)}{2}}\right)\log\left(2\lambda_{\min}(\sigma)^{-1}\right) \\ &\quad + (1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right) + 2\log\left(1 + \frac{2\lambda_{\min}(\sigma)^{-1}\delta}{1 - \frac{\lambda_{\min}(\sigma)}{2} + \delta}\right). \end{aligned} \tag{75}$$

The RHS of inequality (60) gives a more manageable estimate that outperforms (75).

Remark 4. Note that the right hand sides of (57) of Theorem 5, (60) of Theorem 6, and (69) of Corollary 1, respectively, are generally not monotonically increasing in ε, and hence it is important to note that we assumed in all cases that the trace distance $\frac{1}{2}\|\rho_1 - \rho_2\|_1$ is exactly equal to ε and not just upper bounded by it. If one wants to obtain a result where one only assumes an upper bound $\frac{1}{2}\|\rho_1 - \rho_2\|_1 \leq \varepsilon'$, one will have to take a supremum of the RHS of (60) over all $\varepsilon \leq \varepsilon'$. It is easy to see that the RHS of (60) is monotonically increasing until it hits its maximum at $\varepsilon = 1 - 1/d_A^2$, and hence one then obtains

$$\begin{aligned} &|S(A|B)_{\rho_1} - S(A|B)_{\rho_2}| \\ &\leq \begin{cases} \varepsilon'\log(d_A^2 - 1) + h(\varepsilon'), & \varepsilon' < 1 - 1/d_A^2 \\ \log(d_A^2), & \varepsilon' \geq 1 - 1/d_A^2 \end{cases}, \end{aligned} \tag{76}$$

which is similar to the result obtained in [25, Theorem 5]. Analogously, for Theorem 6 one gets

$$\begin{aligned} &|D(\rho_1\|\sigma) - D(\rho_2\|\sigma)| \\ &\leq \begin{cases} \varepsilon'\log(M - 1) + h(\varepsilon'), & \varepsilon' < 1 - 1/M \\ \log(M), & \varepsilon' \geq 1 - 1/M \end{cases}, \end{aligned} \tag{77}$$

with $M := e^{\max\{D_{\max}(\rho_+\|\sigma), D_{\max}(\rho_-\|\sigma)\}}$, which is similar to the result obtained in [25, Theorem 1]. Finally, for Corollary 1, note that this problem is not present in the factor involving σ_1 and σ_2, for which we only need to assume $\frac{1}{2}\|\sigma_1 - \sigma_2\|_1 \leq \delta$. Therefore, as a consequence of this and (77), one gets

$$|D(\rho_1\|\sigma_1) - D(\rho_2\|\sigma_2)| \leq \begin{cases} \varepsilon' \log(M-1) + \log\big(1 + \delta\lambda_{\min}(\sigma)^{-1}\big) + h(\varepsilon'), & \varepsilon' < 1 - 1/M \\ \log(M) + \log\big(1 + \delta\lambda_{\min}(\sigma)^{-1}\big), & \varepsilon' \geq 1 - 1/M \end{cases}. \tag{78}$$

VII. Extension of the fundamental inequality (Theorem 1) to the infinite-dimensional setting

The fundamental inequality (Theorem 1) is also valid in the infinite-dimensional setting in following form[2]:

$$S(\rho_1) + \varepsilon S(\rho_-) \leq S(\rho_2) + \varepsilon S(\rho_+) + h(\varepsilon)\,,$$

where one or both sides may equal $+\infty$. This can be derived from the finite-dimensional version by approximation.

We will use the homogeneous extension of the von Neumann entropy $S(\rho) = \mathrm{Tr}(\eta(\rho))$ (where $\eta(x) = -x\log x$) to the positive cone of trace class operators $\mathcal{T}_+(\mathcal{H})$ on a separable Hilbert space $\mathcal{H}$, defined as

$$S(\rho) := (\mathrm{Tr}\,\rho) S(\rho/\mathrm{Tr}\,\rho) = \mathrm{Tr}(\eta(\rho)) - \eta(\mathrm{Tr}\,\rho) \tag{79}$$

for any nonzero operator ρ in $\mathcal{T}_+(\mathcal{H})$ and equal to 0 at the zero operator [34].

It is easy to see (cf., f.i., [35, p.1541]) that

$$S(c\rho) = cS(\rho), \quad c \geq 0, \tag{80}$$

and

$$S(\rho + \sigma) \leq S(\rho) + S(\sigma) + h(\mathrm{Tr}\,\rho, \mathrm{Tr}\,\sigma), \tag{81}$$

for any ρ and σ in $\mathcal{T}_+(\mathcal{H})$, where $h(\mathrm{Tr}\,\rho, \mathrm{Tr}\,\sigma) = \eta(\mathrm{Tr}\,\rho) + \eta(\mathrm{Tr}\,\sigma) - \eta(\mathrm{Tr}(\rho+\sigma))$ is the homogeneous extension of the binary entropy to the positive cone in $\mathbb{R}^2$.

Note: An equality holds in (81) if and only if $\rho\sigma = 0$.

Assume that ρ_1, ρ_2 are arbitrary states in $\mathcal{D}(\mathcal{H})$ (where in infinite dimensions $\mathcal{D}(\mathcal{H})$ denotes the positive trace class operators with unit trace) and fix $\rho_\pm$ to be the unique states defined through the Jordan-Hahn decomposition $\rho_1 - \rho_2 = \Delta_+ - \Delta_- = \varepsilon(\rho_+ - \rho_-)$, where $\varepsilon = \frac{1}{2}\|\rho_1 - \rho_2\|_1 \neq 0$. We want to show that

$$S(\rho_1) + \varepsilon S(\rho_-) \leq S(\rho_2) + \varepsilon S(\rho_+) + h(\varepsilon) \tag{82}$$

(where one or both sides may be equal to $+\infty$) using our Theorem 1 valid for finite rank states.

Let $\{P_n\}$ and $\{Q_n\}$ be non-decreasing sequences of spectral projectors of the operators Δ_+ and Δ_- strongly converging, respectively, to the projectors P_* and Q_* onto the supports of Δ_+ and Δ_-. Let further $R = I_{\mathcal{H}} - P_* - Q_*$.

[2]We are very grateful to Maksim Shirokov for pointing this out to us.

Let $\mathcal{H}'$ be any finite-dimensional space and ω be a given pure state in $\mathcal{S}(\mathcal{H}')$. Consider the sequences of states

$$\hat{\rho}_{1,n} = (P_n + Q_n + R)\rho_1(P_n + Q_n + R) \oplus p_n\omega,$$

and

$$\hat{\rho}_{2,n} = (P_n + Q_n + R)\rho_2(P_n + Q_n + R) \oplus q_n\omega,$$

in $\mathcal{S}(\mathcal{H}\oplus\mathcal{H}')$, where $p_n = 1 - \mathrm{Tr}((P_n + Q_n + R)\rho_1)$ and $q_n = 1 - \mathrm{Tr}((P_n + Q_n + R)\rho_2)$. Employing Jordan-Hahn again we decompose $\hat{\rho}_{1,n} - \hat{\rho}_{2,n}$ into $\hat{\Delta}_{\pm,n}$ and note that

$$\hat{\Delta}_{+,n} = P_n\Delta_+ \oplus [p_n - q_n]_+\omega,$$

and

$$\hat{\Delta}_{-,n} = Q_n\Delta_- \oplus [p_n - q_n]_-\omega.$$

($[x]_+ = \max\{x, 0\}$, $[x]_- = \max\{-x, 0\}$). Indeed, it is clear that the positive operators in the RHS of the above expressions have orthogonal supports. Note also that the difference between these operators is equal to

$$\begin{aligned}(P_n + &Q_n + R)(\Delta_+ - \Delta_-)(P_n + Q_n + R) \\ &\oplus ([p_n - q_n]_+ - [p_n - q_n]_-)\omega \\ &= (P_n + Q_n + R)(\rho_1 - \rho_2)(P_n + Q_n + R) \\ &\oplus (p_n - q_n)\omega \\ &= \hat{\rho}_{1,n} - \hat{\rho}_{2,n}\,.\end{aligned}$$

Let $\hat{\rho}_{\pm,n} = \varepsilon_n^{-1}\hat{\Delta}_{\pm,n}$ be states in $\mathcal{S}(\mathcal{H}\oplus\mathcal{H}')$, where

$$\begin{aligned}\varepsilon_n &= \frac{1}{2}\|\hat{\rho}_{1,n} - \hat{\rho}_{2,n}\|_1 \\ &= \frac{1}{2}\big(\|(P_n + Q_n + R)(\rho_1 - \rho_2)(P_n + Q_n + R)\|_1 \\ &\quad + |p_n - q_n|\big)\,.\end{aligned}$$

It follows from (80), (81) and the remark after (81) that

$$\begin{aligned}S(\hat{\rho}_{1,n}) &= S((P_n + Q_n + R)\rho_1(P_n + Q_n + R)) \\ &\quad + h(p_n, (1 - p_n))\,, \\ S(\hat{\rho}_{2,n}) &= S((P_n + Q_n + R)\rho_2(P_n + Q_n + R)) \\ &\quad + h(q_n, (1 - q_n))\,,\end{aligned}$$

and that

$$\begin{aligned}S(\hat{\rho}_{+,n}) &= \varepsilon_n^{-1}(S(P_n\Delta_+) + h(c_n, [p_n - q_n]_+))\,, \\ S(\hat{\rho}_{-,n}) &= \varepsilon_n^{-1}(S(Q_n\Delta_-) + h(d_n, [p_n - q_n]_-))\,,\end{aligned}$$

where

$$\begin{aligned}c_n &= \mathrm{Tr}(P_n\Delta_+)\,, \\ d_n &= \mathrm{Tr}(Q_n\Delta_-)\,.\end{aligned}$$

Since $\hat{\rho}_{1,n}$ and $\hat{\rho}_{2,n}$ are finite-rank states for each n, Theorem 1 implies that

$$S(\hat{\rho}_{1,n}) + \varepsilon_n S(\hat{\rho}_{-,n}) \leq S(\hat{\rho}_{2,n}) + \varepsilon_n S(\hat{\rho}_{+,n}) + h(\varepsilon_n) \quad \forall n. \tag{83}$$

By Lemma 4 in [34] we have

$$\lim_{n\to+\infty} S((P_n + Q_n + R)\rho_1(P_n + Q_n + R)) = S(\rho_1) \leq +\infty$$

and

$$\lim_{n\to+\infty} S((P_n+Q_n+R)\rho_2(P_n+Q_n+R)) = S(\rho_2) \le +\infty\,.$$

It is also clear that

$$\lim_{n\to+\infty} S(P_n\Delta_+) = S(P_*\Delta_+) = S(\Delta_+) \le +\infty$$

and

$$\lim_{n\to+\infty} S(Q_n\Delta_-) = S(Q_*\Delta_-) = S(\Delta_-) \le +\infty\,.$$

Since ε_n tends to ε, p_n, q_n tend to 0 and c_n, d_n tend to 1, using these limits relations and the expressions before (83) one can prove (82) by taking the limit in (83) as $n \to +\infty$.

VIII. Strengthened Inequalities

Note that for all our applications in Sections V and VI we relied on the inequality (19) of Theorem 1 instead of the sharper inequality given by (45) of Theorem 3. This is because (19) is simpler. However, the entropic inequalities of Sections V and VI can be easily strengthened by making use of Theorem 3 instead of Theorem 1. For example, the strengthened forms of (51) of Theorem 4, (57) of Theorem 5 and (60) of Theorem 6, are respectively given by

$$|S(\rho_1)-S(\rho_2)| \le \varepsilon\log\left(d\max\{\lambda_{\max}(\rho_-),\lambda_{\max}(\rho_+)\}-1\right) + c\,h\left(\frac{\varepsilon}{c}\right);$$

$$|S(A|B)_{\rho_1}-S(A|B)_{\rho_2}| \le \varepsilon\log\left(d_A^2-1\right) + c\,h\left(\frac{\varepsilon}{c}\right);$$

$$|D(\rho_1\|\sigma)-D(\rho_2\|\sigma)| \le \varepsilon\log\left(e^{\max\{D_{\max}(\rho_+\|\sigma),D_{\max}(\rho_-\|\sigma)\}}-1\right) + c\,h\left(\frac{\varepsilon}{c}\right),$$

where $c = \mathrm{Tr}\left(\rho_2|_{\mathrm{supp}\,\rho_-}\right)$ in all three cases.

Note: Our result, in the finite-dimensional setting, on the uniform continuity bound for the conditional entropy in the case of equal marginals (Theorem 5) was obtained with different techniques by Berta, Lami, and Tomamichel [25].

Acknowledgments

ND would like to thank Mohammad Alhejji for igniting her curiosity in the bound (56). She is grateful to Milan Mosonyi for the invitation to the *Focused Workshop on Rényi Divergences* at the Erdös Center in Budapest, in July 2024, during which the proof of Theorem 1 was completed. It was also at that workshop that Peter Frenkel raised a question which led us to prove Theorem 3, which is a sharper version of Theorem 1, so we would like to thank him too. ND would also like to thank Ludovico Lami, Mario Berta and Marco Tomamichel for helpful discussions in Budapest and Cambridge, where they shared their preliminary results with her, which motivated her and her coauthors to prove Theorems 5 and 6. BB acknowledges support from the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/V52024X/1. MGJ acknowledges support from the Fonds de la Recherche Scientifique – FNRS (Belgium) and by the Wiener-Anspach Foundation (Belgium), and is grateful for the hospitality of the Centre for Mathematical Sciences at the University of Cambridge where this work was performed during an extended research stay. AC and PG acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 470903074 – TRR 352. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) license to any Accepted Manuscript version arising. The authors would like to thank Maksim Shirokov for the proof of the extension of our fundamental inequality (Theorem 1) to the infinite-dimensional setting (see Section VII).

References

[1] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 4, pp. 623–656, 1948.
[2] J. Von Neumann, *Mathematische grundlagen der quantenmechanik*. Springer-Verlag, 2013, vol. 38.
[3] M. Fannes, “A continuity property of the entropy density for spin lattice systems,” *Commun. Math. Phys.*, vol. 31, pp. 291–294, 1973.
[4] D. Petz, *Quantum information theory and quantum statistics*. Berlin, Heidelberg: Springer, 2008.
[5] K. M. R. Audenaert, “A sharp continuity estimate for the von Neumann entropy,” *J. Phys. A: Math. Theor.*, vol. 40, no. 28, p. 8127, 2007.
[6] R. Alicki and M. Fannes, “Continuity of quantum conditional information,” *J. Phys. A: Math. Gen.*, vol. 37, no. 5, pp. L55–L57, 2004.
[7] A. Winter, “Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints,” *Commun. Math. Phys.*, vol. 347, pp. 291–313, 2016.
[8] M. E. Shirokov, “Advanced Alicki-Fannes-Winter method for energy-constrained quantum systems and its use,” *Quantum Inf. Process.*, vol. 19, p. 164, 2020.
[9] ——, “Quantifying continuity of characteristics of composite quantum systems,” *Physica Scripta*, vol. 98, no. 4, p. 042002, mar 2023.
[10] A. Bluhm, Á. Capel, P. Gondolf, and A. Pérez-Hernández, “Continuity of Quantum Entropic Quantities via Almost Convexity,” *IEEE Transactions on Information Theory*, vol. 69, no. 9, pp. 5869–5901, Sep. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10129917/
[11] A. Bluhm, Á. Capel, P. Gondolf, and A. Pérez-Hernández, “General continuity bounds for quantum relative entropies,” *2023 IEEE International Symposium on Information Theory (ISIT)*, pp. 162–167, 2023.
[12] B. Synak-Radtke and M. Horodecki, “On asymptotic continuity of functions of quantum states,” *J. Phys. A: Math. Gen.*, vol. 39, no. 26, p. L423, 2006.
[13] G. Gour and M. Tomamichel, “Entropy and relative entropy from information-theoretic principles,” *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6313–6327, 2021.
[14] A. E. Rastegin, “Upper continuity bounds on the relative q-entropy for $q > 1$,” *J. Math. Phys.*, vol. 52, p. 062203, 2011.
[15] D. Reeb and M. M. Wolf, “Tight bound on relative entropy by entropy difference,” *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1458–1473, 2015.
[16] M. E. Shirokov, “Uniform continuity bounds for information characteristics of quantum channels depending on input dimension and on input energy,” *J. Phys. A: Math. Theor.*, vol. 52, no. 1, p. 014001, 2019.
[17] ——, “Adaptation of the Alicki-Fannes-Winter method for the set of states with bounded energy and its use,” *Reports on Mathematical Physics*, vol. 81, no. 1, pp. 81–104, 2018.
[18] ——, “Quantum relative entropy: general convergence criterion and preservation of convergence under completely positive linear maps,” *Matematicheskii Sbornik*, vol. 213, no. 12, p. 137–174, 2022. [Online]. Available: http://dx.doi.org/10.4213/sm9794
[19] M. Mosonyi and F. Hiai, “On the quantum Rényi relative entropies and related capacity formulas,” *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2474–2487, 2011.
[20] A. Bluhm, Á. Capel, P. Gondolf, and T. Möbus, “Unified framework for continuity of sandwiched Rényi divergences,” *arXiv:2308.12425*, 2023.
[21] A. Gilyén and A. Poremba, “Improved quantum algorithms for fidelity estimation,” *arXiv:2203.15993*, 2022.
[22] I. H. Kim, “Modulus of convexity for operator convex functions,” *J. Math. Phys.*, vol. 55, p. 082201, 2014.
[23] I. Kim and M. Ruskai, “Bounds on the concavity of quantum entropy,” *J. Math. Phys.*, vol. 55, p. 092201, 2014.

[24] N. Datta, "Min- and max- relative entropies and a new entanglement monotone," *IEEE Transactions on Information Theory*, vol. 55, pp. 2816–2826, 2009.
[25] M. Berta, L. Lami, and M. Tomamichel, "Continuity of entropies via integral representations," *arXiv:2408.15226*, 2024. [Online]. Available: http://arxiv.org/abs/2408.15226
[26] M. M. Wilde, "Optimal uniform continuity bound for conditional entropy of classical–quantum states," *Quantum Information Processing*, vol. 19, no. 2, Jan. 2020. [Online]. Available: http://dx.doi.org/10.1007/s11128-019-2563-4
[27] M. A. Alhejji, *Some Problems Concerning Quantum Channels and Entropies*. PhD thesis, University of Colorado Boulder, 2023.
[28] Z. Zhang, "Estimating mutual information via kolmogorov distance," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3280–3282, 2007.
[29] M. A. Alhejji and G. Smith, "A tight uniform continuity bound for equivocation," *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 2270–2274, 2020.
[30] M. Tomamichel, *Quantum Information Processing with Finite Resources*. Springer International Publishing, 2016. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-21891-5
[31] H. Umegaki, "Conditional expectation in an operator algebra. IV. Entropy and information," *Kodai Mathematical Journal*, vol. 14, no. 2, Jan. 1962.
[32] K. M. R. Audenaert and J. Eisert, "Continuity bounds on the quantum relative entropy — II," *J. Math. Phys.*, vol. 52, p. 112201, 2011.
[33] G. Gour and M. Tomamichel, "Optimal extensions of resource measures and their applications," *Phys. Rev. A*, vol. 102, p. 062401, Dec 2020.
[34] G. Lindblad, "Expectations and entropy inequalities for finite quantum systems," *Communications in Mathematical Physics*, vol. 39, pp. 111–119, 1974.
[35] M. E. Shirokov, "The continuity of the output entropy of positive maps," *Sbornik: Mathematics*, vol. 202, no. 10, p. 1537, 2011.

Koenraad Audenaert Koenraad Audenaert obtained a "Burgerlijk Ingenieur" degree in Electrical Engineering in 1989, a PhD in Computer Science in 1995, and a Master's Degree in Physics in 1999, all from the University of Ghent, Belgium. From 1999 to 2007 he was a postdoctoral researcher at the Catholic University of Leuven, Belgium, and at Imperial College London. In 2007 he took up a Lectureship at the Mathematics Department of Royal Holloway, University of London, and was promoted to Professor in 2016. Since 2022 works at the Institute for Theoretical Physics at the University of Ulm, Germany. His main research interests lie at the intersection of Matrix Analysis and Quantum Information Theory, including entanglement theory, quantum Shannon theory and quantum metrology.

Bjarne Bergh received his PhD degree in mathematics of information from the University of Cambridge (United Kingdom) under the supervision of Nilanjana Datta in 2024. He is now a Postdoctoral Researcher at the University of Cambridge working on problems in Quantum Information Theory.

Nilanjana Datta obtained a PhD in mathematical physics from ETH Zurich, Switzerland, in 1996. From 1997 to 2000, she was a Postdoctoral Researcher at the Dublin Institute of Advanced Studies, C.N.R.S. Marseille, and EPFL in Lausanne. In 2001 she joined the University of Cambridge, U.K., as a Lecturer in Mathematics of Pembroke College, and a member of the Statistical Laboratory in the Centre for Mathematical Sciences. She is currently a Professor of Quantum Information Theory in the Department of Applied Mathematics and Theoretical Physics of the University of Cambridge, and a Fellow of Pembroke College. Her scientific interests include quantum information theory and mathematical physics.

Michael G. Jabbour obtained a Ph.D. degree in engineering sciences from École polytechnique de Bruxelles, Brussels, Belgium, in 2018. From 2019 to 2024, he was a Postdoctoral Researcher at the University of Cambridge (U.K.), DTU Denmark, and École polytechnique de Bruxelles. Since 2024, he has been an Associate Professor at Télécom SudParis, Institut Polytechnique de Paris, Palaiseau, France. His scientific interests include quantum information theory, quantum optics, and mathematical physics.

Ángela Capel obtained a Ph.D. degree in mathematics from Universidad Autónoma de Madrid, Spain, in 2019. From 2020 to 2021, she was a Postdoctoral Researcher at the Technical University of Munich, hosted by the Chair of Mathematical Physics. From 2021 to 2024, she was a Junior Professor at the Department of Mathematics of Eberhard Karls University of Tübingen, Germany. Since 2024, she is an Assistant Professor of Quantum Information Theory and Theoretical Quantum Computation at the Department of Applied Mathematics and Theoretical Physics of the University of Cambridge. Her research interests include quantum information theory, mathematical physics, and quantum many-body systems.

Paul Gondolf received his Bachelor's degree in Physics from the University in Heidelberg, Germany in 2020 and then pursued a Master's degree in Mathematical Physics at the Eberhard-Karls University of Tübingen, Germany. He graduated there in 2022 and then began his PhD in mathematics under the supervision of Ángela Capel, Marius Lemm and Cambyse Rouzé in the same year. His research interests lie in matrix analysis, convex analysis and bosonic open quantum systems.

Ann. Henri Poincaré *Online First*

https://doi.org/10.1007/s00023-024-01519-x

Annales Henri Poincaré

Unified Framework for Continuity of Sandwiched Rényi Divergences

Andreas Bluhm, Ángela Capel, Paul Gondolf and Tim Möbus

Abstract. In this work, we prove uniform continuity bounds for entropic quantities related to the sandwiched Rényi divergences such as the sandwiched Rényi conditional entropy. We follow three different approaches: The first one is the "almost additive approach", which exploits the sub-/superadditivity and joint concavity/convexity of the exponential of the divergence. In our second approach, termed the "operator space approach", we express the entropic measures as norms and utilize their properties for establishing the bounds. These norms draw inspiration from interpolation space norms. We not only demonstrate the norm properties solely relying on matrix analysis tools but also extend their applicability to a context that holds relevance in resource theories. By this, we extend the strategies of Marwah and Dupuis as well as Beigi and Goodarzi employed in the sandwiched Rényi conditional entropy context. Finally, we merge the approaches into a mixed approach that has some advantageous properties and then discuss in which regimes each bound performs best. Our results improve over the previous best continuity bounds or sometimes even give the first continuity bounds available. In a separate contribution, we use the ALAFF method, developed in a previous article by some of the authors, to study the stability of approximate quantum Markov chains.

Contents

1. Introduction
2. Main Results
3. Preliminaries
   3.1. Notation and Basic Concepts
   3.2. Divergences and Entropic Quantities
4. Technical Tools and Main Theorems
   4.1. Almost Additive Approach

Published online: 26 December 2024

Birkhäuser

4.2. Operator Spaces Approach
4.3. Mixed Approach
5. Continuity Bounds for Sandwiched Rényi Divergences
5.1. Continuity Bounds for the Sandwiched Rényi Conditional Entropy
5.2. Continuity Bounds for the Sandwiched Rényi Mutual Information
5.3. Continuity Bounds for the Sandwiched Rényi Conditional Mutual Information
5.4. Continuity Bounds in the First Argument for Sandwiched Rényi Divergences
5.5. Divergence Bounds for Sandwiched Rényi Divergences
5.6. Distance to Separable States
5.7. Distance Measures in Resource Theories
5.8. Generalized Sandwiched Rényi Mutual Information
6. α-Approximate Quantum Markov Chains
6.1. Continuity Bounds for Non-Variational Rényi Divergences via the ALAFF method
6.2. Application: Approximate Quantum Markov Chains
6.3. General Continuity Bounds Via the ALAFF Method
7. Discussion
Acknowledgements
References

1. Introduction

Entropic quantities are indispensable for classical and quantum information theory to characterize information-theoretic tasks. Examples of such quantities include various forms of (conditional) entropies, (conditional) mutual information, and many others. In applications, one often has especially convenient quantum states for which the entropic quantity of interest can be evaluated explicitly and one would therefore like to reduce nearby quantum states to this case. This is why continuity bounds for entropic quantities have become a ubiquitous tool. They provide upper bounds on

$$\sup\{|g(\rho) - g(\sigma)| : \rho, \sigma \in \mathcal{S}_0, \operatorname{dist}(\rho, \sigma) \leq \varepsilon\}.$$

Here, g is the entropic quantity of interest, $\mathcal{S}_0$ is a suitable subset of the set of quantum states $\mathcal{S}(\mathcal{H})$ in a finite-dimensional Hilbert space $\mathcal{H}$, and $\operatorname{dist}(\cdot, \cdot)$ is an appropriate metric on $\mathcal{S}(\mathcal{H})$ (often the trace distance). If the bound on the supremum only depends on ε and $\mathcal{S}_0$, but not on ρ and σ, then g is *uniformly continuous* on $\mathcal{S}_0$.

One of the earliest continuity statements in quantum information theory is the continuity bound on the von Neumann entropy provided by Fannes in [15], which was later improved in [4,30]. Another well-known bound in quantum information theory is the Alicki-Fannes inequality for the conditional entropy [1], for which an almost tight version was proven by Winter in [42]:

$$|H_\rho(A|B) - H_\sigma(A|B)| \leq 2\varepsilon \log d_A + (1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right),$$

where ε is an upper bound on the trace distance between quantum states ρ and σ and h is the binary entropy. Shirokov and others [27,39] realized that the proof method used for this result does not only work for the conditional entropy but can be generalized [35,36]. Shirokov named this approach the Alicki–Fannes–Winter (AFW) method. Recently, the method was further developed in [8] by some of the authors of the present article and applied to quantities based on the Belavkin–Staszewski relative entropy [6].

In this article, we focus on entropic quantities derived from the sandwiched Rényi divergences [28,41]

$$\widetilde{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\log \operatorname{tr}[(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}})^\alpha],$$

where $\alpha \in [1/2,1)\cup(1,\infty)$. Examples of such entropic quantities include the sandwiched Rényi conditional entropy and the sandwiched Rényi mutual information:

$$\widetilde{H}^\uparrow_\alpha(A|B)_\rho := \sup_{\tau_B\in\mathcal{S}(\mathcal{H}_B)} -\widetilde{D}_\alpha(\rho_{AB}\|\mathbb{1}_A\otimes\tau_B)\,, \quad \widetilde{I}^\uparrow_\alpha(A:B)_\rho = \inf_{\substack{\tau_A\in\mathcal{S}(\mathcal{H}_A),\\ \tau_B\in\mathcal{S}(\mathcal{H}_B)}} \widetilde{D}_\alpha(\rho_{AB}\|\tau_A\otimes\tau_B)\,.$$

Recently, there has been increased interest in continuity bounds for the sandwiched Rényi conditional entropy. In [25], it was shown that for $\alpha \in [1/2,1)$,

$$\left|\widetilde{H}^\uparrow_\alpha(A|B)_\rho - \widetilde{H}^\uparrow_\alpha(A|B)_\sigma\right| \leq \log(1+\varepsilon) + \frac{1}{1-\alpha}\log\left(1+\varepsilon^\alpha d_A^{2(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right) \tag{1}$$

and for $\alpha\in(1,\infty)$, they used duality to infer that

$$\left|\widetilde{H}^\uparrow_\alpha(A|B)_\rho - \widetilde{H}^\uparrow_\alpha(A|B)_\sigma\right| \leq \log\left(1+\sqrt{2\varepsilon}) + \frac{1}{1-\beta}\log(1+\sqrt{2\varepsilon}^\beta d_A^{2(1-\beta)} - \frac{\sqrt{2\varepsilon}}{(1+\sqrt{2\varepsilon})^{1-\beta}}\right), \tag{2}$$

where β is such that $\alpha^{-1}+\beta^{-1}=2$.

Using techniques from the interpolation of operator space, it was shown in [5, Theorem 6.2] that for $\alpha\in(1,\infty)$

$$\left|\widetilde{H}^\uparrow_\alpha(A|B)_\rho - \widetilde{H}^\uparrow_\alpha(A|B)_\sigma\right| \leq \alpha'\log(1+2\varepsilon d_A^{2/\alpha'}), \tag{3}$$

where $\alpha' = \alpha/(\alpha-1)$. The authors of [5] note that their bound is better than Eq. (2) for large α, but that it diverges for $\alpha\to1$.

On a high level, the proof of [25] uses sub-/ superadditivity of the exponential of the sandwiched Rényi divergence, while [5] makes a connection to norms on interpolation spaces. Based on these ideas we develop a unified approach to proving not only continuity bounds for the sandwiched Rényi conditional entropy which improves or extends the ones discussed above, but further allows us to prove bounds for related entropic quantities such as the sandwiched Rényi mutual information. In particular, we introduce a new family of norms defined as optimizations over amalgamations with compact convex sets of positive operators.

Generally, we consider uniform continuity bounds on quantities of the form

$$\widetilde{D}_{\alpha,\mathcal{C}}(\rho) := \inf_{\sigma\in\mathcal{C}} \widetilde{D}_{\alpha}(\rho\|\sigma)\,,$$

where $\mathcal{C}$ is a compact convex subset of the quantum states containing at least one full-rank element. Such quantities have appeared in the context of resource theories (see, e.g., [3,23,32,43]) and were termed entropy of resource. Resource theories provide a framework to answer questions about the interconvertability of states, using only allowed operations. Every resource theory has two main ingredients: (i) the set of free states, i.e., states that do not possess the resource (ii) the set of free operations which map the set of free states to itself, i.e., these operations do not create the resource. The best-known example of such a theory is the theory of entanglement, in which the free states are the separable states and the free operations are the local quantum operations and classical communication (LOCC). Others include, for example, the resource theories of coherence and asymmetry.

One way to quantify the resourcefulness of a quantum state in a given resource theory is to measure its distance to the set of free states. Common choices of distance measure include the relative entropy [3,23] and the sandwiched Rényi entropies [32,43]. In the latter case, for $\mathcal{C}$ the set of free states, $\widetilde{D}_{\alpha,\mathcal{C}}$ is the corresponding resource measure. The results of this article can therefore be used to quantify the continuity of popular resource measures such as the Rényi relative entropies of entanglement and coherence studied, e.g., in [43].

The article is organized as follows: In Sect. 2, we present the main results of this paper, before continuing with the necessary preliminaries on divergences and entropic quantities in Sect. 3. We will follow three different approaches in this article: an almost additive approach, an approach based on operator spaces, and one where both methods are mixed. The tools for all these approaches are developed in Sect. 4, such that all the continuity bounds will follow straightforwardly from the theorems proven in this section. In Sect. 5, we derive and discuss our continuity bounds on the sandwiched Rényi divergences and their derived entropic quantities. In Sect. 6, we showcase how continuity bounds can be useful for studying approximate quantum Markov chains. Notably, the continuity bounds in this section do not stem from the three approaches mentioned previously but use the ALAFF method introduced in [8] by some of the present authors. Finally, the paper finishes with a short discussion in Sect. 7.

2. Main Results

In this section, we summarize the main results—the achieved continuity bounds for the sandwiched Rényi conditional entropy, mutual information, and the divergence itself with a fixed second argument. All quantities are defined via the *sandwiched Rényi divergence*, i.e.,

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log(\widetilde{Q}_\alpha(\rho\|\sigma)) = \frac{1}{\alpha-1}\log \operatorname{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right].$$

Then, the *sandwiched Rényi conditional entropy* is defined as

$$\widetilde{H}^\uparrow_\alpha(A|B)_\rho := -\sup_{\tau_B\in\mathcal{S}(\mathcal{H}_B)} \widetilde{D}_\alpha(\rho_{AB}\|\,\mathbb{1}_A\otimes\tau_B)\,,$$

the *mutual information* by

$$\widetilde{I}^\uparrow_\alpha(A:B)_\rho = \inf_{\substack{\tau_A\in\mathcal{S}(\mathcal{H}_A),\\ \tau_B\in\mathcal{S}(\mathcal{H}_B)}} \widetilde{D}_\alpha(\rho_{AB}\|\tau_A\otimes\tau_B)\,,$$

and the *divergence* itself is just considered as a function in the first argument with a fixed second argument. A precise definition of these quantities can be found in Sect. 3.2.

To prove bounds for these quantities, we explore three different methods (see Sect. 4).

- **The almost additive approach:** This approach is inspired by [25] and uses joint convexity/concavity and super-/subadditivity of $\widetilde{Q}_\alpha$. The name is motivated by the fact that this super-/subadditivity reduces to almost concavity of $\rho\mapsto D(\rho\|\sigma)$, respectively, almost convexity of the von Neuman entropy if $\alpha\to 1$. The main result in this approach, from which the respective continuity bounds follow straightforwardly, is a continuity bound on the distance to a compact convex subset of the quantum states in Theorem 4.4.
- **The operator space approach:** It is inspired by [5] and relates $\widetilde{Q}_\alpha$ to a norm. In this context, showing the norm properties poses a challenge, yet yields continuity bounds directly through the triangle inequality. In particular, we define new quantities for $1\le q'\le p'\le\infty$ such that $\frac{1}{r}=\frac{1}{q'}-\frac{1}{p'}$, namely

$$\|\cdot\|^*_{\mathcal{C},p',q'}:\mathcal{B}(\mathcal{H})\to[0,\infty),\quad X\mapsto\|X\|^*_{\mathcal{C},p',q'} := \inf_{c\in\mathcal{C},c>0}\|c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}}\|_{p'}$$

  where $X\in\mathcal{B}(\mathcal{H})$ and $\|\cdot\|_{p'}=\left(\operatorname{tr}[|\cdot|^{p'}]\right)^{\frac{1}{p'}}$. Usually, $\mathcal{C}$ is considered a subalgebra of $\mathcal{B}(\mathcal{H})$. Our proof, however, extends this to compact convex subsets of $\mathcal{B}(\mathcal{H})$ consisting of positive semidefinite operators containing at least one full-rank state. We prove, without having to resort to interpolation theory, that the dual quantity is subadditive on positive semi-definite elements, i.e.,

$$\|X+Y\|^*_{\mathcal{C},p',q'}\le\|X\|^*_{\mathcal{C},p',q'}+\|Y\|^*_{\mathcal{C},p',q'}$$

for $X, Y \in \mathcal{B}_{\geq 0}(\mathcal{H})$. This yields another continuity bound on the distance to a compact convex subset of the quantum states in Theorem 4.17, which allows us to directly infer all our continuity bounds based on this approach in the following.

- **The mixed approach:** It combines the previous two approaches. The main theorem in this approach is Theorem 4.20.

The table below illustrates the various upper bounds that our work encompasses with each bound having a range in α on which it is superior. Note that one bound was already previously proven in [25], which we marked in the table with the reference. All other bounds are new.

Theorem. *Let $\rho, \sigma, \tau \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$*[1] *with* $\ker \tau \subseteq \ker \rho \cap \ker \sigma$ *and* $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$*, define $\widetilde{m}_\tau$ to be the minimal nonzero eigenvalue of τ, and $m = \min\{d_A, d_B\}$. Then, the following continuity bounds for the sandwiched Rényi conditional entropy, mutual information, and divergence in the first input (3.2) hold, i.e., for an entropic quantity g_α, one finds upper bounds on $|g_\alpha(\rho) - g_\alpha(\sigma)|$:*

Approach	Continuity bound	α
Conditional entropy (5.1)		
A. additive	$\log\left(1+\varepsilon\right) + \frac{1}{1-\alpha}\log(1+\varepsilon^\alpha d_A^{2(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right)$ [25]	$[\frac{1}{2}, 1)$
	$\log(1+\varepsilon) + \frac{1}{\alpha-1}\log(1+\varepsilon d_A^{2(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}})$	$(1, \infty)$
Op. space	$\frac{\alpha}{\alpha-1}\log\left(1+\varepsilon d_A^{2\frac{\alpha-1}{\alpha}}\right)$	$(1, \infty)$
Mixed	$\log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon d_A^{2\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right)$	$(1, \infty)$
Mutual Info. (5.2)		
A. additive	$2\log\left(1+\varepsilon^{\frac{1}{\alpha}}\right) + \frac{1}{1-\alpha}\log\left(1+\varepsilon^\alpha m^{2(1-\alpha)} - \frac{\varepsilon^{\frac{1}{\alpha}}}{(1+\varepsilon^{\frac{1}{\alpha}})^{2(1-\alpha)}}\right)$	$[\frac{1}{2}, 1)$
	$2\log\left(1+\varepsilon^{\frac{1}{\alpha}}\right) + \frac{1}{\alpha-1}\log\left(1+\varepsilon^{\frac{1}{\alpha}} m^{2(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)}}\right)$	$(1, \infty)$
1st Entry of divergence (5.4)		
'A. additive[1]	$\log(1+\varepsilon) + \frac{1}{1-\alpha}\log(1+\varepsilon^\alpha \widetilde{m}_\tau^{\alpha-1} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}})$	$[\frac{1}{2}, 1)$
	$\log(1+\varepsilon) + \frac{1}{\alpha-1}\log\left(1+\varepsilon \widetilde{m}_\tau^{1-\alpha} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\right)$	$(1, \infty)$
Op. space[1]	$\frac{\alpha}{\alpha-1}\log\left(1+\varepsilon \widetilde{m}_\tau^{\frac{1-\alpha}{\alpha}}\right)$	$(1, \infty)$
Mixed[1]	$\log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon \widetilde{m}_\tau^{\frac{1-\alpha}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right)$	$(1, \infty)$

The continuity bound directly implies the same bound as a divergence bound 5.5 by choosing $\sigma = \tau$

For the various results presented in the table, it is important to note that none is universally preferable across the entire interval of α. Rather, each method possesses its strengths and limitations. In the following, we compare our results using the sandwiched Rényi conditional entropy. The accompanying

[1] Note that this includes the case $S(\mathcal{H})$ by choosing $\mathcal{H}_A = \mathcal{H}$ and $\mathcal{H}_B = \mathbb{C}$.

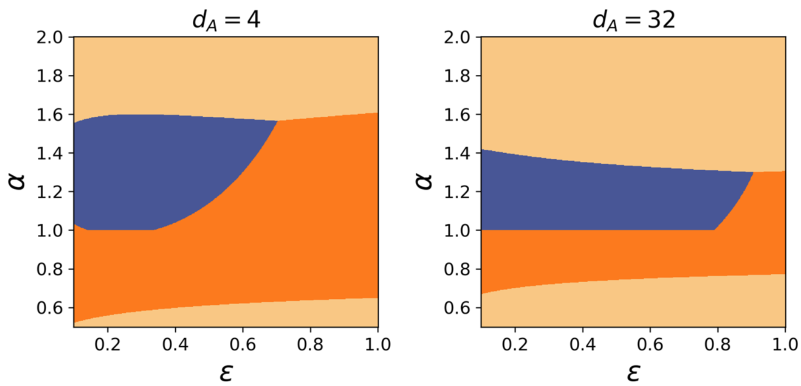


FIGURE 1. Continuity bounds for $\tilde{H}^{\uparrow}_{\alpha}(A|B)_{\rho}$ proven by the almost additive, operator space, and mixed approach, where the visible colour indicates the tightest bound

diagram (Fig. 1) demonstrates the almost additive method is particularly well-suited for α close to 1 and small dimensions. The mixed method, which as the name suggests is a combination of the almost additive and operator space method, can even improve this property in large dimensions, while the operator space method is better suited for large α. Even though the mixed method results in a slightly weaker bound, it performs well across the entire regime $\alpha \in (1, \infty)$. The continuity bounds proven for the mutual information and the first entry of the divergence come with similar scaling. In comparison to the existing results mentioned in Eqs. (1), (2), and (3), our bound proven by the almost additive approach performs in the regime $\alpha \in (1, \infty)$ by a power of two better than the result in [25] for small ε. The operator space approach improves the result in [5] by an order of two. Furthermore, [5,25] treat only the case of the sandwiched Rényi conditional entropy, whereas our versions can also be used for other entropic quantities such as the sandwiched Rényi mutual information and the sandwiched Rényi divergence with fixed second argument.

For even more general quantities, where none of the inputs of the sandwiched Rényi divergence is fixed, we apply the ALAFF method introduced in [8] to estimate the continuity bounds for the $\widetilde{Q}_{\alpha}$ and related quantities of the sandwiched Rényi divergences. The analysis of the bounds is worked out in Sect. 6 and applied in the context of approximate quantum Markov chains.

As covered in the introduction, more results exist for the limiting case $\alpha \to 1$. Here, the sandwiched Rényi conditional entropy converges to the usual quantum conditional entropy, the sandwiched Rényi mutual information to the usual quantum mutual information and the divergence itself to the relative entropy. Similarly, for the limit $\alpha \to \infty$, the quantities converge, respectively, to the min-conditional entropy, the max-mutual information and the max-divergences. Moving forward, we will focus on exploring the limits to compare our methods with each other but also with already established continuity bounds.

Corollary. *Let $\rho, \sigma, \tau, \varepsilon, \widetilde{m}_\tau$, and m be defined as in the beginning of Theorem 2. Then, the obtained continuity bounds converge for $\alpha \to 1$ or $\alpha \to \infty$ in the following way:*

	Approach	$\alpha \to 1$	$\alpha \to \infty$
Conditional Entropy (5.1)	A. additive	$2\varepsilon \log d_A + (1+\varepsilon)h(\frac{\varepsilon}{1+\varepsilon})$	$\log(1+\varepsilon) + 2\log d_A$
	Op. space	∞	$\log(1+\varepsilon d_A^2)$
	Mixed	$2\varepsilon \log d_A + (1+\varepsilon)h(\frac{\varepsilon}{1+\varepsilon})$	$\log(1 + \varepsilon d_A^2 + \varepsilon(1 + \varepsilon(d_A^2-1)))$
Mutual Info. (5.2)	A. additive	$2\varepsilon \log m + 2(1+\varepsilon)h(\frac{\varepsilon}{1+\varepsilon})$	$\log 4m^2$
1^{st} Entry of Divergence (5.4)	A. additive	$\varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon)h(\frac{\varepsilon}{1+\varepsilon})$	$\log(1+\varepsilon) + \log(\widetilde{m}_\tau^{-1})$
	Op. space	∞	$\log(1+\varepsilon\widetilde{m}_\tau^{-1})$
	Mixed	$\varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon)h(\frac{\varepsilon}{1+\varepsilon})$	$\log(1 + \varepsilon\frac{1}{\widetilde{m}_\tau} + \varepsilon(1 + \varepsilon(\frac{1}{\widetilde{m}_\tau} - 1)))$

The corollary shows that the almost additive and mixed approach for the sandwiched Rényi conditional entropy converge for $\alpha \to 1$ to the established bound by Winter [42] and thereby recover the best-known results. For the other quantities, we achieve similar bounds. Moreover, it is noteworthy that both the operator space and mixed method converge to a similar limit, which is markedly superior to that of the almost additive method, as the latter does not vanish for $\varepsilon \to 0$. Furthermore, the almost additive and mixed approach applied to the first entry of the divergence reduce for $\alpha \to 1$ to the bound proven in [8].

Thus, our three new approaches have allowed us to prove good continuity bounds for many quantities of interest related to sandwiched Rényi divergences. Additionally, the ALAFF method allows us to derive slightly worse, albeit more general, continuity bounds with applications in the context of approximate quantum Markov chains.

3. Preliminaries

3.1. Notation and Basic Concepts

We start by fixing the notation used in this paper. We denote by $\mathcal{H}$ a complex Hilbert space with an inner product linear in the second argument. Throughout this paper $\mathcal{H}$ is finite-dimensional with dimension $d \in \mathbb{N}$. For a bipartite or tripartite system, we will always use indices with capital letters to refer to the different subsystems. If we have, for example, the bipartite space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and consider the dimension of $\mathcal{H}_A$, we write d_A.

The set of bounded linear operators on the Hilbert space $\mathcal{H}$ will be denoted by $\mathcal{B}(\mathcal{H})$ and its convex subset of positive semi-definite operators with trace one, i.e. the quantum states, by $\mathcal{S}(\mathcal{H})$.

We use $\operatorname{tr}[\,\cdot\,]$ for the usual matrix trace and $\|\cdot\|_1$ and $\|\cdot\|_\infty$ to denote the trace norm and the spectral norm on $\mathcal{B}(\mathcal{H})$, respectively. More generally, we set $\|X\|_p := \operatorname{tr}[|X|^p]^{1/p}$, which coincides with the Schatten p-norms for $p \in [1, \infty]$.

Moreover, for a state ρ on a bipartite system $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, we set $\rho_A \in \mathcal{B}(\mathcal{H}_A)$ to be the output-state of the partial trace. The partial trace is a completely positive trace-preserving (CPTP) map. Finally, we denote by $\mathbb{1}_A$ the identity matrix on A and, with a slight abuse of notation, we denote by $\operatorname{tr}_A[\cdot]$ both the partial trace of A as well as the corresponding map on $\mathcal{H}_{AB}$ by tensorizing with $\mathbb{1}_A$. In the first case, we mean just the usual definition of the partial trace as a map from $\mathcal{B}(\mathcal{H}_{AB})$ to $\mathcal{B}(\mathcal{H}_B)$ while in the second case we mean $\mathbb{1}_A \otimes \operatorname{tr}_A[\cdot]$

Throughout the text, we will use the logarithm in natural basis and will denote it by log.

3.2. Divergences and Entropic Quantities

In this section, we will introduce the entropic quantities considered in the present article. We start with the sandwiched Rényi divergence, which is the base for all subsequently defined quantities.

Definition 3.1 (*Sandwiched Rényi divergence*). Let X, $Y \in \mathcal{B}(\mathcal{H})$ be positive semi-definite operators with $\ker X \supseteq \ker Y$. For $\alpha \in [1/2, 1) \cup (1, \infty)$, we define

$$\widetilde{Q}_\alpha(X\|Y) := \operatorname{tr}[(Y^{\frac{1-\alpha}{2\alpha}} X Y^{\frac{1-\alpha}{2\alpha}})^\alpha] = \|Y^{\frac{1-\alpha}{2\alpha}} X Y^{\frac{1-\alpha}{2\alpha}}\|_\alpha^\alpha.$$

In case the power in $Y^{\frac{1-\alpha}{2\alpha}}$ becomes negative these quantities have to be understood as positive powers of the pseudoinverse of Y. Then, the *sandwiched Rényi divergence* of X and Y is

$$\widetilde{D}_\alpha(X\|Y) := \frac{1}{\alpha - 1} \log \widetilde{Q}_\alpha(X\|Y)\,.$$

Remark 3.2. Alternatively, we can write

$$\widetilde{Q}_\alpha(X\|Y) = \operatorname{tr}\left[\left(X^{\frac{1}{2}} Y^{\frac{1-\alpha}{\alpha}} X^{\frac{1}{2}}\right)^\alpha\right] = \|X^{\frac{1}{2}} Y^{\frac{1-\alpha}{\alpha}} X^{\frac{1}{2}}\|_\alpha^\alpha,$$

because the operator $f(A^*A)$ for $f(0) = 0$ is defined by applying the continuous real function f on the singular values of A. The simple trick $A^*Av = \lambda v \implies AA^*Av = \lambda Av$ shows that the singular values of A^*A and AA^* are equal such that $\operatorname{tr}[f(A^*A)] = \operatorname{tr}[f(AA^*)]$.

It is known that the sandwiched Rényi divergences converge in the limits $\alpha \to 1$ and $\alpha \to \infty$ to well-known quantities [40, Section 4.3.2]:

Proposition 3.3. *Let ρ, $\sigma \in \mathcal{S}(\mathcal{H})$. Then, $\widetilde{D}_\alpha(\rho\|\sigma)$ converges to*

- *the relative entropy $D(\rho\|\sigma) := tr[\rho(\log\rho - \log\sigma)]$ for $\alpha \to 1$.*
- *the max-entropy $D_\infty(\rho\|\sigma) := \inf\{\lambda > 0 \,:\, \rho \leq e^\lambda \sigma\}$ (see also Eq. (3.10)) for $\alpha \to \infty$.*

For $\alpha = \frac{1}{2}$*, it holds that* $\widetilde{D}_{1/2}(\rho\|\sigma) = -\log F(\rho,\sigma)$*, where* F *is the fidelity* $F(\rho,\sigma) = \left(tr[|\sqrt{\rho}\sqrt{\sigma}|]\right)^2$.

Next, we define the sandwiched Rényi conditional entropy in the same spirit as the conditional entropy in terms of the relative entropy.

Definition 3.4 (Sandwiched Rényi conditional entropy). Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, for $\alpha \in [1/2, 1) \cup (1, \infty)$, the *sandwiched Rényi conditional entropy* is given by

$$\widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\rho} := \sup_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \frac{1}{1-\alpha} \log \widetilde{Q}_{\alpha}(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B)\,.$$

Mimicking how the mutual information arises from the relative entropy, we arrive at the sandwiched Rényi mutual information:

Definition 3.5 [*Sandwiched Rényi mutual information*]. Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, for $\alpha \in [1/2, 1) \cup (1, \infty)$, we define the *sandwiched Rényi mutual information* as

$$\widetilde{I}^{\uparrow}_{\alpha}(A : B)_{\rho} := \inf_{\sigma_A, \sigma_B} \widetilde{D}_{\alpha}(\rho_{AB} \| \sigma_A \otimes \sigma_B),$$

where the infimum is taken over $\sigma_A \in \mathcal{S}(\mathcal{H}_A)$ and $\sigma_B \in \mathcal{S}(\mathcal{H}_B)$.

Finally, we define the sandwiched Rényi conditional mutual information in the same spirit as the conditional mutual information in terms of the relative entropy. In this case, we base our definition on the difference between (sandwiched Rényi) conditional entropies.

Definition 3.6 (*Sandwiched Rényi conditional mutual information*). Let $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$. Then, for $\alpha \in [1/2, 1) \cup (1, \infty)$, the *sandwiched Rényi conditional mutual information* is given by

$$\widetilde{I}^{\uparrow}_{\alpha}(A : C|B)_{\rho} := \widetilde{H}^{\uparrow}_{\alpha}(C|B)_{\rho} - \widetilde{H}^{\uparrow}_{\alpha}(C|AB)_{\rho}\,.$$

Note that the infimum in the definition of the mutual information and the conditional entropy is attained at the reduced state of ρ_{AB} (see, e.g., [40, Section 5.1]). We will often use that the sandwiched Rényi conditional entropy is bounded by [40, Lemma 5.2]

$$-\log \min\{d_A, d_B\} \leq \widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\rho} \leq \log d_A. \tag{4}$$

As already observed, the entropic quantities defined for the sandwiched Rényi divergences converge for $\alpha \to 1$ to the ones defined by the Umegaki relative entropy. We recall its definition:

Definition 3.7 (*Umegaki relative entropy*). Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ be quantum states with $\ker \sigma \subseteq \ker \rho$. Then, the *Umegaki relative entropy* of ρ and σ is given by

$$D(\rho\|\sigma) := \mathrm{tr}[\rho \log \rho - \rho \log \sigma].$$

We can furthermore define the quantum conditional entropy

Definition 3.8. (Quantum conditional entropy) Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, the *quantum conditional entropy* is given by

$$H(A|B)_\rho := \sup_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} -D(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B) = -D(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B)\,.$$

and the quantum mutual information

Definition 3.9 (*Quantum mutual information*). Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, the *quantum mutual information* is given by

$$I(A:B)_\rho := \inf_{\sigma_A, \sigma_B} D(\rho_{AB} \| \sigma_A \otimes \sigma_B) = D(\rho_{AB} \| \rho_A \otimes \rho_B)\,,$$

where the infimum is taken over $\sigma_A \in \mathcal{S}(\mathcal{H}_A)$ and $\sigma_B \in \mathcal{S}(\mathcal{H}_B)$.

Finally, we consider explicitly the limit of $\alpha \to \infty$ of the sandwiched Rényi divergences:

Definition 3.10 (*Max-divergence*). Let $X, Y \in \mathcal{B}(\mathcal{H})$ be positive semi-definite operators with $X \neq 0$. Then, the *max-divergence* of X and Y is given by

$$D_\infty(X\|Y) := \inf\{\lambda > 0 \,:\, X \leq e^\lambda Y\}\,.$$

Note that it admits the following equivalent representation:

$$D_\infty(X\|Y) := \log \|Y^{-1/2} X Y^{-1/2}\|_\infty\,.$$

Definition 3.11 (*Min-conditional entropy*). Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, the *min-conditional entropy* is given by

$$H^\uparrow_\infty(A|B)_\rho := \sup_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} -D_\infty(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B)\,.$$

Likewise, we can define a max-mutual information:

Definition 3.12 (*Max-mutual information*). Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, we define the *max-mutual information* as

$$I^\uparrow_\infty(A:B)_\rho := \inf_{\sigma_A, \sigma_B} D_\infty(\rho_{AB} \| \sigma_A \otimes \sigma_B),$$

where the infimum is taken over $\sigma_A \in \mathcal{S}(\mathcal{H}_A)$ and $\sigma_B \in \mathcal{S}(\mathcal{H}_B)$.

It has been shown that $\alpha \to \widetilde{D}_\alpha(\rho\|\sigma)$ is monotonically increasing [40, Corollary 4.2]. Thus, in particular,

$$\widetilde{D}_\alpha(\rho\|\sigma) \leq D_\infty(\rho\|\sigma) \qquad \forall \alpha \in [1/2, 1) \cup (1, \infty). \tag{5}$$

4. Technical Tools and Main Theorems

In this section, we will introduce the technical tools and prove the main theorems that form the cornerstones for our proofs of the continuity bounds in Sect. 5, which are just corollaries of the former.

4.1. Almost Additive Approach

We start by reviewing some tools that are related to the almost additive approach to sandwiched Rényi divergences (see [28,40] for an overview). One property is that $\widetilde{Q}_\alpha$ is jointly convex/concave:

Lemma 4.1 *(Joint convexity and concavity of $\widetilde{Q}_\alpha$). For quantum states ρ, $\sigma \in \mathcal{S}(\mathcal{H})$, the function*

$$(\rho, \sigma) \mapsto \widetilde{Q}_\alpha(\rho\|\sigma)$$

is jointly concave for $\alpha \in [1/2, 1)$ and jointly convex for $\alpha \in (1, \infty)$.

Proof. For a proof, see [40, Proposition 4.7 and Theorem 4.1].

Another useful property of the $\widetilde{Q}_\alpha$ is that it behaves nicely under addition, unlike the sandwiched Rényi divergences:

Lemma 4.2 *(*Sub- and Superadditivity of *$\widetilde{Q}_\alpha$). For X_1, X_2, $Y \in \mathcal{B}(\mathcal{H})$ positive semi-definite with $\ker Y \subseteq \ker X_1 \cap \ker X_2$, we find that for $\alpha \in (0, 1)$*

$$\widetilde{Q}_\alpha(X_1 + X_2\|Y) \leq \widetilde{Q}_\alpha(X_1\|Y) + \widetilde{Q}_\alpha(X_2\|Y)$$

and further for $\alpha \in (1, \infty)$

$$\widetilde{Q}_\alpha(X_1\|Y) + \widetilde{Q}_\alpha(X_2\|Y) \leq \widetilde{Q}_\alpha(X_1 + X_2\|Y)\,.$$

Proof. The proof of the first claim can be found in [25], based on an inequality from [26]. For the second, we use that one can write

$$\widetilde{Q}_\alpha(X_1 + X_2\|Y) = \mathrm{tr}[(X_1' + X_2')^\alpha] = \|(X_1' + X_2')^\alpha\|_1\,,$$

with $X_i' := Y^{\frac{1-\alpha}{2\alpha}} X_i Y^{\frac{1-\alpha}{2\alpha}}$, $i = 1, 2$. As before Y^{-1} is the pseudoinverse of Y. Since $\|\cdot\|_1$ is unitarily invariant and since the map $\mathbb{R}_+ \ni x \mapsto x^\alpha$ is convex for $\alpha > 1$ and vanishes at zero, we can apply [10, Theorem 1.2] to conclude

$$\|(X_1'+X_2')^\alpha\|_1 \geq \|(X_1')^\alpha + (X_2')^\alpha\|_1 = \mathrm{tr}[(X_1')^\alpha + (X_2')^\alpha] = \widetilde{Q}_\alpha(X_1\|Y) + \widetilde{Q}_\alpha(X_2\|Y)\,.$$

The trace and 1-norm agree as all involved operators are positive semi-definite.

One might be tempted to conjecture that super- and subadditivity holds more generally for the analogues of the $\widetilde{Q}_\alpha$, in the case of Petz or geometric Renyi divergences. However, this is not the case. One can relatively easily construct counterexamples:

Example 4.3. In the following, we present an example which contradicts the superadditivity of the Petz and the geometric Rényi divergence. For ρ, $\tau \in \mathcal{S}(\mathcal{H})$ with $\ker \rho \supseteq \ker \tau$ and $\alpha \in (0, 1) \cup (1, \infty)$,

$$\overline{Q}_\alpha(\rho\|\tau) := \mathrm{tr}[\rho^\alpha \tau^{1-\alpha}]\,, \quad \overline{D}_\alpha(\rho\|\tau) := \frac{1}{\alpha - 1} \log \overline{Q}_\alpha(\rho\|\tau)$$

is the Petz Rényi divergence and for $\alpha \in (1, 2]$, let

$$\widehat{Q}_\alpha(\rho\|\tau) := \mathrm{tr}[\tau^{\frac{1}{2}}(\tau^{-\frac{1}{2}}\rho\tau^{-\frac{1}{2}})^\alpha \tau^{\frac{1}{2}}]\,, \quad \widehat{D}_\alpha(\rho\|\tau) := \frac{1}{\alpha - 1} \log \widehat{Q}_\alpha(\rho\|\tau)$$

be the geometric (maximal) Rényi divergence. For the density matrices

$$\rho_1 = \begin{pmatrix} 0.8 & 0.3 \\ 0.3 & 0.2 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 0.1 & 0.2 \\ 0.2 & 0.9 \end{pmatrix}, \quad \tau = \begin{pmatrix} 0.45 & 0.49 \\ 0.49 & 0.55 \end{pmatrix},$$

one can calculate

$$\overline{Q}_{1.5}(\rho_1\|\tau) + \overline{Q}_{1.5}(\rho_2\|\tau) > 6 > 5.9 > \overline{Q}_{1.5}(\rho_1 + \rho_2\|\tau)$$

and

$$\widehat{Q}_{1.5}(\rho_1\|\tau) + \widehat{Q}_{1.5}(\rho_2\|\tau) > 9 > 6 > \widehat{Q}_{1.5}(\rho_1 + \rho_2\|\tau)$$

which contradicts the superadditivity.

We conclude the subsection on the almost additive approach by proving our fundamental technical result from which all our continuity bounds following the almost additive approach will be obtained.

Theorem 4.4 *(*Distance to a convex, compact set*). Let $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ be a convex, compact set that contains at least one positive definite state. Then the map*

$$\widetilde{D}_{\alpha,\mathcal{C}} : \mathcal{S}(\mathcal{H}) \to \mathbb{R}, \qquad \rho \mapsto \widetilde{D}_{\alpha,\mathcal{C}}(\rho) := \inf_{\tau\in\mathcal{C}} \widetilde{D}_\alpha(\rho\|\tau)$$

is uniformly continuous (cf. [33, Definition 4.18]) for $\alpha \in [1/2, 1) \cup (1, \infty)$. For ρ, $\sigma \in \mathcal{S}(\mathcal{H})$ with $\frac{1}{2}\|\rho - \sigma\|_1 \le \varepsilon$, $\alpha \in [1/2, 1)$ and κ (see Remark 4.5) such that $\sup_{\rho\in\mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \le \log(\kappa) < \infty$ we get

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \le \log(1+\varepsilon) + \frac{1}{1-\alpha}\log\left(1 + \varepsilon^\alpha\kappa^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right)$$

Further for $\alpha \in (1, \infty)$ and κ (see Remark 4.5) such that $\sup_{\rho\in\mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \le \log(\kappa) < \infty$

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \le \log(1+\varepsilon) + \frac{1}{\alpha-1}\log\left(1 + \varepsilon\kappa^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\right).$$

Proof. We will only demonstrate here the proof for the second inequality ($\alpha > 1$), as the proof for the first ($\alpha < 1$) is almost completely analogous. The proof is inspired by [25]. Without loss of generality, we can assume that $\frac{1}{2}\|\rho-\sigma\|_1 = \varepsilon$, as the bound is monotone in ε. We have

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| = \frac{1}{\alpha-1}\left|\log\frac{\inf_{\tau\in\mathcal{C}}\widetilde{Q}_\alpha(\rho\|\tau)}{\inf_{\tau\in\mathcal{C}}\widetilde{Q}_\alpha(\sigma\|\tau)}\right|$$

using the monotonicity of the logarithm. Hence the result reduces to an upper bound on

$$\frac{\inf_{\tau\in\mathcal{C}}\widetilde{Q}_\alpha(\rho\|\tau)}{\inf_{\tau\in\mathcal{C}}\widetilde{Q}_\alpha(\sigma\|\tau)} \quad \text{and} \quad \frac{\inf_{\tau\in\mathcal{C}}\widetilde{Q}_\alpha(\sigma\|\tau)}{\inf_{\tau\in\mathcal{C}}\widetilde{Q}_\alpha(\rho\|\tau)}.$$

We only upper bound the first fraction as the bound on the second is achieved by swapping the roles of ρ and σ. First, there are orthonormal quantum states

μ, ν such that $\rho + \varepsilon\mu = \sigma + \varepsilon\nu$. Using the superadditivity of $\widetilde{Q}_\alpha$ (Lemma 4.2) and the α-homogenity in the first argument gives

$$\widetilde{Q}_{\alpha,\mathcal{C}}(\rho) + \varepsilon^\alpha \widetilde{Q}_{\alpha,\mathcal{C}}(\mu) \le \inf_{\tau\in\mathcal{C}} \Big(\widetilde{Q}_\alpha(\rho\|\tau) + \varepsilon^\alpha \widetilde{Q}_\alpha(\mu\|\tau) \Big) \le \widetilde{Q}_{\alpha,\mathcal{C}}(\rho + \varepsilon\mu)$$

where we set $\widetilde{Q}_{\alpha,\mathcal{C}}(X) := \inf_{\tau\in\mathcal{C}} \widetilde{Q}_\alpha(X\|\tau)$ for a positive semi-definite operator X. Joint convexity of the $\widetilde{Q}_\alpha$ and their α-homogenity further allows us to write

$$\begin{aligned}
\widetilde{Q}_{\alpha,\mathcal{C}}(\sigma + \varepsilon\nu) &= (1+\varepsilon)^\alpha \inf_{\tau_1,\tau_2\in\mathcal{C}} \widetilde{Q}_\alpha\Big(\tfrac{1}{1+\varepsilon}\sigma + \tfrac{\varepsilon}{1+\varepsilon}\nu \Big\| \tfrac{1}{1+\varepsilon}\tau_1 + \tfrac{\varepsilon}{1+\varepsilon}\tau_2 \Big) \\
&\le (1+\varepsilon)^{\alpha-1} \left(\inf_{\tau_1\in\mathcal{C}} \widetilde{Q}_\alpha(\sigma\|\tau_1) + \inf_{\tau_2\in\mathcal{C}} \varepsilon \widetilde{Q}_\alpha(\nu\|\tau_2) \right) \\
&= (1+\varepsilon)^{\alpha-1} \Big(\widetilde{Q}_{\alpha,\mathcal{C}}(\sigma) + \varepsilon \widetilde{Q}_{\alpha,\mathcal{C}}(\nu) \Big)
\end{aligned}$$

where we split the infimum in the second argument into an equivalent infimum over a convex combination $\frac{1}{1+\varepsilon}\tau_1 + \frac{\varepsilon}{1+\varepsilon}\tau_2$ which allows us to split the infimum later. Moreover, it holds that $1 \le \widetilde{Q}_{\alpha,\mathcal{C}}(\nu) \le \kappa^{\alpha-1}$ for any state ν, where the lower bound stems from the non-negativity of the sandwiched α-Rényi divergences on quantum states and the upper bound holds by assumption. Putting these estimates together, using $\rho + \varepsilon\mu = \sigma + \varepsilon\nu$, we find

$$\widetilde{Q}_{\alpha,\mathcal{C}}(\rho) \le (1+\varepsilon)^{\alpha-1} \left(\widetilde{Q}_{\alpha,\mathcal{C}}(\sigma) + \varepsilon\kappa^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}} \right).$$

We therefore obtain

$$\frac{\widetilde{Q}_{\alpha,\mathcal{C}}(\rho)}{\widetilde{Q}_{\alpha,\mathcal{C}}(\sigma)} \le (1+\varepsilon)^{\alpha-1} \left(1 + \frac{\varepsilon\kappa^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}}{\widetilde{Q}_{\alpha,\mathcal{C}}(\sigma)} \right) \le (1+\varepsilon)^{\alpha-1} \left(1 + \varepsilon\kappa^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}} \right)$$

where in the second inequality we lower bound $\widetilde{Q}_{\alpha,\mathcal{C}}(\sigma)$ by 1. This is valid, since for $\varepsilon \in (0,1)$, $\varepsilon\kappa^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}} \ge \varepsilon - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}} \ge 0$. Repeating the procedure for the other fraction gives the same bound and hence concludes the claim.

Remark 4.5 (Existence of (uniform) κ). A $\kappa < \infty$ upper bound on $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot)$ and uniform in all $\alpha \in [1/2, 1) \cup (1, \infty)$ always exists, since we can upper bound $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot) \le \log \|\tau^{-1}\|_\infty$ independently of α, where τ is a full-rank state in $\mathcal{C}$ (which by assumption exists). For the continuity bounds the κ can, however, also depend on α.

Before considering the limit cases, we prove that limits can be exchanged with the infimum in the above definition of $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot)$:

Lemma 4.6. *Let $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ be a compact set that contains at least one positive definite state. Then the following identities hold:*

$$\lim_{\alpha\to 1} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) = \inf_{\tau\in\mathcal{C}} \lim_{\alpha\to 1} \widetilde{D}_\alpha(\rho\|\tau) = \inf_{\tau\in\mathcal{C}} D(\rho\|\tau) =: \widetilde{D}_{1,\mathcal{C}}(\rho)$$

and

$$\lim_{\alpha\to\infty} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) = \inf_{\tau\in\mathcal{C}} \lim_{\alpha\to\infty} \widetilde{D}_\alpha(\rho\|\tau) = \inf_{\tau\in\mathcal{C}} D_\infty(\rho\|\tau) =: \widetilde{D}_{\infty,\mathcal{C}}(\rho).$$

Proof. Let $\eta \in \mathcal{C}$ be positive definite, which exists by assumption. We note first that the infimum $\inf_{\tau\in\mathcal{C}} \widetilde{D}_\alpha(\rho\|\tau)$ is attained [2, Theorem 2.43] since $\mathcal{C}$ is compact and $\sigma \mapsto \widetilde{D}(\rho\|\sigma)$ is lower semi-continuous for any fixed $\rho \in \mathcal{S}(\mathcal{H})$ [20, Proposition 4.5]. Then,

$$\widetilde{D}_\infty(\rho\|\eta) =: c_\infty < \infty.$$

Moreover,

$$\inf_{\tau\in\mathcal{C}} \widetilde{D}_\alpha(\rho\|\tau) \le \widetilde{D}_\alpha(\rho\|\eta) \le c_\infty$$

for all $\alpha \in (1,\infty]$, since the sandwiched Rényi divergences are monotonically increasing in α [40, Corollary 4.2]. Next, we define $\mathcal{C}(\rho) = \{\tau \in \mathcal{C} \mid \widetilde{D}_\infty(\rho\|\tau) \le c_\infty\}$ which satisfies by the above

$$\inf_{\tau\in\mathcal{C}} \widetilde{D}_\alpha(\rho\|\tau) = \inf_{\tau\in\mathcal{C}(\rho)} \widetilde{D}_\alpha(\rho\|\tau)$$

for all $\alpha \in (1,\infty]$. The set $\mathcal{C}(\rho)$ is compact, because the preimage of $(-\infty, c_\infty]$ under a lower semi-continuous function is closed. Hence, also the infimum on the right-hand side is attained. Moreover, the function $\mathcal{C}(\rho) \ni \tau \mapsto D_\alpha(\rho\|\tau)$ is continuous for all $\alpha \in (1,\infty]$. Therefore, Dini's theorem [2, Theorem 2.66] shows that $\mathcal{C}(\rho) \ni \tau \mapsto D_\alpha(\rho\|\tau)$ converges uniformly for $\alpha \to 1$, $\alpha \to \infty$, since we have pointwise convergence by [40, Section 4.3.2] (see also Proposition 3.3). Thus, the assertion follows from [14, Lemma I.7.6].

Lemma 4.7 (Limits). *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ with $\frac{1}{2}\|\rho-\sigma\|_1 \le \varepsilon$ and κ a bound on $\widetilde{D}_{\alpha,\mathcal{C}}$ independent of α (see Remark 4.5), then the limit $\alpha \to 1$ of the bounds obtained in Theorem 4.4 gives*

$$|\widetilde{D}_{1,\mathcal{C}}(\rho) - \widetilde{D}_{1,\mathcal{C}}(\sigma)| \le \varepsilon \log\kappa + (1+\varepsilon) h\Big(\frac{\varepsilon}{1+\varepsilon}\Big)$$

where $h(\cdot)$ is the binary entropy. Unless ε is trivial, i.e. $\varepsilon = 0$, we find for $\alpha \to \infty$

$$|\widetilde{D}_{\infty,\mathcal{C}}(\rho) - \widetilde{D}_{\infty,\mathcal{C}}(\sigma)| \le \log(1+\varepsilon) + \log\kappa$$

which is no longer a continuity bound

Proof. Using l'Hospital's rule, we find that

$$\begin{aligned}
&\lim_{\alpha\nearrow 1}\left[\frac{1}{1-\alpha}\log(1+\varepsilon^\alpha\kappa^{(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}})\right]\\
&\quad= -\lim_{\alpha\nearrow 1}\left[\frac{\log(\varepsilon)\varepsilon^\alpha\kappa^{(1-\alpha)} - \varepsilon^\alpha\kappa^{(1-\alpha)}\log\kappa - \log(1+\varepsilon)\frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}}{1+\varepsilon^\alpha\kappa^{(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}}\right]\\
&\quad= -\varepsilon\log(\varepsilon) + \varepsilon\log\kappa + \varepsilon\log(1+\varepsilon)
\end{aligned}$$

and similarly

$$\begin{aligned}
&\lim_{\alpha\searrow 1}\left[\frac{1}{\alpha-1}\log(1+\varepsilon\kappa^{(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}})\right]\\
&\quad= \lim_{\alpha\searrow 1}\left[\frac{\varepsilon\kappa^{(\alpha-1)}\log\kappa - \log(\varepsilon)\frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}} + \log(1+\varepsilon)\frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}}{1+\varepsilon\kappa^{(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}}\right]
\end{aligned}$$

$$= \varepsilon \log \kappa - \varepsilon \log(\varepsilon) + \varepsilon \log(1+\varepsilon)$$

which proves the limit $\alpha \to 1$. Another use of l'Hospital's rule shows that

$$\lim_{\alpha\to\infty}\left[\frac{1}{\alpha-1}\log(1+\varepsilon\kappa^{(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}})\right] = \log\kappa.$$

4.2. Operator Spaces Approach

In this section, we will construct a family of norms inspired by the norms on interpolation spaces over von Neumann algebras defined in [5, Theorem 4.5]. These norms have an explicit characterization given by a supremum of amalgamations with elements from a different von Neumann subalgebra in a Schatten p-norm. In this section, we show that in the finite-dimensional setting, this construction is not limited to amalgamation with a von Neumann subalgebra, but can be generalized to a convex compact set of positive semi-definite operators, which contains at least one positive definite state. We further give a direct proof of a triangle-like inequality for a map arising from these norms, previously only shown in an abstract and more restrictive setting. To be more precise, our goal is to show that for positive semi-definite operators the map

$$X \mapsto \inf_{c\in\mathcal{C},c>0} \|c^{-\frac{1}{2r}} X c^{-\frac{1}{2r}}\|_p$$

satisfies a triangle inequality and some other properties that come in handy when proving continuity bounds later. Here r is implicitly constrained by the explicit p and an implicit q satisfying $p \geq q \geq 1$. The choice to not use r as a defining parameter will become clear later when we establish duality relations and p and q transform to their Hölder conjugates while r is unaffected. $\mathcal{C}$ is a convex, compact set of positive semi-definite operators containing at least one positive definite element.

We begin with the construction of the auxiliary norms.

Definition 4.8 (*The $\mathcal{C}, p, q$ norm*). Let $\mathcal{C} \subset \mathcal{B}_{\geq 0}(\mathcal{H})$ be a convex, compact set containing at least one positive definite state. Then for $1 \leq p \leq q \leq \infty$, $\frac{1}{r} := \frac{1}{p} - \frac{1}{q}$, we define

$$\|\cdot\|_{\mathcal{C},p,q} : \mathcal{B}(\mathcal{H}) \to [0,\infty), \quad X \mapsto \|X\|_{\mathcal{C},p,q} := \sup_{c\in\mathcal{C}} \|c^{\frac{1}{2r}} X c^{\frac{1}{2r}}\|_p$$

where $\|\cdot\|_p = (\mathrm{tr}[|\cdot|^p])^{\frac{1}{p}}$ is the Schatten-p-norm.

Lemma 4.9. $\|\cdot\|_{\mathcal{C},p,q} : \mathcal{B}(\mathcal{H}) \to [0,\infty)$ *defines a norm on* $\mathcal{B}(\mathcal{H})$.

Proof. The map is finite for all $X \in \mathcal{B}(\mathcal{H})$, because $\mathcal{C}$ is compact and $c \mapsto c^{\frac{1}{2r}} X c^{\frac{1}{2r}}$ continuous on positive semi-definite matrices. Further, it is positive definite since $\mathcal{C}$ contains a positive-definite state, by assumption. Finally, it satisfies positive homogeneity and triangle inequality because the Schatten-p-norm has this property and the supremum is subadditive.

We will proceed to define the maps we are interested in:

Definition 4.10. Let $\mathcal{C} \subset \mathcal{B}_{\geq 0}(\mathcal{H})$ a convex, compact set containing at least one positive definite state. Then for $1 \leq q' \leq p' \leq \infty$ and $\frac{1}{r} = \frac{1}{q'} - \frac{1}{p'}$, we define

$$\| \cdot \|^*_{\mathcal{C},p',q'} : \mathcal{B}(\mathcal{H}) \to [0, \infty), \quad X \mapsto \|X\|^*_{\mathcal{C},p',q'} := \inf_{c \in \mathcal{C}, c>0} \|c^{-\frac{1}{2r}} X c^{-\frac{1}{2r}}\|_{p'}$$

where $X \in \mathcal{B}(\mathcal{H})$ and $\| \cdot \|_{p'} = \left(\mathrm{tr}[| \cdot |^{p'}]\right)^{\frac{1}{p'}}$.

Remark 4.11. For the above maps, the norm properties are no longer apparent. We get that the map is finite, positive homogenous, and further positive definite (requires a bit more work but relates back to the compactness of $\mathcal{C}$). However, the triangle inequality we can only prove for positive semi-definite operators.

We will now show that on the positive semi-definite operators $\| \cdot \|^*_{\mathcal{C},p',q'}$ satisfies triangle inequality by showing that $\| \cdot \|^*_{\mathcal{C},p',q'}$ agrees with the dual norm of $\| \cdot \|_{\mathcal{C},p,q}$ on this set. Here p is given by $1 = \frac{1}{p'} + \frac{1}{p}$ and q by $1 = \frac{1}{q} + \frac{1}{q'}$. Note that the latter norm is indeed consistent with Definition 4.8 (as choosing q via $1 = \frac{1}{q} + \frac{1}{q'}$ leaves r invariant and satisfies the requirements, i.e. $1 \leq p \leq q \leq \infty$). As a first step, we derive a Hölder-inequality for these maps.

Lemma 4.12 (Hölder-inequality)*. Let $\| \cdot \|_{\mathcal{C},p,q}$ as defined in Definition 4.8 be given, then for all $X, Y \in \mathcal{B}(\mathcal{H})$*

$$|tr[XY]| \leq \|X\|_{\mathcal{C},p,q} \|Y\|^*_{\mathcal{C},p',q'}$$

where $\| \cdot \|^*_{\mathcal{C},p',q'}$ as defined in Definition 4.10 with p' given by $\frac{1}{p} + \frac{1}{p'} = 1$ and q' via $\frac{1}{q} + \frac{1}{q'} = 1$.

Proof. For $c \in \mathcal{C}$ with $c > 0$, using Hölder inequality on Schatten-p-norms, we have

$$\begin{aligned} |\mathrm{tr}[XY]| &= \left|\mathrm{tr}[c^{\frac{1}{2r}} X c^{\frac{1}{2r}} c^{-\frac{1}{2r}} Y c^{-\frac{1}{2r}}]\right| \\ &\leq \|c^{\frac{1}{2r}} X c^{\frac{1}{2r}}\|_p \|c^{-\frac{1}{2r}} Y c^{-\frac{1}{2r}}\|_{p'} . \end{aligned}$$

Now using that $\|c^{\frac{1}{2r}} X c^{\frac{1}{2r}}\|_p \leq \|X\|_{\mathcal{C},p,q}$ we get

$$|\mathrm{tr}[XY]| \leq \|X\|_{\mathcal{C},p,q} \|c^{-\frac{1}{2r}} Y c^{-\frac{1}{2r}}\|_{p'} .$$

Finally taking the infimum over all $c \in \mathcal{C}$, $c > 0$ of the above inequality

$$|\mathrm{tr}[XY]| \leq \|X\|_{\mathcal{C},p,q} \|Y\|^*_{\mathcal{C},p',q'} .$$

We can now characterize $\| \cdot \|_{\mathcal{C},p,q}$ via $\| \cdot \|^*_{\mathcal{C},p',q'}$ as follows.

Lemma 4.13. *Let $\| \cdot \|_{\mathcal{C},p,q}$ as defined in Definition 4.8 be given. Then we obtain*

$$\sup_{Y \in \mathcal{B}(\mathcal{H}),\, \|Y\|^*_{\mathcal{C},p',q'} \leq 1} |tr[XY]| = \|X\|_{\mathcal{C},p,q}$$

where $\| \cdot \|^*_{\mathcal{C},p',q'}$ is the map from Definition 4.10 with p' defined via $\frac{1}{p} + \frac{1}{p'} = 1$ and q' via $\frac{1}{q} + \frac{1}{q'} = 1$.

Proof. Let $X \in \mathcal{B}(\mathcal{H})$ be given. Due to the continuity of $c \mapsto c^{\frac{1}{2r}} X c^{\frac{1}{2r}}$ on $\mathcal{C}$ and compactness of $\mathcal{C}$, there exists an $c_* \in \mathcal{C}$ such that

$$\|X\|_{\mathcal{C},p,q} = \|c_*^{\frac{1}{2r}} X c_*^{\frac{1}{2r}}\|_p$$

Due to the convexity, the existence of a positive definite state τ in $\mathcal{C}$ and the continuity of $c \mapsto c^{\frac{1}{2r}} X c^{\frac{1}{2r}}$, we have that for all $\varepsilon > 0$ there exists a $\delta \in [0, 1]$ such that $c_\delta = (1-\delta)c_* + \delta\tau \in \mathcal{C}$, $c_\delta > 0$ and

$$\|X\|_{\mathcal{C},p,q} \le \|c_\delta^{\frac{1}{2r}} X c_\delta^{\frac{1}{2r}}\|_p + \varepsilon\,.$$

Regarding $\|\cdot\|_{p'}$ as dual norm to $\|\cdot\|_p$ (see e.g. [40, Lemma 3.3]) where $1 = \frac{1}{p'} + \frac{1}{p}$, we find

$$\begin{aligned}\|X\|_{\mathcal{C},p,q} &\le \sup_{W \in \mathcal{B}(\mathcal{H}), \|W\|_{p'} \le 1} \left|\mathrm{tr}[c_\delta^{\frac{1}{2r}} X c_\delta^{\frac{1}{2r}} W]\right| + \varepsilon \\ &= \sup_{Y \in \mathcal{B}(\mathcal{H}), \|c_\delta^{-\frac{1}{2r}} Y c_\delta^{-\frac{1}{2r}}\|_{p'} \le 1} |\mathrm{tr}[XY]| + \varepsilon \\ &\le \sup_{Y \in \mathcal{B}(\mathcal{H}), \|Y\|^*_{\mathcal{C},p',q'} \le 1} |\mathrm{tr}[XY]| + \varepsilon\,,\end{aligned}$$

where the last inequality holds, since $\|c_\delta^{-\frac{1}{2r}} Y c_\delta^{-\frac{1}{2r}}\|_{p'} \le 1$ implies $\|Y\|^*_{\mathcal{C},p',q'} \le 1$. Now taking $\varepsilon \to 0$ gives

$$\|X\|_{\mathcal{C},p,q} \le \sup_{Y \in \mathcal{B}(\mathcal{H}), \|Y\|^*_{\mathcal{C},p',q'} \le 1} |\mathrm{tr}[XY]|\,.$$

The reverse inequality is a direct consequence of Lemma 4.12. We, hence, conclude the claim.

Now we come to the theorem which is at the heart of our continuity bounds.

Theorem 4.14 (A dual formula for $\|\cdot\|^*_{\mathcal{C},p',q'}$). *Let $\|\cdot\|^*_{\mathcal{C},p',q'}$ be as defined in Definition 4.10 and $X \ge 0$. Then*

$$\sup_{Y \in \mathcal{B}_{\ge 0}(\mathcal{H}), \|Y\|_{\mathcal{C},p,q} \le 1} |tr[XY]| = \|X\|^*_{\mathcal{C},p',q'}\,,$$

for p given by $\frac{1}{p} + \frac{1}{p'} = 1$, q' via $\frac{1}{q} + \frac{1}{q'} = 1$ and $\|Y\|_{\mathcal{C},p,q}$ as in Definition 4.8. That is, on positive semi-definite states, $\|\cdot\|^*_{\mathcal{C},p',q'}$ agrees with the dual norm of $\|\cdot\|_{\mathcal{C},p,q}$.

Proof. Let $X \in \mathcal{B}_{\ge 0}(\mathcal{H})$. Then we immediately get that

$$\sup_{Y \in \mathcal{B}_{\ge 0}(\mathcal{H}), \|Y\|_{\mathcal{C},p,q} \le 1} |\mathrm{tr}[XY]| \le \|X\|^*_{\mathcal{C},p',q'}$$

by Lemma 4.12. This inequality furthermore holds for $X \in \mathcal{B}(\mathcal{H})$. To prove the reverse inequality, we define the auxiliary function

$$\begin{aligned}&f_X : \mathcal{B}_{\ge 0}(\mathcal{H}) \times \mathcal{C} \to \mathbb{R}, \\ &\qquad (Y, c) \mapsto f_X(Y, c) = p'\mathrm{tr}[XY] - \frac{p'}{p}\mathrm{tr}[(c^{\frac{1}{2r}} Y c^{\frac{1}{2r}})^p]\,,\end{aligned}$$

where p is given by $\frac{1}{p}+\frac{1}{p'}=1$. Note that the function is lower semi-continuous in c and upper semi-continuous in Y (as it is continuous in both of its arguments). In addition, we will now show, that it is concave in Y and convex in c. The concavity in Y is straightforward. We have that $Y \mapsto \mathrm{tr}[YX]$ is linear in Y, hence in particular concave. Further for a fixed c, we have that $Y \mapsto -\frac{p'}{p}\mathrm{tr}[(c^{\frac{1}{2r}}Yc^{\frac{1}{2r}})^p]$ is concave, since $x \mapsto -x^p$ for $p \geq 1$ is. To show convexity in c, we have to show that for a fixed Y, $c \mapsto \mathrm{tr}[(c^{\frac{1}{2r}}Yc^{\frac{1}{2r}})^p]$ is concave. We first rewrite

$$\mathrm{tr}[(c^{\frac{1}{2r}}Yc^{\frac{1}{2r}})^p] = \mathrm{tr}[(\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y})^p]$$

and then use [40, Eq. (3.16)] to obtain

$$\mathrm{tr}[(\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y})^p] = \sup_{Z\geq 0} p\mathrm{tr}[\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y}Z] - \frac{p}{p'}\mathrm{tr}[Z^{p'}] = \sup_{Z\geq 0} p\mathrm{tr}[\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y}Z^{\frac{1}{p'}}] - \frac{p}{p'}\mathrm{tr}[Z]\,.$$

Since $0 \leq \frac{1}{p'}, \frac{1}{r} \leq 1$ and $\frac{1}{p'}+\frac{1}{r} = \frac{1}{p'}+\frac{1}{q'}-\frac{1}{p'} = \frac{1}{q'} \leq 1$, Lieb's concavity theorem [24] gives us joint concavity of the map

$$(Z,c) \mapsto p\mathrm{tr}\left[\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y}Z^{\frac{1}{p'}}\right] - \frac{p}{p'}\mathrm{tr}[Z]$$

and therefore concavity of

$$c \mapsto \sup_{Z\geq 0} p\mathrm{tr}\left[\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y}Z^{\frac{1}{p'}}\right] - \frac{p}{p'}\mathrm{tr}[Z] = \mathrm{tr}\left[\left(\sqrt{Y}c^{\frac{1}{r}}\sqrt{Y}\right)^p\right] = \mathrm{tr}\left[\left(c^{\frac{1}{2r}}Yc^{\frac{1}{2r}}\right)^p\right]\,.$$

Knowing that f_X is a map from the Cartesian product of a convex set with a convex compact set to the reals, being upper semi-continuous and convex in its first and lower semi-continuous and concave in its second argument, we can employ Sions minimax theorem [22, Theorem 2.18] to find that

$$\sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H})} \inf_{c\in\mathcal{C}} f_X(Y,c) = \inf_{c\in\mathcal{C}} \sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H})} f_X(Y,c)\,.$$

Rewriting the RHS of the above equation gives for $\ker c \subseteq \ker X$

$$\begin{aligned}\sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H})} f_X(Y,c) &= \sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H})} p'\mathrm{tr}[XY] - \frac{p'}{p}\mathrm{tr}[(c^{\frac{1}{2r}}Yc^{\frac{1}{2r}})^p] \\ &= \sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H})} p'\mathrm{tr}[c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}}Y] - \frac{p'}{p}\mathrm{tr}[Y^p] \\ &= \|c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}}\|_{p'}^{p'}\end{aligned}$$

with $\cdot^{-1}$ in this context being the pseudoinverse, and ∞ if $\ker c \not\subseteq \ker X$. Using the lower semi-continuity of $c \mapsto \|c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}}\|_{p'}^{p'}$ and that for every $c \in \mathcal{C}$ we can approximate it via $c_\delta = (1-\delta)c + \delta\tau \in \mathcal{C}$ a positive definite sequence, due to $\tau \in \mathcal{C}$ chosen positive definite and $\delta \in (0,1)$, we have finally

$$\inf_{c\in\mathcal{C}} \sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H})} f_X(Y,c) = \inf_{c\in\mathcal{C},\ker c\subseteq\ker X} \|c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}}\|_{p'}^{p'} = \|X\|_{\mathcal{C},p',q'}^{*}{}^{p'}\,. \quad (6)$$

The LHS becomes

$$\inf_{c\in\mathcal{C}} f_X(Y,c) = p'\mathrm{tr}[XY] - \frac{p'}{p}\sup_{c\in\mathcal{C}}\mathrm{tr}[(c^{\frac{1}{2r}}Yc^{\frac{1}{2r}})^p]$$
$$= p'\mathrm{tr}[XY] - \frac{p'}{p}\|Y\|^p_{\mathcal{C},p,q}$$

As of Equation (6) we have that for all $\varepsilon > 0$ there exists $Y \in \mathcal{B}_{\geq 0}(\mathcal{H})$, s.t.

$$p'\mathrm{tr}[XY] - \frac{p'}{p}\|Y\|^p_{\mathcal{C},p,q} + \varepsilon \geq \|X\|^{*}_{\mathcal{C},p',q'}{}^{p'}$$

This immediately[2] gives that for all $\varepsilon > 0$ there exists $Y \in \mathcal{B}_{\geq 0}(\mathcal{H})$ (w.l.o.g. $Y \neq 0$), s.t.

$$\left(\frac{\mathrm{tr}[XY]}{\|Y\|_{\mathcal{C},p,q}}\right)^{p'} + \varepsilon \geq \|X\|^{*}_{\mathcal{C},p',q'}{}^{p'} . \tag{7}$$

and hence taking an upper bound on the LHS,

$$\left(\sup_{Y\in\mathcal{B}_{\geq 0}(\mathcal{H}),\|Y\|_{\mathcal{C},p,q}\leq 1} |\mathrm{tr}[XY]|\right)^{p'} + \varepsilon \geq \|X\|^{*}_{\mathcal{C},p',q'}{}^{p'}$$

Letting $\varepsilon \to 0$ and taking the p'th root on both sides concludes the claim.

The following corollary is a consequence.

Corollary 4.15. *Let $X, Y \in \mathcal{B}_{\geq 0}(\mathcal{H})$, then*

$$\|X+Y\|^{*}_{\mathcal{C},p',q'} \leq \|X\|^{*}_{\mathcal{C},p',q'} + \|Y\|^{*}_{\mathcal{C},p',q'}$$

Proof. This follows directly from the formula derived in Theorem 4.14 and the subadditivity of the supremum.

We further have the following lemma.

Lemma 4.16. *For $X, Y \in \mathcal{B}_{\geq 0}(\mathcal{H})$ with $X \leq Y$, we have that*

$$\|X\|^{*}_{\mathcal{C},p',q'} \leq \|Y\|^{*}_{\mathcal{C},p',q'} .$$

Proof. We have that for every $c \in \mathcal{C}$, $c > 0$

$$c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}} \leq c^{-\frac{1}{2r}}Yc^{-\frac{1}{2r}}$$

and hence, since the Schatten p-norms preserve order on positive semi-definite operators, we have

$$\|c^{-\frac{1}{2r}}Xc^{-\frac{1}{2r}}\|_p \leq \|c^{-\frac{1}{2r}}Yc^{-\frac{1}{2r}}\|_p$$

concluding the claim.

Putting together the above results, we can now show the following.

[2] To see this we introduce a nonnegative parameter λ and take $g(\lambda) = p'\mathrm{tr}[X(\lambda\cdot Y)] - \frac{p'}{p}\|(\lambda\cdot Y)\|^p_{\mathcal{C},p,q}$. Clearly $\sup_{\lambda\geq 0} g(\lambda) \geq p'\mathrm{tr}[XY] - \frac{p'}{p}\|Y\|^p_{\mathcal{C},p,q}$ with the supremum being achieved at $\lambda^{p-1} = \frac{\mathrm{tr}[XY]}{\|Y\|^p_{\mathcal{C},p,q}}$. Inserting this immediately gives Equation (7).

Theorem 4.17 (Distance to convex, compact set). *Let $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ be a convex, compact set that contains at least one positive definite state. Then the map*

$$\widetilde{D}_{\alpha,\mathcal{C}} : \mathcal{S}(\mathcal{H}) \to \mathbb{R}, \qquad \rho \mapsto \widetilde{D}_{\alpha,\mathcal{C}}(\rho) := \inf_{\tau \in \mathcal{C}} \widetilde{D}_\alpha(\rho\|\tau)$$

is uniformly continuous (cf. [33, Definition 4.18]) for $\alpha \in [1/2, 1) \cup (1, \infty)$. For ρ, $\sigma \in \mathcal{S}(\mathcal{H})$ with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$, $\alpha \in [1/2, 1)$ and κ (see Remark 4.5) such that $\sup_{\rho \in \mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \leq \log(\kappa) < \infty$,

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \leq \frac{1}{1-\alpha} \log(1 + \varepsilon^\alpha \kappa^{1-\alpha}) \,.$$

Further for $\alpha \in (1, \infty)$ and κ (see Remark 4.5) such that $\sup_{\rho \in \mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \leq \log(\kappa) < \infty$ we have

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \leq \frac{\alpha}{\alpha-1} \log\left(1 + \varepsilon \kappa^{\frac{\alpha-1}{\alpha}}\right) \,.$$

Proof. The proof strategy is inspired by [5]. Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Without loss of generality, we can assume that $\frac{1}{2}\|\rho - \sigma\|_1 = \varepsilon$, as both bounds are monotonically increasing in ε. We will begin with the first bound, i.e. the bound for $\alpha < 1$ and note that

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \leq \frac{1}{1-\alpha} \left| \log\left(\frac{\sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\rho\|c)}{\sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\sigma\|c)} \right) \right| .$$

Now, we can use that there exists $\nu, \mu \in \mathcal{S}(\mathcal{H})$ such that $\rho + \varepsilon\nu = \sigma + \varepsilon\mu$ and hence $\rho \leq \sigma + \varepsilon\mu$. Furthermore $\widetilde{Q}_\alpha(\cdot\|c)$ is monotone for all $c \in \mathcal{C}$ and subadditive (c.f. Lemma 4.2), which gives us

$$\sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\rho\|c) \leq \sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\sigma + \varepsilon\mu\|c) \leq \sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\sigma\|c) + \varepsilon^\alpha \sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\mu\|c) \,.$$

If we use that $\kappa^{\alpha-1} \leq \sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\mu\|c)$ and $\sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\sigma\|c) \leq 1$, we find

$$\frac{\sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\rho\|c)}{\sup_{c \in \mathcal{C}} \widetilde{Q}_\alpha(\sigma\|c)} \leq 1 + \varepsilon^\alpha \kappa^{1-\alpha} \,.$$

Repeating the same steps for the inverse fraction gives the claim.

For $\alpha > 1$, we find

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| = \frac{\alpha}{\alpha-1} \left| \log\left(\frac{\|\rho\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}} \right) \right| .$$

Using now that there exist $\nu, \mu \in \mathcal{S}(\mathcal{H})$ such that $\rho + \varepsilon\nu = \sigma + \varepsilon\mu$ and hence $\rho \leq \sigma + \varepsilon\mu$, we can employ Lemma 4.16 and Corollary 4.15

$$\begin{aligned}\frac{\|\rho\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}} &\leq \frac{\|\sigma + \varepsilon\mu\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}} \\ &\leq \frac{\|\sigma\|^*_{\mathcal{C},\alpha,1} + \varepsilon\|\mu\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}} \\ &\leq 1 + \varepsilon\frac{\|\mu\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}}\end{aligned}$$

We can now bound $\|\mu\|^*_{\mathcal{C},\alpha,1}$ from above with $\kappa^{\frac{\alpha-1}{\alpha}}$ and lower bound $\|\sigma\|^*_{\mathcal{C},\alpha,1}$ with 1 from below, as $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot) \geq 0$ on quantum states. We hence obtain

$$\frac{\|\rho\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}} \leq 1 + \varepsilon\kappa^{\frac{\alpha-1}{\alpha}}$$

and the same bound for the inverse quotient. This proves the claim.

Remark 4.18. In [31], the authors prove a continuity bound for $\alpha \in [1/2, 1)$ and $0 \leq \varepsilon \leq (\widetilde{Q}_{\alpha,\mathcal{C}}(\rho))^{1/\alpha}$ given as

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \leq \frac{1}{\alpha - 1}\log\left(1 - \frac{\varepsilon^\alpha}{\widetilde{Q}_{\alpha,\mathcal{C}}(\rho)}\right).$$

This form is convenient for them as they work in the context of resource theories and are looking for dimension independent bounds. This bound immediately implies a continuity bound of the kind that we are looking for, namely

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \leq \frac{1}{\alpha - 1}\log(1 - \varepsilon^\alpha\kappa^{1-\alpha}).$$

for $0 \leq \varepsilon \leq \kappa^{(\alpha-1)/\alpha}$. Using the inequality $\log(1+x) \leq -\log 1 - x$ for $0 \leq x \leq 1$, it can be seen that this is worse than the bound we obtain in Theorem 4.17. Altering the proof in [31] slightly one could, however, also derive the bound in Theorem 4.17, since both proofs are almost identical.

In the following lemma, we investigate the limiting behaviour for the bounds derived above.

Lemma 4.19 (Limits). *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$ and κ (see Remark 4.5) a uniform bound on $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot)$ independent of α, then the limit $\alpha \to \infty$ of the bound derived in Theorem 4.17 is stable and we find that*

$$|\widetilde{D}_{\infty,\mathcal{C}}(\rho) - \widetilde{D}_{\infty,\mathcal{C}}(\sigma)| \leq \log(1 + \varepsilon\kappa).$$

For $\alpha \to 1$ the bound diverges unless it is trivial, i.e. $\varepsilon = 0$.

Proof. Both conclusions are obtained straightforwardly.

4.3. Mixed Approach

For $\alpha > 1$ we notice that both approaches have one limit ($\alpha \to 1$ or $\alpha \to \infty$) in which they perform well, while for the other one they either diverge or do not give a continuity bound anymore. The purpose of this section, therefore, is to combine both approaches and to obtain a bound which performs well in both limits.

Theorem 4.20. (Distance to convex, compact set) *Let $\mathcal{C} \subseteq \mathcal{S}(\mathcal{H})$ be a convex, compact set that contains at least one positive definite state. Then the map*

$$\widetilde{D}_{\alpha,\mathcal{C}} : \mathcal{S}(\mathcal{H}) \to \mathbb{R}, \qquad \rho \mapsto \widetilde{D}_{\alpha,\mathcal{C}}(\rho) := \inf_{\tau \in \mathcal{C}} \widetilde{D}_\alpha(\rho\|\tau)$$

is uniformly continuous (cf. [33, Definition 4.18]) for $\alpha \in (1,\infty)$. For ρ, $\sigma \in \mathcal{S}(\mathcal{H})$ satisfying $\frac{1}{2}\|\rho-\sigma\|_1 \le \varepsilon$ and κ (see Remark 4.5) such that $\sup_{\rho\in\mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \le \log(\kappa) < \infty$ we find

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| \le \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon\kappa^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{\frac{2\alpha-1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right).$$

Proof. Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Without loss of generality, we can assume that $\frac{1}{2}\|\rho - \sigma\|_1 = \varepsilon$, as the bound is monotonically increasing in ε. Since the logarithm and the map $x \mapsto x^{\frac{1}{\alpha}}$ are monotone, we find

$$|\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)| = \frac{\alpha}{\alpha-1}\left|\log\left(\frac{\|\rho\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}}\right)\right|.$$

We further get $\mu, \nu \in \mathcal{S}(\mathcal{H})$, such that $\rho + \varepsilon\mu = \sigma + \varepsilon\nu$. Due to Corollary 4.15 and $\|\nu\|^*_{\mathcal{C},\alpha,1} \le \kappa^{\frac{\alpha-1}{\alpha}}$, we have

$$\|\sigma + \varepsilon\nu\|^*_{\mathcal{C},\alpha,1} \le \|\sigma\|^*_{\mathcal{C},\alpha,1} + \varepsilon\|\nu\|^*_{\mathcal{C},\alpha,1} \le \|\sigma\|^*_{\mathcal{C},\alpha,1} + \varepsilon\kappa^{\frac{\alpha-1}{\alpha}}.$$

We then also have that for $X \ge 0$, $\|X\|^*_{\mathcal{C},\alpha,1} = \left(\inf_{c\in\mathcal{C}} \widetilde{Q}_\alpha(X\|c)\right)^{\frac{1}{\alpha}}$, which gives

$$\begin{aligned}
\|\rho + \varepsilon\mu\|^*_{\mathcal{C},\alpha,1} &= \left(\inf_{c\in\mathcal{C}} \widetilde{Q}_\alpha(\rho+\varepsilon\mu\|c)\right)^{\frac{1}{\alpha}} \\
&\ge \left(\inf_{c\in\mathcal{C}} \widetilde{Q}_\alpha(\rho\|c) + \varepsilon^\alpha \inf_{c\in\mathcal{C}} \widetilde{Q}_\alpha(\mu\|c)\right)^{\frac{1}{\alpha}} \\
&= (1+\varepsilon)^{\frac{1}{\alpha}}\left(\frac{1}{1+\varepsilon}\|\rho\|^{*\,\alpha}_{\mathcal{C},\alpha,1} + \frac{\varepsilon}{1+\varepsilon}\varepsilon^{\alpha-1}\|\mu\|^{*\,\alpha}_{\mathcal{C},\alpha,1}\right)^{\frac{1}{\alpha}} \\
&\ge (1+\varepsilon)^{\frac{1-\alpha}{\alpha}}\left(\|\rho\|^*_{\mathcal{C},\alpha,1} + \varepsilon^{\frac{2\alpha-1}{\alpha}}\|\mu\|^*_{\mathcal{C},\alpha,1}\right) \\
&\ge (1+\varepsilon)^{\frac{1-\alpha}{\alpha}}\left(\|\rho\|^*_{\mathcal{C},\alpha,1} + \varepsilon^{\frac{2\alpha-1}{\alpha}}\right)
\end{aligned}$$

where we used Lemma 4.2, the concavity of $x \mapsto x^{\frac{1}{\alpha}}$ and that $\|\mu\|^*_{\mathcal{C},\alpha,1} \ge 1$, since $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot) \ge 0$ on quantum states. Combining these two bounds, we find

$$\|\rho\|^*_{\mathcal{C},\alpha,1} \le (1+\varepsilon)^{\frac{\alpha-1}{\alpha}}\left(\|\sigma\|^*_{\mathcal{C},\alpha,1} + \varepsilon\kappa^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{\frac{2\alpha-1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right)$$

and as a consequence (since again $\|\sigma\|^*_{\mathcal{C},\alpha,1} \geq 1$ and $\varepsilon\kappa^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{\frac{2\alpha-1}{\alpha}}}{(1+\alpha)^{\frac{\alpha-1}{\alpha}}} > 0$)

$$\frac{\|\rho\|^*_{\mathcal{C},\alpha,1}}{\|\sigma\|^*_{\mathcal{C},\alpha,1}} \leq (1+\varepsilon)^{\frac{\alpha-1}{\alpha}} \left(1 + \varepsilon\kappa^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{\frac{2\alpha-1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right).$$

Repeating the same steps for the inverse fraction proves the claimed bound.

Lemma 4.21 *(Limits). Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$, and κ (see Remark 4.5) be a bound on $\widetilde{D}_{\alpha,\mathcal{C}}(\cdot)$ independent of $\alpha \in (1, \infty)$. Then, the limit $\alpha \to 1$ of the bounds obtained in Theorem 4.20 gives*

$$|\widetilde{D}_{1,\mathcal{C}}(\rho) - \widetilde{D}_{1,\mathcal{C}}(\sigma)| \leq \varepsilon \log \kappa + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right)$$

and for $\alpha \to \infty$

$$|\widetilde{D}_{\infty,\mathcal{C}}(\rho) - \widetilde{D}_{\infty,\mathcal{C}}(\sigma)| \leq \log((1+\varepsilon)(1+\varepsilon\kappa) - \varepsilon^2)$$

Proof. With $\beta = \frac{\alpha-1}{\alpha}$ l'Hospital's rule lets us infer that

$$\lim_{\beta\to 0} (\beta)^{-1} \log(1 + \varepsilon\kappa^\beta - \frac{\varepsilon^{1+\beta}}{(1+\varepsilon)^\beta}) = \lim_{\beta\to 0} \frac{\varepsilon\kappa^\beta \log\kappa - \log(\varepsilon)\frac{\varepsilon^{1+\beta}}{(1+\varepsilon)^\beta} + \log(1+\varepsilon)\frac{\varepsilon^{1+\beta}}{(1+\varepsilon)^\beta}}{1 + \varepsilon\kappa^\beta - \frac{\varepsilon^{1+\beta}}{(1+\varepsilon)^\beta}}$$
$$= \varepsilon \log\kappa - \varepsilon\log\varepsilon + \varepsilon\log(1+\varepsilon),$$

and

$$\lim_{\alpha\to\infty} \frac{\alpha}{\alpha-1} \log\left(1 + \varepsilon\kappa^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right) = \log\left(1 + \varepsilon\kappa - \frac{\varepsilon^2}{(1+\varepsilon)}\right).$$

5. Continuity Bounds for Sandwiched Rényi Divergences

5.1. Continuity Bounds for the Sandwiched Rényi Conditional Entropy

In this section, we prove and compare the bounds for the sandwiched Rényi conditional entropy using the three different approaches: almost-additive, operator space, and mixed. We will start with a discussion of these bounds for the example of the sandwiched Rényi conditional entropies, but our conclusions carry over to the other quantities in the following sections as well.

The comparison is shown in Fig. 2. As we discussed in Sect. 2, the strength and weaknesses of each bound depend on the combination of parameters α, d_A, and ε. Our analysis reveals that the almost additive approach performs best in the low d_A regime with small α, followed by a region where the mixed approach is superior, and then with increasing α, the operator-space approach outperforms the other two. As we increase d_A, the almost additive approach becomes progressively weaker compared to the mixed and operator-space approaches. Specifically, in the low α regime, the mixed approach dominates, while in the high α regime, the operator-space approach performs best. This improvement in performance with increasing d_A for the mixed and operator-space approaches is due to their scaling with $d_A^{2\frac{\alpha-1}{\alpha}}$, which is more favourable than the scaling with $d_A^{2(\alpha-1)}$ of the bound we got using the almost additive

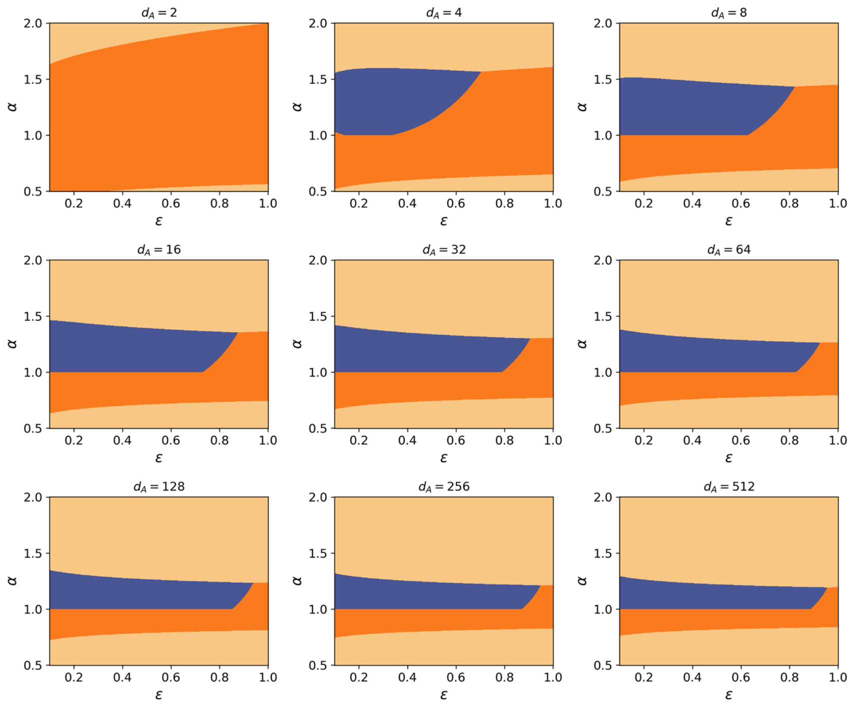


FIGURE 2. A comparison of the continuity bounds for $\tilde{H}^{\uparrow}_{\alpha}(A|B)_\rho$ proven by the almost-additive, operator space, and mixed approach. The value of d_A is according to the title of each plot. The visible colour indicates where the respective bound outperforms (is tighter than) the others

approach. Notably, the superior scaling of the mixed and operator-space approaches allows us to take the limit $\alpha \to \infty$ and obtain a continuity bound. However, the almost additive approach fails to have this property and does not vanish for $\varepsilon \to 0$ in this limit.

Corollary 5.1. *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, with $\frac{1}{2}\|\rho - \sigma\| \leq \varepsilon$, then for $\alpha \in [1/2, 1)$*

$$|\widetilde{H}^{\uparrow}_{\alpha}(A|B)_\rho - \widetilde{H}^{\uparrow}_{\alpha}(A|B)_\sigma| \leq \log(1+\varepsilon) + \frac{1}{1-\alpha} \log\left(1 + \varepsilon^{\alpha} d_A^{2(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right), \tag{8}$$

which is the bound from [25] (see Eq. 1), and for $\alpha \in (1, \infty)$

$$|\widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\rho} - \widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\sigma}| \leq \min \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log\left(1+\varepsilon d_A^{2(\alpha-1)} - \frac{\varepsilon^{\alpha}}{(1+\varepsilon)^{\alpha-1}}\right), \\ \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon d_A^{2\frac{\alpha-1}{\alpha}}\right), \\ \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon d_A^{2\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right). \end{cases} \tag{9}$$

Proof. We have that $\mathcal{C} = \{d_A^{-1}\mathbb{1}_A \otimes \sigma_B \ : \ \sigma_B \in \mathcal{S}(\mathcal{H}_B)\}$ is clearly a convex and compact set, containing a positive definite state. Using this definition and the $(1-\alpha)$-homogeneity of the $\widetilde{Q}_{\alpha}$ in the second argument, we get

$$\widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\rho} = -\widetilde{D}_{\alpha,\mathcal{C}}(\rho) + \log d_A$$

and hence

$$|\widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\rho} - \widetilde{H}^{\uparrow}_{\alpha}(A|B)_{\sigma}| = |\widetilde{D}_{\alpha,\mathcal{C}}(\rho) - \widetilde{D}_{\alpha,\mathcal{C}}(\sigma)|$$

We further have that $\sup_{\rho\in\mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \leq 2\log d_A$ using Eq. (4) and hence can directly apply Theorem 4.4, Theorem 4.17, and Theorem 4.20 to obtain the bounds in the assertion.

We can now study the limits $\alpha \to 1$ and $\alpha \to \infty$ of the new bound in Corollary 5.1. We find that the former limit coincides with the bounds by Alicki, Fannes and Winter and the latter bound with the one found by Marwah & Dupuis in the appendix of [25, Theorem 2].

Corollary 5.2 (Limits). *Let* $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, *with* $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$. *Then, the limit* $\alpha \to 1$ *of Equation* (9) *yields*

$$|H(A|B)_{\rho} - H(A|B)_{\sigma}| \leq 2\varepsilon \log d_A + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right),$$

and of Equation (9)

$$|H(A|B)_{\rho} - H(A|B)_{\sigma}| \leq \min \begin{cases} 2\varepsilon \log d_A + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right), \\ \infty, \\ 2\varepsilon \log d_A + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right), \end{cases}$$

where $H_{\rho}(A|B)$ *is the usual quantum conditional entropy, which is the bound derived in [42]. The limit* $\alpha \to \infty$ *of Equation* (9) *yields*

$$|H^{\uparrow}_{\infty}(A|B)_{\rho} - H^{\uparrow}_{\infty}(A|B)_{\sigma}| \leq \min \begin{cases} \log(1+\varepsilon) + 2\log d_A, \\ \log(1+\varepsilon d_A^2), \\ \log((1+\varepsilon)(1+\varepsilon d_A^2) - \varepsilon^2). \end{cases}$$

Proof. The proof is a direct application of Lemma 4.7, Lemma 4.19, Lemma 4.21 in the context of Corollary 5.1.

5.2. Continuity Bounds for the Sandwiched Rényi Mutual Information

Now, we can bring our techniques to bear on the sandwiched Rényi mutual information. We, unfortunately, cannot employ the theorems established in the beginning as $\{\rho_A \otimes \rho_B \ : \ \rho_A \in \mathcal{S}(\mathcal{H}_A), \rho_B \in \mathcal{S}(\mathcal{H}_B)\}$ is not a convex set, however, will use techniques very similar to those presented already.

Corollary 5.3. *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$. Then, for $\alpha \in [1/2, 1)$ we find*

$$|\widetilde{I}^{\uparrow}_{\alpha}(A:B)_{\rho} - \widetilde{I}^{\uparrow}_{\alpha}(A:B)_{\sigma}| \leq 2\log\left(1+\varepsilon^{\frac{1}{\alpha}}\right) + \frac{1}{1-\alpha}\log\left(1+\varepsilon^{\alpha}m^{2(1-\alpha)} - \frac{\varepsilon^{\frac{1}{\alpha}}}{(1+\varepsilon^{\frac{1}{\alpha}})^{2(1-\alpha)}}\right), \tag{10}$$

and for $\alpha \in (1, \infty)$ we have

$$|\widetilde{I}^{\uparrow}_{\alpha}(A:B)_{\rho} - \widetilde{I}^{\uparrow}_{\alpha}(A:B)_{\sigma}| \\ \leq 2\log\left(1+\varepsilon^{\frac{1}{\alpha}}\right) + \frac{1}{\alpha-1}\log\left(1+\varepsilon^{\frac{1}{\alpha}}m^{2(\alpha-1)} - \frac{\varepsilon^{\alpha}}{(1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)}}\right), \tag{11}$$

where in both bounds $m = \min\{d_A, d_B\}$.

Proof. We will only demonstrate the proof for $\alpha \in (1, \infty)$ as the case $\alpha \in [1/2, 1)$ is almost completely analogous up to reversing the inequalities. We further will only cover the case where $\frac{1}{2}\|\rho - \sigma\|_1 = \varepsilon$, since the proposed bounds are monotone in ε. Let μ, ν be orthogonal quantum states such that $\rho + \varepsilon\mu = \sigma + \varepsilon\nu$. It is straightforward to see that deriving the claimed bound boils down to deriving upper bounds on

$$\frac{\inf_{\tau_A,\tau_B} \widetilde{Q}_{\alpha}(\rho\|\tau_A \otimes \tau_B)}{\inf_{\tau_A,\tau_B} \widetilde{Q}_{\alpha}(\sigma\|\tau_A \otimes \tau_B)} \quad \text{and} \quad \frac{\inf_{\tau_A,\tau_B} \widetilde{Q}_{\alpha}(\sigma\|\tau_A \otimes \tau_B)}{\inf_{\tau_A,\tau_B} \widetilde{Q}_{\alpha}(\rho\|\tau_A \otimes \tau_B)}.$$

Since the proof for both fractions is exactly the same, we will only demonstrate it for the first one here. Employing joint convexity of the $\widetilde{Q}_{\alpha}$ (c.f. Lemma 4.1), we get that

$$\inf_{\tau_A,\tau_B} \widetilde{Q}_{\alpha}(\sigma + \varepsilon\nu\|\tau_A \otimes \tau_B) \\ \leq (1+\varepsilon)^{\alpha-1}\left(\inf_{\tau_A}\left(\inf_{\tau_B}\widetilde{Q}_{\alpha}(\sigma\|\tau_A \otimes \tau_B) + \varepsilon\inf_{\tau_B}\widetilde{Q}_{\alpha}(\nu\|\tau_A \otimes \tau_B)\right)\right)$$

where we could split the first infimum by writing the optimization as an optimization over $\tau_{B,1}$ and $\tau_{B,2}$, replacing $\tau_B = \frac{1}{1+\varepsilon}\tau_{B,1} + \frac{\varepsilon}{1+\varepsilon}\tau_{B,2}$ and then apply joint convexity. The second infimum is divided in $\tau_A = \frac{1}{1+\varepsilon^{\frac{1}{\alpha}}}\tau_{A,1} + \frac{\varepsilon^{\frac{1}{\alpha}}}{1+\varepsilon^{\frac{1}{\alpha}}}\tau_{A,2}$ and bounded via the anti-monotonicity of $\widetilde{Q}_{\alpha}$ in the second argument (see [40, Lemma 4.10]) and $\tau_A \geq \frac{1}{1+\varepsilon^{\frac{1}{\alpha}}}\tau_{A,1}$, $\tau_A \geq \frac{\varepsilon^{\frac{1}{\alpha}}}{1+\varepsilon^{\frac{1}{\alpha}}}\tau_{A,2}$. This gives

$$(1+\varepsilon)^{\alpha-1}(\inf_{\tau_A}(\inf_{\tau_B}\widetilde{Q}_{\alpha}(\sigma\|\tau_A \otimes \tau_B) + \varepsilon\inf_{\tau_B}\widetilde{Q}_{\alpha}(\nu\|\tau_A \otimes \tau_B))) \\ \leq (1+\varepsilon)^{\alpha-1}(1+\varepsilon^{\frac{1}{\alpha}})^{\alpha-1}(\inf_{\tau_A,\tau_B}\widetilde{Q}_{\alpha}(\sigma\|\tau_A \otimes \tau_B)$$

$$+\varepsilon^{1+\frac{1-\alpha}{\alpha}} \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\nu\|\tau_A\otimes\tau_B))$$
$$\leq (1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)}(\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\sigma\|\tau_A\otimes\tau_B) + \varepsilon^{\frac{1}{\alpha}} \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\nu\|\tau_A\otimes\tau_B)).$$

Now, using again the superadditivity (c.f. Lemma 4.2), we get

$$\begin{aligned}\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\sigma+\varepsilon\nu\|\tau_A\otimes\tau_B) &= \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\rho+\varepsilon\mu\|\tau_A\otimes\tau_B)\\ &\geq \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\rho\|\tau_A\otimes\tau_B) + \varepsilon^\alpha \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\mu\|\tau_A\otimes\tau_B)\end{aligned}$$

It holds that for a quantum state ξ, $1 \leq \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\xi\|\tau_A\otimes\tau_B) \leq m^{2(\alpha-1)}$, where $m = \min\{d_A, d_B\}$. Indeed, assuming without loss of generality $m = d_A$, we can estimate $\widetilde{I}_\alpha(A:B)_\xi \leq -\widetilde{H}(A|B)_\xi + \log d_A$ and use the bound in Eq. (4). Combining both estimates with this bound, we get

$$\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\rho\|\tau_A\otimes\tau_B) \leq (1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)}(\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\sigma\|\tau_A\otimes\tau_B)+\varepsilon^{\frac{1}{\alpha}} m^{2(\alpha-1)}) - \varepsilon^\alpha.$$

Subsequently, we divide by $\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\sigma\|\tau_A\otimes\tau_B)$ and use again that for ξ a quantum state $1 \leq \inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\xi\|\tau_A\otimes\tau_B)$. In addition, we use that $(1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)}\varepsilon^{\frac{1}{\alpha}} m^{2(\alpha-1)} - \varepsilon^\alpha \geq \varepsilon^{1/\alpha} - \varepsilon^\alpha \geq 0$. This leads to

$$\frac{\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\rho\|\tau_A\otimes\tau_B)}{\inf_{\tau_A,\tau_B} \widetilde{Q}_\alpha(\sigma\|\tau_A\otimes\tau_B)} \leq (1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)} \left(1+\varepsilon^{\frac{1}{\alpha}} m^{2(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon^{\frac{1}{\alpha}})^{2(\alpha-1)}}\right)$$

Applying the logarithm, multiplying with $\frac{1}{\alpha-1}$ and then repeating the whole procedure for the other fraction gives the claimed result.

Corollary 5.4 (Limits). *Let $\rho,\sigma \in \mathcal{S}(\mathcal{H}_A\otimes\mathcal{H}_B)$ with $\frac{1}{2}\|\rho-\sigma\|_1 \leq \varepsilon$. Then, taking $\alpha \to 1$ in Eqs. (10) and (11) yields*

$$|I(A:B)_\rho - I(A:B)_\sigma| \leq 2\varepsilon\log m + 2(1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right),$$

where $I(A:B)_\rho$ is the usual mutual information. This is the bound from [34]. Taking the limit $\alpha\to\infty$ leads to

$$|I^\uparrow_\infty(A:B)_\rho - I^\uparrow_\infty(A:B)_\sigma| \leq \log 4m^2,$$

which is no longer a continuity bound.

Proof. The first limit can be obtained using l'Hospital's rule as in Lemma 4.21. The second limit can be obtained by isolating $m^{2(\alpha-1)}$ in the logarithm.

5.3. Continuity Bounds for the Sandwiched Rényi Conditional Mutual Information

Let us recall that given a tripartite space $\mathcal{H}_{ABC} = \mathcal{H}_A\otimes\mathcal{H}_B\otimes\mathcal{H}_C$ and $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_{ABC})$, the sandwiched Rényi conditional mutual information of ρ_{ABC} is given by

$$\widetilde{I}^\uparrow_\alpha(A:C|B)_\rho := \widetilde{H}^\uparrow_\alpha(A|B)_\rho - \widetilde{H}^\uparrow_\alpha(A|BC)_\rho.$$

We also define the max conditional mutual information

$$\widetilde{I}^{\uparrow}_{\infty}(A:C|B)_\rho := \widetilde{H}^{\uparrow}_{\infty}(A|B)_\rho - \widetilde{H}^{\uparrow}_{\infty}(A|BC)_\rho \,.$$

and the quantum conditional mutual information

$$\widetilde{I}(A:C|B)_\rho := \widetilde{H}(A|B)_\rho - \widetilde{H}(A|BC)_\rho \,. \tag{12}$$

As a consequence of the definition, we can derive continuity bounds for the sandwiched Rényi conditional mutual information in terms of the continuity bounds for sandwiched Rényi conditional entropies obtained in Sect. 5.1. This is the content of the next result.

Corollary 5.5 (Continuity bound for sandwiched Rényi conditional mutual information). *Let ρ, $\sigma \in \mathcal{S}(\mathcal{H}_{ABC})$ be quantum states, with $\frac{1}{2}\|\rho-\sigma\|_1 \leq \varepsilon$. Then, for $\alpha \in [1/2, 1)$ we have*

$$\begin{aligned} |\widetilde{I}^{\uparrow}_{\alpha}(A:C|B)_\rho - \widetilde{I}^{\uparrow}_{\alpha}(A:C|B)_\sigma| &\leq 2\log(1+\varepsilon) \\ &+ \frac{2}{1-\alpha}\log\left(1+\varepsilon^\alpha d_A^{2(1-\alpha)} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right)\,, \end{aligned} \tag{13}$$

and for $\alpha \in (1,\infty)$,

$$|\widetilde{I}^{\uparrow}_{\alpha}(A:C|B)_\rho - \widetilde{I}^{\uparrow}_{\alpha}(A:C|B)_\sigma| \leq \min \begin{cases} 2\log\left(1+\varepsilon) + \frac{2}{\alpha-1}\log(1+\varepsilon d_A^{2(\alpha-1)} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\right)\,, \\ \frac{2\alpha}{\alpha-1}\log(1+\varepsilon d_A^{2\frac{\alpha-1}{\alpha}})\,, \\ 2\log(1+\varepsilon) + \frac{2\alpha}{\alpha-1}\log(1+\varepsilon d_A^{2\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}})\,. \end{cases} \tag{14}$$

Proof. Note that

$$\begin{aligned} |\widetilde{I}^{\uparrow}_{\alpha}(A:C|B)_\rho - \widetilde{I}^{\uparrow}_{\alpha}(A:C|B)_\sigma| &\leq |\widetilde{H}^{\uparrow}_{\alpha}(A|B)_\rho - \widetilde{H}^{\uparrow}_{\alpha}(A|B)_\sigma| \\ &+ |\widetilde{H}^{\uparrow}_{\alpha}(A|BC)_\rho - \widetilde{H}^{\uparrow}_{\alpha}(A|BC)_\sigma|\,. \end{aligned}$$

Thus, a continuity bound for the sandwiched Rényi conditional mutual information follows as a continuity bound for the sandwiched Rényi conditional entropy, with a factor of 2. We conclude using the bounds from Corollary 5.1.

For the sake of consistency, we state in the following the limits $\alpha \to 1$ and $\alpha \to \infty$ which directly follows from Corollary 5.2 by the same argument as in the proof of the above corollary:

Corollary 5.6 (Limits). *Let ρ, $\sigma \in \mathcal{S}(\mathcal{H}_{ABC})$ be quantum states, with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$. Then, the limit $\alpha \to 1$ of Equation (13) yields*

$$|I(A:C|B)_\rho - I(A:C|B)_\sigma| \leq 4\varepsilon\log d_A + 2(1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right)\,,$$

and the limit of Equation (14)

$$|I(A:C|B)_\rho - I(A:C|B)_\sigma| \leq \min \begin{cases} 4\varepsilon\log d_A + 2(1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right)\,, \\ \infty\,, \\ 4\varepsilon\log d_A + 2(1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right)\,, \end{cases}$$

where $I(A:C|B)_\rho$ is the usual quantum conditional information (see Eq. (12)). The limit $\alpha \to \infty$ of equation (14) yields

$$|I_\infty^\uparrow(A:C|B)_\rho - I_\infty^\uparrow(A:C|B)_\sigma| \le \min \begin{cases} 2\log(1+\varepsilon) + 4\log d_A\,, \\ 2\log(1+\varepsilon d_A^2)\,, \\ 2\log((1+\varepsilon)(1+\varepsilon d_A^2) - \varepsilon^2)\,. \end{cases}$$

Proof. The proof is a direct application of Corollary 5.2.

5.4. Continuity Bounds in the First Argument for Sandwiched Rényi Divergences

Corollary 5.7 (Continuity bound in the first argument). *Let ρ, σ, $\tau \in \mathcal{S}(\mathcal{H})$ be quantum states, with $\ker\tau \subseteq \ker\rho \cap \ker\sigma$, $\frac{1}{2}\|\rho-\sigma\|_1 \le \varepsilon$ and $\alpha \in [1/2, 1)$. Then*

$$|\widetilde{D}_\alpha(\rho\|\tau) - \widetilde{D}_\alpha(\sigma\|\tau)| \le \log\left(1+\varepsilon) + \frac{1}{1-\alpha}\log(1+\varepsilon^\alpha \widetilde{m}_\tau^{\alpha-1} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right)\,, \tag{15}$$

where $\widetilde{m}_\tau$ is the smallest nonzero eigenvalue of τ. For $\alpha \in (1,\infty)$ we find

$$|\widetilde{D}_\alpha(\rho\|\tau) - \widetilde{D}_\alpha(\sigma\|\tau)| \le \min \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log\left(1+\varepsilon\widetilde{m}_\tau^{1-\alpha} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\right)\,, \\ \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon\widetilde{m}_\tau^{\frac{1-\alpha}{\alpha}}\right)\,, \\ \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon\widetilde{m}_\tau^{\frac{1-\alpha}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right)\,. \end{cases} \tag{16}$$

Proof. We first restrict $\mathcal{H}$ to the support of τ which gives us a Hilbert space $\widetilde{\mathcal{H}}$ on which τ is positive definite. Clearly ρ, $\sigma \in \mathcal{S}(\widetilde{\mathcal{H}})$ as $\ker\tau \subseteq \ker\rho \cap \ker\sigma$. We further have that $\mathcal{C} := \{\tau\}$ is a convex, compact set containing a positive definite state and

$$\sup_{\rho\in\mathcal{S}(\widetilde{\mathcal{H}})} \widetilde{D}_{\alpha,\mathcal{C}}(\rho) \le \log\frac{1}{\widetilde{m}_\tau}\,,$$

which follows from $\widetilde{D}_\alpha(\nu\|\tau) \le D_\infty(\nu\|\tau) \le -\log\widetilde{m}_\tau$ for $\nu \in \mathcal{S}(\widetilde{\mathcal{H}})$ (see Eq. (5)). Lastly, we find

$$\widetilde{D}_\alpha(\eta\|\tau) = \widetilde{D}_{\alpha,\mathcal{C}}(\eta) \qquad \eta \in \{\rho,\sigma\}\,,$$

allowing us to use Theorem 4.4, Theorem 4.17, Theorem 4.20. This yields both assertions.

We can again consider the limits $\alpha \to 1$ and $\alpha \to \infty$.

Corollary 5.8 (Limits). *Let $\rho, \sigma, \tau \in \mathcal{S}(\mathcal{H})$ be quantum states, with $\ker\tau \subseteq \ker\rho \cap \ker\sigma$, $\frac{1}{2}\|\rho-\sigma\|_1 \le \varepsilon$. For the limit $\alpha \to 1$ in Equation (16), we obtain*

$$|D(\rho\|\tau) - D(\sigma\|\tau)| \le \varepsilon\log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon)h\left(\frac{\varepsilon}{1+\varepsilon}\right)\,,$$

where $\widetilde{m}_\tau$ is the smallest nonzero eigenvalue of τ, and for Equation (16)

$$|D(\rho\|\tau) - D(\sigma\|\tau)| \leq \min \begin{cases} \varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) , \\ \infty , \\ \varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) , \end{cases}$$

which is exactly the continuity bound in [8]. For $\alpha \to \infty$ Equation (16) gives

$$|D_\infty(\rho\|\tau) - D_\infty(\sigma\|\tau)| \leq \min \begin{cases} \log(1+\varepsilon) + \log(\widetilde{m}_\tau^{-1}) , \\ \log(1+\varepsilon\widetilde{m}_\tau^{-1}) , \\ \log((1+\varepsilon)(1+\varepsilon\widetilde{m}_\tau^{-1}) - \varepsilon^2) . \end{cases}$$

Proof. The proof is a direct application of Lemma 4.7, Lemma 4.19, Lemma 4.21 in the context of Corollary 5.7.

5.5. Divergence Bounds for Sandwiched Rényi Divergences

The continuity bounds in the first argument directly give us divergence bounds of the sandwiched Rényi divergences.

Corollary 5.9. (Divergence bound) *Let ρ, $\tau \in \mathcal{S}(\mathcal{H})$ with $\ker\tau \subseteq \ker\rho$ and $\frac{1}{2}\|\rho-\tau\|_1 \leq \varepsilon$ and $\alpha \in [1/2, 1)$*

$$\widetilde{D}_\alpha(\rho\|\tau) \leq \log(1+\varepsilon) + \frac{1}{1-\alpha} \log\left(1 + \varepsilon^\alpha \widetilde{m}_\tau^{\alpha-1} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right) , \tag{17}$$

where $\widetilde{m}_\tau$ is the smallest nonzero eigenvalue of τ. For $\alpha \in (1,\infty)$ we find

$$\widetilde{D}_\alpha(\rho\|\tau) \leq \min \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1} \log\left(1 + \varepsilon \widetilde{m}_\tau^{1-\alpha} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\right) , \\ \frac{\alpha}{\alpha-1} \log(1 + \varepsilon \widetilde{m}_\tau^{\frac{1-\alpha}{\alpha}}) , \\ \log(1+\varepsilon) + \frac{\alpha}{\alpha-1} \log\left(1 + \varepsilon \widetilde{m}_\tau^{\frac{1-\alpha}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right) . \end{cases} \tag{18}$$

Proof. The proof is a direct application of Corollary 5.7, with $\sigma = \tau$.

Corollary 5.10 (Limits). *Let ρ, $\tau \in \mathcal{S}(\mathcal{H})$ with $\ker\tau \subseteq \ker\rho$ and $\frac{1}{2}\|\rho-\tau\|_1 \leq \varepsilon$. For the limit $\alpha \to 1$ in Equation (17), we obtain*

$$D(\rho\|\tau) \leq \varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) ,$$

where $\widetilde{m}_\tau$ is the smallest nonzero eigenvalue of τ, and for Equation (18)

$$D(\rho\|\tau) \leq \min \begin{cases} \varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) , \\ \infty , \\ \varepsilon \log(\widetilde{m}_\tau^{-1}) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) , \end{cases}$$

which is exactly the divergence bound in [8]. For $\alpha \to \infty$ in Equation (18) gives

$$D_\infty(\rho\|\tau) \leq \min \begin{cases} \log(1+\varepsilon) + \log(\widetilde{m}_\tau^{-1}) , \\ \log(1+\varepsilon\widetilde{m}_\tau^{-1}) , \\ \log((1+\varepsilon)(1+\varepsilon\widetilde{m}_\tau^{-1}) - \varepsilon^2) . \end{cases}$$

Proof The proof is a direct application of Lemma 4.7, Lemma 4.19, Lemma 4.21 in the context of Corollary 5.7.

5.6. Distance to Separable States

We consider the distance to the set of separable states in terms of the sandwiched Rényi divergence and prove that it is uniformly continuous in the spirit of [13,42]. Note that the $\alpha \to 1$ limit recovers the Corollary 8 of [42].

Corollary 5.11 (Distance to the set of separable states). *Let* $\mathrm{SEP}_{AB} \subseteq \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ *be the set of separable states. Then the map*

$$\widetilde{D}_{\alpha,\mathrm{SEP}} : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathbb{R}, \qquad \rho \mapsto \widetilde{D}_{\alpha,\mathrm{SEP}}(\rho) := \inf_{\tau \in \mathrm{SEP}_{AB}} \widetilde{D}_\alpha(\rho\|\tau)$$

is uniformly continuous for $\alpha \in [1/2,1) \cup (1,\infty)$ *and for* ρ, $\sigma \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ *with* $\frac{1}{2}\|\rho - \sigma\|_1 \le \varepsilon$, *and* $\alpha \in [1/2,1)$

$$|\widetilde{D}_{\alpha,\mathrm{SEP}}(\rho) - \widetilde{D}_{\alpha,\mathrm{SEP}}(\sigma)| \le \log(1+\varepsilon) + \frac{1}{1-\alpha}\log(1 + \varepsilon^\alpha m^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}})$$

and for $\alpha \in (1,\infty)$

$$|\widetilde{D}_{\alpha,\mathrm{SEP}}(\rho) - \widetilde{D}_{\alpha,\mathrm{SEP}}(\sigma)| \le \min \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log(1+\varepsilon m^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}})\,, \\ \frac{\alpha}{\alpha-1}\log(1+\varepsilon m^{\frac{\alpha-1}{\alpha}})\,, \\ \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log(1+\varepsilon m^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}})\,, \end{cases}$$

where $m = \min\{d_A, d_B\}$.

Proof The set of separable states is known to be compact and convex and contains the maximally mixed state, which is positive definite. Consider $\alpha > 1$. Due to the joint convexity of $\widetilde{Q}_\alpha$ we can reduce to pure states and further using Remark 3.2 shows

$$\begin{aligned} \inf_{\tau \in \mathrm{SEP}_{AB}} \widetilde{Q}_\alpha(\rho\|\tau) &\le \sup_{|\psi\rangle} \inf_{\tau \in \mathrm{SEP}_{AB}} \widetilde{Q}_\alpha(|\psi\rangle\langle\psi|\|\tau) \\ &= \sup_{|\psi\rangle} \inf_{\tau \in \mathrm{SEP}_{AB}} (\langle\psi|\tau^{\frac{1-\alpha}{\alpha}}|\psi\rangle)^\alpha, \end{aligned}$$

where $|\psi\rangle\langle\psi|$ is a rank-1 projection. Let

$$|\psi\rangle = \sum_{i=1}^m \lambda_i |e_i\rangle_A |f_i\rangle_B$$

be the Schmidt decomposition of $|\psi\rangle$ such that in particular $\lambda_i \ge 0$ for all $i \in \{1,\dots,m\}$ and $\sum_i \lambda_i^2 = 1$. Both $\{|e_i\rangle\}_i$ and $\{|f_j\rangle\}_j$ are orthonormal sets. Choose

$$\tau_0 := \frac{1}{m}\sum_{i=1}^m |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|.$$

Then,

$$\inf_{\tau \in \mathrm{SEP}_{AB}} \widetilde{Q}_\alpha(\rho\|\tau) \le \sup_{|\psi\rangle}(\langle\psi|\tau_0^{\frac{1-\alpha}{\alpha}}|\psi\rangle)^\alpha$$

$$
\begin{aligned}
&= \sup_{|\psi\rangle} m^{\alpha-1} \left(\sum_{i=1}^{m} \lambda_i^2 \right)^{\alpha} \\
&= m^{\alpha-1} \, .
\end{aligned}
$$

The proof for $\alpha \in [1/2, 1)$ proceeds analogously, using the same τ_0. Employing Theorem 4.4, Theorem 4.17 and Theorem 4.20 gives the claimed bounds.

We can conclude the limits $\alpha \to 1$ and $\alpha \to \infty$ as before using Lemma 4.7, Lemma 4.19 and Lemma 4.21.

Corollary 5.12 *(Limits). Let ρ, $\sigma \in \mathcal{S}(\mathcal{H})$ and $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$. For the limit $\alpha \to 1$ in Equation (17), we obtain*

$$
|D_{\mathrm{SEP}}(\rho) - D_{\mathrm{SEP}}(\sigma)| \leq \min \begin{cases} \varepsilon \log(m) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) , \\ \infty \, , \\ \varepsilon \log(m) + (1+\varepsilon) h\left(\frac{\varepsilon}{1+\varepsilon}\right) , \end{cases}
$$

which is exactly the bound in [42]. For $\alpha \to \infty$ we infer

$$
|D_{\infty,\mathrm{SEP}}(\rho) - D_{\infty,\mathrm{SEP}}(\sigma)| \leq \min \begin{cases} \log(1+\varepsilon) + \log(m^{-1}) \, , \\ \log(1+\varepsilon m) \, , \\ \log((1+\varepsilon)(1+\varepsilon m) - \varepsilon^2) \, . \end{cases}
$$

5.7. Distance Measures in Resource Theories

While the previous section already gave an application to entanglement theory of our tools developed in Sect. 4, we can apply our techniques to any resource theory that includes a set of free states $\mathcal{F}$ that is compact and convex and contains a positive definite state. Then $\widetilde{D}_{\alpha,\mathcal{F}}(\rho)$ quantifies the resourcefulness of the state ρ in terms of its distance to the set of free states (see, e.g., [43]). This resource measure is known as the *Rényi relative entropy of resource*. Especially, the requirements for $\mathcal{F}$ hold for the resource theories falling into the framework of [3], requiring the sets of free states to be compact and convex and further contain the maximally mixed state. These theories include the resource theories of

1. entanglement (where the free states are the separable states).
2. coherence (where the free states are the states diagonal in a fixed basis).
3. asymmetry (where the free states are those invariant under some group).
4. nonuniformity and purity (where the only free state is the maximally mixed state).
5. thermodynamics (where the only free state is the Gibbs state for a fixed Hamiltonian and temperature; for this to be the maximally mixed state, we can take the temperature to be infinite).
6. contextuality (where the set of free states are non-contextual probability distributions).

7. stabilizer computation (where the free states are the convex hull of states that can be produced with a Clifford unitary from a standard state such as $|0\rangle$).

For more details on the resource theories in question, we refer the reader to [3] and the references therein as well as to the textbook [17]. We now present the continuity bounds we obtain on resource measures:

Corollary 5.13 *(Distance to the set of free states). Let $\mathcal{F} \subset \mathcal{S}(\mathcal{H})$ be the free states of a resource theory such that $\mathbb{1}/d \in \mathcal{F}$. Then the map*

$$\widetilde{D}_{\alpha,\mathcal{F}} : \mathcal{S}(\mathcal{H}) \to \mathbb{R}, \qquad \rho \mapsto \widetilde{D}_{\alpha,\mathcal{F}}(\rho) := \inf_{\tau\in\mathcal{F}} \widetilde{D}_\alpha(\rho\|\tau)$$

is uniformly continuous for $\alpha \in [1/2, 1) \cup (1, \infty)$ and for ρ, $\sigma \in \mathcal{S}(\mathcal{H})$ with $\frac{1}{2}\|\rho - \sigma\|_1 \le \varepsilon$, and $\alpha \in [1/2, 1)$

$$|\widetilde{D}_{\alpha,\mathcal{F}}(\rho) - \widetilde{D}_{\alpha,\mathcal{F}}(\sigma)| \le \log(1+\varepsilon) + \frac{1}{1-\alpha}\log(1 + \varepsilon^\alpha d^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}})$$

and for $\alpha \in (1, \infty)$

$$|\widetilde{D}_{\alpha,\mathcal{F}}(\rho) - \widetilde{D}_{\alpha,\mathcal{F}}(\sigma)| \le \min \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log(1 + \varepsilon d^{\alpha-1} - \frac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}})\,, \\ \frac{\alpha}{\alpha-1}\log(1 + \varepsilon d^{\frac{\alpha-1}{\alpha}})\,, \\ \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log(1 + \varepsilon d^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}})\,. \end{cases}$$

Proof We observe that, using $\mathbb{1}/d \in \mathcal{F}$,

$$\sup_{\rho\in\mathcal{S}(\mathcal{H})} \widetilde{D}_{\alpha,\mathcal{F}}(\rho) \le \log(d) + \sup_{\rho\in\mathcal{S}(\mathcal{H})} \frac{1}{\alpha-1}\log \operatorname{tr}[\rho^\alpha] \le \log(d).$$

Then, the claimed bounds follow from Theorem 4.4, Theorem 4.17, and Theorem 4.20.

Remark 5.14 The bounds presented here use no special knowledge of $\mathcal{F}$ other than that it contains the maximally mixed state. Note that this state could be replaced with another full-rank state of choice. As seen in Sect. 5.6 better bounds are in some cases possible for a given resource theory using other states in $\mathcal{F}$ to estimate the κ in Theorem 4.4, Theorem 4.17, and Theorem 4.20. Similar calculations can also provide continuity bounds on the resource measure in the resource theory of thermodynamics at finite temperature. We leave further explorations of concrete resource theories to future work.

5.8. Generalized Sandwiched Rényi Mutual Information

In this section, we are interested in continuity bounds for the *generalized sandwiched Rényi mutual information* defined as

$$\widetilde{I}^{\uparrow}_\alpha(\rho_{AB}\|\tau_A) := \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \widetilde{D}_\alpha(\rho_{AB}\|\tau_A \otimes \sigma_B) \tag{19}$$

for $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ and $\tau_A \in \mathcal{S}(\mathcal{H}_A)$ such that $\ker(\tau_A) \subseteq \ker \rho_A$.

This quantity appears in the context of hypothesis testing [18]. More concretely, in appears in the strong converse exponent of the hypothesis test where the null hypothesis is that the state is $\rho_{AB}^{\otimes n}$ and the alternative hypothesis that the state is $\tau_A^{\otimes n} \otimes \sigma_{B^n}$ for some state σ_{B^n}, not necessarily product. In this case, we impose that the error of the second kind goes to zero exponentially fast with a rate exceeding the mutual information $I(\rho_{AB}\|\tau_A)$. Here, $I(\rho_{AB}\|\tau_A)$ is defined as in Eq. (19), but with the Umegaki relative entropy. Then, the authors of [18] prove that the error of the first kind converges to 1 exponentially fast and the rate is determined by $\widetilde{I}_\alpha^{\uparrow}(\rho_{AB}\|\tau_A)$ with $\alpha > 1$. Recently, the generalized sandwiched Rényi mutual information has also appeared in the context of convex splitting [11, 12].

For this quantity, we can prove the following continuity bounds:

Corollary 5.15 *Let ρ_{AB}, $\sigma_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ and $\tau_A \in \mathcal{S}(\mathcal{H}_A)$ such that $\ker \tau_A \subseteq \ker \rho_A \cap \ker \sigma_A$. Moreover, let $\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\| \leq \epsilon$ and let $\widetilde{m}_\tau$ be the minimal nonzero eigenvalue of τ_A. Then the function $\eta_{AB} \mapsto \widetilde{I}_\alpha^{\uparrow}(\eta_{AB}\|\tau_A)$ is uniformly continuous on $\mathcal{S}(\mathcal{H}_{AB})$ with the following continuity bounds: for $\alpha \in [1/2, 1)$, we find*

$$|\widetilde{I}_\alpha^{\uparrow}(\rho_{AB}\|\tau_A) - \widetilde{I}_\alpha^{\uparrow}(\sigma_{AB}\|\tau_A)| \leq \log(1+\varepsilon) + \frac{1}{1-\alpha}\log\left(1 + \varepsilon^\alpha \left(\frac{m}{\widetilde{m}_\tau}\right)^{1-\alpha} - \frac{\varepsilon}{(1+\varepsilon)^{1-\alpha}}\right)$$

and for $\alpha \in (1, \infty)$

$$|\widetilde{I}_\alpha^{\uparrow}(\rho_{AB}\|\tau_A) - \widetilde{I}_\alpha^{\uparrow}(\sigma_{AB}\|\tau_A)| \leq \min \begin{cases} \log(1+\varepsilon) + \frac{1}{\alpha-1}\log\left(1 + \varepsilon\left(\frac{m}{\widetilde{m}_\tau}\right)^{\alpha-1} - \frac{\varepsilon}{(1+\varepsilon)^{\alpha-1}}\right), \\ \frac{\alpha}{\alpha-1}\log(1+\varepsilon\left(\frac{m}{\widetilde{m}_\tau}\right)^{\frac{\alpha-1}{\alpha}}), \\ \log(1+\varepsilon) + \frac{\alpha}{\alpha-1}\log\left(1 + \varepsilon\left(\frac{m}{\widetilde{m}_\tau}\right)^{\frac{\alpha-1}{\alpha}} - \frac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right), \end{cases}$$

where $m = \min\{d_A, d_B\}$.

Proof We first restrict the Hilbert space $\mathcal{H}_{AB}$ to the support of $\tau_A \otimes \mathbb{1}_B$ and get that this is a Hilbert space $\widetilde{\mathcal{H}}_{AB}$ on which $\tau_A \otimes \mathbb{1}$ is positive definite. Clearly $\rho_{AB}, \sigma_{AB} \in \mathcal{S}(\widetilde{\mathcal{H}}_{AB})$ as $\ker \tau_A \subseteq \ker \rho_A \cap \ker \sigma_A$, a fact that can be verified using a purification of ρ_{AB} and an appropriate Schmidt decomposition that for P_A the orthogonal projection onto the support of τ_A, it holds that $(P_A \otimes \mathbb{1})\rho_{AB}(P_A \otimes \mathbb{1}) = \rho_{AB}$. Hence

$$\mathcal{C}(\tau_A) := \{\tau_A \otimes \sigma_B : \sigma_B \in \mathcal{S}(\mathcal{H}_B)\}$$

is a compact convex subset of the state space $\mathcal{S}(\mathcal{H}_{AB})$ which contains a positive definite state (e.g. $\tau_A \otimes \frac{\mathbb{1}_B}{d_B}$). Finally, we have that

$$0 \leq \widetilde{I}_\alpha^{\uparrow}(\rho_{AB}\|\tau_A) \leq \inf_{\sigma_B} \widetilde{D}_\alpha(\rho_{AB}\|\widetilde{m}_\tau \mathbb{1}_A \otimes \sigma_B) = -\log(\widetilde{m}_\tau) + \inf_{\sigma_B} \widetilde{D}_\alpha(\rho_{AB}\|\mathbb{1}_A \otimes \sigma_B)$$

where we have used Lemma 4.3 of [40] in the inequality. Thus

$$\sup_{\rho_{AB}} |\widetilde{I}_\alpha^{\uparrow}(\rho_{AB}\|\tau_A)| \leq -\log(\widetilde{m}_\tau) - \widetilde{H}_\alpha^{\uparrow}(A|B)_\rho \leq \log \frac{m}{\widetilde{m}_\tau}.$$

The assertion now follows from applying Theorem 4.4, Theorem 4.17 and 4.20.

For $\tau_A = \mathbb{1}/d_A$, we retrieve the bounds on the sandwiched Rényi conditional entropy in Sect. 5.1 as expected. Therefore, they also have an interpretation in a hypothesis-testing scenario.

Corollary 5.16 *(Limits). Let ρ_{AB}, $\sigma_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ and $\tau \in \mathcal{S}(\mathcal{H}_A)$ such that $\ker \tau_A \subseteq \ker \rho_A \cap \ker \sigma_A$. Moreover, let $\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\| \leq \epsilon$ and $\widetilde{m}_\tau$ be the minimal nonzero eigenvalue of τ_A. Then we find*

$$|I(\rho_{AB}\|\tau_A) - I(\sigma_{AB}\|\tau_A)| \leq \varepsilon \log\left(\frac{m}{\widetilde{m}_\tau}\right) + (1+\varepsilon)h\Big(\frac{\varepsilon}{1+\varepsilon}\Big),$$

$$|\widetilde{I}^{\uparrow}_{\infty}(\rho_{AB}\|\tau_A) - \widetilde{I}^{\uparrow}_{\infty}(\sigma_{AB}\|\tau_A)| \leq \log\left(1 + \varepsilon\frac{m}{\widetilde{m}_\tau}\right).$$

Proof This follows from Lemmas 4.7 and 4.19.

6. α-Approximate Quantum Markov Chains

The main aim of this section is to study α-approximate quantum Markov chains, namely positive states ρ_{ABC} on a tripartite space whose (non-variational) sandwiched Rényi conditional mutual information is small enough. Here, we show that this notion is equivalent to that of approximate quantum Markov chains, i.e. states for which the conditional mutual information is small enough. Beforehand, we need to introduce some technical results, which concern continuity bounds for sandwiched Rényi divergences in both inputs.

6.1. Continuity Bounds for Non-Variational Rényi Divergences via the ALAFF method

In contrast to our main results, we prove in this section some continuity bounds for the sandwiched Rényi divergence and its derived quantities with respect to both inputs. The drawback of this more general approach is that the bounds obtained are less tight compared to those previously proven by techniques tailored to continuity bounds where the second input is fixed or optimized over.

In comparison to the quantities studied before, we no longer optimize over the second state but consider it to be the marginal of the input. Let us consider a bipartite Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$. Then, for $\alpha \in [1/2, 1) \cup (1, \infty)$, the *(non-variational) sandwiched Rényi conditional entropy* is given by

$$\widetilde{H}_\alpha(A|B)_\rho := \frac{1}{1-\alpha} \log \widetilde{Q}_\alpha(\rho_{AB}\| \mathbb{1}_A \otimes \rho_B).$$

Note that we will also use the notation $\widetilde{Q}_\alpha(A|B)_\rho := \widetilde{Q}_\alpha(\rho_{AB}\| \mathbb{1} \otimes \rho_B)$. Analogously, for the *(non-variational) sandwiched Rényi mutual information* we set

$$\widetilde{I}_\alpha(A : B)_\rho := \widetilde{D}_\alpha(\rho_{AB}\|\rho_A \otimes \rho_B),$$

and lastly, we define the *(non-variational) sandwiched Rényi conditional mutual information* for $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ by

$$\widetilde{I}_\alpha(A:C|B)_\rho := \widetilde{H}_\alpha(C|B)_\rho - \widetilde{H}_\alpha(C|AB)_\rho \,.$$

The approach to derive continuity bounds will be based on the *ALAFF method* [8,9]. This is a generalization of the *Alicki-Fannes-Winter (AFW) method*, introduced under this name by Shirokov [35,36] and based on the seminal results for continuity bounds of entropies by the authors of [1,42].

Let $\mathcal{H}$ denote a finite-dimensional Hilbert space and f be a real-valued function on the convex set $\mathcal{S}_0 \subseteq \mathcal{S}(\mathcal{H})$. We say that f is *almost locally affine (ALAFF)*, if there exist continuous functions $a_f, b_f : [0,1] \to \mathbb{R}$, that are non-decreasing on $[0, \frac{1}{2}]$, vanish as $p \to 0^+$ and satisfy

$$-a_f(p) \leq f(p\rho + (1-p)\sigma) - pf(\rho) - (1-p)f(\sigma) \leq b_f(p) \tag{20}$$

for all $p \in [0,1]$ and $\rho, \sigma \in \mathcal{S}_0$. Moreover, the set $\mathcal{S}_0$ is called *perturbed* Δ-*invariant* with perturbation parameter $s \in [0,1)$ if for all $\rho, \sigma \in \mathcal{S}_0$ with $\rho \neq \sigma$, there is a state τ such that both states

$$\Delta^{\pm}(\rho,\sigma,\tau) = s\tau + (1-s)\varepsilon^{-1}[\rho-\sigma]_\pm \in \mathcal{S}_0 \,,$$

where we fix $\varepsilon = \frac{1}{2}\|\rho - \sigma\|_1$, and denote by $[\,\cdot\,]_\pm$ the positive and negative parts of a self-adjoint operator, respectively.

In the following result, we prove that $\widetilde{Q}_\alpha(\cdot\|\cdot)$ is almost locally affine for $\alpha \in [1/2,1) \cup (1,+\infty)$. Note that we only need to prove the almost joint concavity part for $\alpha \in (1,+\infty)$ (respectively, convexity, for $\alpha \in [1/2,1)$), since $\widetilde{Q}_\alpha(\cdot\|\cdot)$ is already jointly convex (resp. jointly concave).

Theorem 6.1 *Let $(\rho_1,\sigma_1),(\rho_2,\sigma_2) \in \mathcal{S}_{\ker} := \{(\rho,\sigma) \in \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \,:\, \ker\sigma \subseteq \ker\rho\}$, $p \in [0,1]$, and define $\rho := p\rho_1 + (1-p)\rho_2$ and $\sigma := p\sigma_1 + (1-p)\sigma_2$, respectively. Let us denote by m_{σ_1} and m_{σ_2} the minimal nonzero eigenvalue of σ_1 and σ_2, respectively. Then, for $\alpha \in [1/2,1)$, we have*

$$\widetilde{Q}_\alpha(\rho\|\sigma) \leq p\,\widetilde{Q}_\alpha(\rho_1\|\sigma_1) + (1-p)\widetilde{Q}_\alpha(\rho_2\|\sigma_2) + \xi(\alpha,p,\sigma_1,\sigma_2)\,, \tag{21}$$

and for $\alpha \in (1,+\infty)$,

$$\widetilde{Q}_\alpha(\rho\|\sigma) \geq p\widetilde{Q}_\alpha(\rho_1\|\sigma_1) + (1-p)\widetilde{Q}_\alpha(\rho_2\|\sigma_2) + \xi(\alpha,p,\sigma_1,\sigma_2)\,, \tag{22}$$

where for $\alpha \in [1/2,1)$ the error term ξ is positive and if $\alpha \in (1,\infty)$ it is negative. Moreover,

$$\begin{aligned}\xi(\alpha,p,\sigma_1,\sigma_2) &:= -1 + p^\alpha(p + (1-p)m_{\sigma_1}^{-1})^{1-\alpha} + (1-p)^\alpha(pm_{\sigma_2}^{-1} + (1-p))^{1-\alpha}\\ &\leq (1-\alpha)\sqrt{p}\left(\left(\log(m_{\sigma_1}^{-1})+1\right)(m_{\sigma_1}^{-1})^{1-\alpha} + \left(m_{\sigma_2}^{-1}+1\right)(m_{\sigma_2}^{-1})^{1-\alpha}\right) =: u_\alpha(p)\end{aligned}$$

for $\alpha \in [1/2,1)$ and

$$\begin{aligned}\xi(\alpha,p,\sigma_1,\sigma_2) &:= -(p - p^\alpha(p+(1-p)m_{\sigma_1}^{-1})^{1-\alpha})m_{\sigma_1}^{1-\alpha}\\ &\quad - ((1-p) - (1-p)^\alpha(pm_{\sigma_2}^{-1} + (1-p))^{1-\alpha})m_{\sigma_2}^{1-\alpha}\\ &\geq (1-\alpha)\sqrt{p}\left(\left(\log(m_{\sigma_1}^{-1})+1\right)m_{\sigma_1}^{1-\alpha} + \left(m_{\sigma_2}^{-1}+1\right)m_{\sigma_2}^{1-\alpha}\right) =: v_\alpha(p)\end{aligned}$$

for $\alpha \in (1,\infty)$.

Proof We start with the case $\alpha \in [1/2, 1)$. Here, we first separate ρ using the superadditivity of $\widetilde{Q}_\alpha$ (cf. Lemma 4.2)

$$\widetilde{Q}_\alpha(\rho\|\sigma) = \mathrm{tr}[(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}})^\alpha] \\ \leq p^\alpha \mathrm{tr}[(\sigma^{\frac{1-\alpha}{2\alpha}}\rho_1\sigma^{\frac{1-\alpha}{2\alpha}})^\alpha] + (1-p)^\alpha \mathrm{tr}[(\sigma^{\frac{1-\alpha}{2\alpha}}\rho_2\sigma^{\frac{1-\alpha}{2\alpha}})^\alpha],$$

To split σ, let P_1 be the projection onto the support of σ_1. Then, we can upper bound $P_1\sigma P_1$ by σ_1 using $P_1\sigma_2 P_1 \leq P_1 \leq m_{\sigma_1}^{-1}\sigma_1$ to obtain

$$P_1\sigma P_1 = p\sigma_1 + (1-p)P_1\sigma_2 P_1 \leq \big(p + (1-p)m_{\sigma_1}^{-1}\big)\sigma_1 =: c_{\sigma_1}\sigma_1\,.$$

Then, we rewrite and upper bound by

$$\mathrm{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho_1\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right] = \mathrm{tr}\left[\left(\rho_1^{1/2}P_1\sigma^{\frac{1-\alpha}{\alpha}}P_1\rho_1^{1/2}\right)^\alpha\right] \\ \leq \mathrm{tr}\left[\left(\rho_1^{1/2}(P_1\sigma P_1)^{\frac{1-\alpha}{\alpha}}\rho_1^{1/2}\right)^\alpha\right] \\ \leq c_{\sigma_1}^{1-\alpha}\mathrm{tr}\left[\left(\rho_1^{1/2}\sigma_1^{\frac{1-\alpha}{\alpha}}\rho_1^{1/2}\right)^\alpha\right],$$

which uses that $x \mapsto x^s$ is operator monotone and operator concave for $s \in [0,1]$ and, in particular, $\frac{1-\alpha}{\alpha} \in (0,1]$ for all $\alpha \in [\frac{1}{2}, 1)$. Therefore, the first inequality follows from [7, Theorem V.2.3] Repeating the same steps for σ_2, this time inserting the projection P_2 onto the support of σ_2, and using the inequalities $\widetilde{Q}_\alpha(\rho_1\|\sigma_1) \leq 1$ and $\widetilde{Q}_\alpha(\rho_2\|\sigma_2) \leq 1$, we obtain

$$\widetilde{Q}_\alpha(\rho\|\sigma) \leq p^\alpha c_{\sigma_1}^{1-\alpha}\widetilde{Q}_\alpha(\rho_1\|\sigma_1) + (1-p)^\alpha c_{\sigma_2}^{1-\alpha}\widetilde{Q}_\alpha(\rho_2\|\sigma_2) \\ \leq p\widetilde{Q}_\alpha(\rho_1\|\sigma_1) + (1-p)\widetilde{Q}_\alpha(\rho_2\|\sigma_2) + \xi(\alpha, p, \sigma_1, \sigma_2)\,,$$

for

$$\xi(\alpha, p, \sigma_1, \sigma_2) = (p^\alpha c_{\sigma_1}^{1-\alpha} - p) + ((1-p)^\alpha c_{\sigma_2}^{1-\alpha} - (1-p)) \\ \geq (p^\alpha c_{\sigma_1}^{1-\alpha} - p)\widetilde{Q}_\alpha(\rho_1\|\sigma_1) + ((1-p)^\alpha c_{\sigma_2}^{1-\alpha} - (1-p))\widetilde{Q}_\alpha(\rho_2\|\sigma_2)\,,$$

which proves the first bound. For $p \in \{0,1\}$ the bound is clear so that w.l.o.g. we consider $p \in (0,1)$ in the following. Next, we upper bound $\xi(\alpha, p, \sigma_1, \sigma_2)$ further:

$$\xi(\alpha, p, \sigma_1, \sigma_2) = -p + p^\alpha(p + (1-p)m_{\sigma_1}^{-1})^{1-\alpha} \\ - (1-p) + (1-p)^\alpha(pm_{\sigma_2}^{-1} + (1-p))^{1-\alpha} \\ = \int_0^1 p\frac{d}{ds}(1 + p^{-1}(1-p)m_{\sigma_1}^{-1})^{s(1-\alpha)} \\ + (1-p)\frac{d}{ds}((1-p)^{-1}pm_{\sigma_2}^{-1} + 1)^{s(1-\alpha)}ds \\ = (1-\alpha)\int_0^1 p^{1-s(1-\alpha)}\left(\log(p + (1-p)m_{\sigma_1}^{-1}) - \log(p)\right) \\ (p + (1-p)m_{\sigma_1}^{-1})^{s(1-\alpha)} \\ + (1-p)^{1-s(1-\alpha)}\left(\log(pm_{\sigma_2}^{-1} + (1-p)) - \log(1-p)\right) \\ (pm_{\sigma_2}^{-1} + (1-p))^{s(1-\alpha)}ds\,,$$

where the natural logarithm is considered. Next, we bound the following term separately

$$\begin{aligned}-\int_0^1 p^{1-s(1-\alpha)}\log(p)(p+(1-p)m_{\sigma_1}^{-1})^{s(1-\alpha)} \\ \leq -(p+(1-p)m_{\sigma_1}^{-1})^{1-\alpha} \\ \int_0^1 p^{1-s/2}\log(p)ds \\ \leq 2(m_{\sigma_1}^{-1})^{1-\alpha}p(\sqrt{p}-p) \\ \leq (m_{\sigma_1}^{-1})^{1-\alpha}\sqrt{p}\end{aligned}$$

where we used $p(\sqrt{p}-p)\leq\frac{1}{2}\sqrt{p}$. Similarly,

$$\begin{aligned}-\int_0^1 (1-p)^{1-s(1-\alpha)}\log(1-p)(pm_{\sigma_2}^{-1}+(1-p))^{s(1-\alpha)} \\ \leq -(pm_{\sigma_2}^{-1}+(1-p))^{1-\alpha}\int_0^1(1-p)^{1-s/2}\log(1-p)ds \\ \leq 2(m_{\sigma_1}^{-1})^{1-\alpha}(1-p)(\sqrt{1-p}-(1-p)) \\ \leq (m_{\sigma_1}^{-1})^{1-\alpha}\sqrt{p}\end{aligned}$$

which uses that $\sqrt{1-p}-(1-p)\leq\frac{1}{2}\sqrt{p}$ in the last inequality. Therefore,

$$\begin{aligned}&\xi(\alpha,p,\sigma_1,\sigma_2)\\ &=(1-\alpha)\int_0^1 p^{1-s(1-\alpha)}\left(\log(p+(1-p)m_{\sigma_1}^{-1})-\log(p)\right)(p+(1-p)m_{\sigma_1}^{-1})^{s(1-\alpha)}\\ &\quad+(1-p)^{1-s(1-\alpha)}\left(\log(pm_{\sigma_2}^{-1}+(1-p))-\log(1-p)\right)(pm_{\sigma_2}^{-1}+(1-p))^{s(1-\alpha)}ds\\ &\leq(1-\alpha)\left(\sqrt{p}\left(\log(m_{\sigma_1}^{-1})+1\right)(m_{\sigma_1}^{-1})^{1-\alpha}+\left(\log(pm_{\sigma_2}^{-1}+(1-p))+\sqrt{p}\right)(m_{\sigma_2}^{-1})^{1-\alpha}\right)\\ &\leq(1-\alpha)\left(\sqrt{p}\left(\log(m_{\sigma_1}^{-1})+1\right)(m_{\sigma_1}^{-1})^{1-\alpha}+\left(\log(pm_{\sigma_2}^{-1}+1)+\sqrt{p}\right)(m_{\sigma_2}^{-1})^{1-\alpha}\right)\\ &\leq(1-\alpha)\sqrt{p}\left(\left(\log(m_{\sigma_1}^{-1})+1\right)(m_{\sigma_1}^{-1})^{1-\alpha}+\left(m_{\sigma_2}^{-1}+1\right)(m_{\sigma_2}^{-1})^{1-\alpha}\right)\end{aligned}$$

finishes the bounds for the case $\alpha\in[1/2,1)$. Next, we consider the case of $\alpha\in(1,+\infty)$, which follows a similar line of reasoning. The superadditivity of $\widetilde{Q}_\alpha$ (cf. Lemma 4.2) gives

$$\widetilde{Q}_\alpha(\rho\|\sigma)\geq p^\alpha\mathrm{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho_1\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]+(1-p)^\alpha\mathrm{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho_2\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right].$$

Then, we use the bounds $P_1\sigma P_1\leq c_{\sigma_1}\sigma_1$ and $P_2\sigma P_2\leq c_{\sigma_2}\sigma_2$ again. Since $\alpha\in(1,+\infty)$, the fraction $\frac{1-\alpha}{\alpha}\in(-1,0)$ so that $x\mapsto -x^{\frac{1-\alpha}{\alpha}}$ is operator monotone and operator concave, allowing us to use [7, Exercise V.2.2] and find

$$\begin{aligned}\widetilde{Q}_\alpha(\rho\|\sigma)&\geq p^\alpha c_{\sigma_1}^{1-\alpha}\widetilde{Q}_\alpha(\rho_1\|\sigma_1)+(1-p)^\alpha c_{\sigma_2}^{1-\alpha}\widetilde{Q}_\alpha(\rho_2\|\sigma_2)\\ &=p\widetilde{Q}_\alpha(\rho_1\|\sigma_1)+(1-p)\widetilde{Q}_\alpha(\rho_2\|\sigma_2)+\xi(\alpha,p,\sigma_1,\sigma_2),\end{aligned}$$

Since $\widetilde{Q}_\alpha(\rho_i\|\sigma_i)\leq m_{\sigma_i}^{1-\alpha}$ for $i=1,2$, we find that

$$\xi(\alpha,p,\sigma_1,\sigma_2)=-[(p-p^\alpha c_{\sigma_1}^{1-\alpha})m_{\sigma_1}^{1-\alpha}+((1-p)-(1-p)^\alpha c_{\sigma_2}^{1-\alpha})m_{\sigma_2}^{1-\alpha}]$$

$$\leq -[(p - p^{\alpha} c_{\sigma_1}^{1-\alpha})\widetilde{Q}_{\alpha}(\rho_1\|\sigma_1) + ((1-p) - (1-p)^{\alpha} c_{\sigma_2}^{1-\alpha})\widetilde{Q}_{\alpha}(\rho_2\|\sigma_2)]\,,$$

giving the third bound in the assertion. At last, we simplify this bound as follows;

$$\begin{aligned}
&-\xi(\alpha, p, \sigma_1, \sigma_2)\\
&\quad= (p - p^{\alpha}(p + (1-p)m_{\sigma_1}^{-1})^{1-\alpha})m_{\sigma_1}^{1-\alpha} + ((1-p) - (1-p)^{\alpha}(pm_{\sigma_2}^{-1} + (1-p))^{1-\alpha})m_{\sigma_2}^{1-\alpha}\\
&\quad= (\alpha-1)\int_0^1 p^{1-s(1-\alpha)}\left(\log(p + (1-p)m_{\sigma_1}^{-1}) - \log(p)\right)(p + (1-p)m_{\sigma_1}^{-1})^{s(1-\alpha)}m_{\sigma_1}^{1-\alpha}\\
&\qquad\qquad + (1-p)^{1-s(1-\alpha)}\left(\log(pm_{\sigma_2}^{-1} + (1-p)) - \log(1-p)\right)(pm_{\sigma_2}^{-1} + (1-p))^{s(1-\alpha)}m_{\sigma_2}^{1-\alpha}ds\\
&\quad\overset{(i)}{\leq} (\alpha-1)\left(p\left(\log(m_{\sigma_1}^{-1}) - \log(p)\right)m_{\sigma_1}^{1-\alpha} + \left(\log(pm_{\sigma_2}^{-1} + (1-p)) + \sqrt{p}\right)m_{\sigma_2}^{1-\alpha}\right)\\
&\quad\overset{(ii)}{\leq} (\alpha-1)\left(\sqrt{p}\left(\log(m_{\sigma_1}^{-1}) + 1\right)m_{\sigma_1}^{1-\alpha} + \left(\log(pm_{\sigma_2}^{-1} + 1) + \sqrt{p}\right)m_{\sigma_2}^{1-\alpha}\right)\\
&\quad\leq (\alpha-1)\sqrt{p}\left(\left(\log(m_{\sigma_1}^{-1}) + 1\right)m_{\sigma_1}^{1-\alpha} + \left(m_{\sigma_2}^{-1} + 1\right)m_{\sigma_2}^{1-\alpha}\right)\,,
\end{aligned}$$

additionally, to the inequalities exploited before, we used in (i), the inequality $-(1-p)\log(1-p) \leq \sqrt{p}$ and in (ii) $-p\log p \leq \sqrt{p}$. This finishes the proof.

Remark 6.2 The function $\xi(\alpha, p, \sigma_1, \sigma_2)$ has the following behaviour:

- If $p = 0, 1$, we have $\xi(\alpha, p, \sigma_1, \sigma_2) = 0$. For the bounds $u_\alpha(p), v_\alpha(p)$ from Lemma 6.1 only $p = 0$ makes them vanish.
- If $\sigma_1 = \sigma_2$, we can replace $\xi(\alpha, p, \sigma_1, \sigma_1)$ by the simpler function $\xi(\alpha, p, \sigma_1) = \big(-1 + p^\alpha + (1-p)^\alpha\big)$, for $\alpha \in [1/2, 1)$, and $\xi(\alpha, p, \sigma_1) = -\big(1 - p^\alpha - (1-p)^\alpha\big)m_{\sigma_1}^{1-\alpha}$, for $\alpha \in (1, +\infty)$, as in these cases we can take c_{σ_1} and c_{σ_2} to be 1 in the proof.
- It can be seen from the statement of Lemma 6.1 that $p \mapsto u_\alpha(p)$ and $p \mapsto v_\alpha(p)$ are non-decreasing on $[0, 1/2]$, since the square root is.

We are now in position of using the findings of Lemma 6.1 for $\widetilde{Q}_\alpha(\cdot\|\cdot)$, jointly with a suitable Δ-invariant set, to apply the ALAFF method (cf. [8, Theorem 4.6]) and obtain continuity bounds. For the time being, we focus on $\widetilde{H}_\alpha(A|B)$ and $\widetilde{I}_\alpha(A:C|B)$, however, prove more general continuity bounds for $\widetilde{Q}_\alpha(\cdot\|\cdot)$, and thus $\widetilde{D}_\alpha(\cdot\|\cdot)$, in Sect. 6.3.

Corollary 6.3 *Let $d_{AB}^{-1} > m > 0$ and $\mathcal{S}_0 := \{\rho : \rho \in \mathcal{S}(\mathcal{H}_{AB}), m_\rho \geq m\}$ with m_ρ the minimal eigenvalue of ρ. Then for $\rho, \sigma \in \mathcal{S}_0$ with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$, we find for $\alpha \in [1/2, 1) \cup (1, \infty)$*

$$|\widetilde{H}_\alpha(A|B)_\rho - \widetilde{H}_\alpha(A|B)_\sigma| \leq c(\alpha, m, d_A, d_{AB})\sqrt{\varepsilon}\,.$$

with d_A, d_{AB} the dimensions of $\mathcal{H}_A$ and $\mathcal{H}_{AB}$, respectively, and

$$c(\alpha, m, d_A, d_{AB}) := \begin{cases} \frac{d_A^{2(1-\alpha)}-1}{1-\alpha}\frac{1}{1-md_{AB}} + \frac{\sqrt{2}\,\tilde{c}(\alpha,m)}{(1-md_{AB})}d_A^{2(1-\alpha)} & \alpha \in [1/2, 1)\\ 2\log d_A\frac{1}{1-md_{AB}} + \frac{\sqrt{2}\,\tilde{c}(1,m)}{(1-md_{AB})} & \alpha = 1\\ \frac{d_A^{2(\alpha-1)}-1}{\alpha-1}\frac{1}{1-md_{AB}} + \frac{\sqrt{2}\,\tilde{c}(\alpha,m)}{(1-md_{AB})} & \alpha \in (1, \infty) \end{cases}$$

where $\widetilde{c}(\alpha,m) = \begin{cases} (\log(m^{-1}) + m^{-1} + 2)m^{\alpha-1} & \alpha \in [1/2,1] \\ (\log(m^{-1}) + m^{-1} + 2)m^{1-\alpha} & \alpha \in (1,\infty) \end{cases}$.

Proof We will only demonstrate the proof for $\alpha > 1$ as the one for $\alpha < 1$ is completely analogous. First note that $\mathcal{S}_0$ is a convex, md_{AB}-perturbed Δ-invariant set[3] (see [8, Corollary 6.8] for comparison). We further have that $\rho \mapsto \widetilde{Q}_\alpha(A|B)_\rho$ as a map from $\mathcal{S}_0 \to [0,\infty)$ is convex and almost concave with remainder $v_\alpha(p)$ (cf. Lemma 6.1). Lastly, we get

$$\sup_{\mu,\nu\in\mathcal{S}(\mathcal{H})} |\widetilde{Q}_\alpha(A|B)_\mu - \widetilde{Q}_\alpha(A|B)_\nu| \leq d_A^{2(\alpha-1)} - 1\,,$$

and that $p \mapsto \frac{v_\alpha(p)}{1-p}$ is monotone on $[0,1)$ as clearly $\frac{\sqrt{p}}{1-p}$ is. Employing [8, Theorem 4.6] gives

$$\begin{aligned} |\widetilde{Q}_\alpha(A|B)_\rho - \widetilde{Q}_\alpha(A|B)_\sigma| &\leq (d_A^{2(\alpha-1)} - 1)\frac{\varepsilon}{1-md_{AB}} + (\alpha-1)\frac{\sqrt{1-md_{AB}+\varepsilon}}{(1-md_{AB})}\widetilde{c}(\alpha,m)\sqrt{\varepsilon} \\ &\leq (d_A^{2(\alpha-1)} - 1)\frac{\varepsilon}{1-md_{AB}} + (\alpha-1)\frac{\sqrt{2}\,\widetilde{c}(\alpha,m)}{(1-md_{AB})}\sqrt{\varepsilon} \\ &\leq (\alpha-1)c(\alpha,m,d_A,d_{AB})\sqrt{\varepsilon}\,. \end{aligned}$$

Note that the normalization of $\widetilde{Q}_\alpha(A|B)_\rho$ is cancelled in the difference so that the bounds in Lemma 6.1 can be applied. In the above estimations, we used $\varepsilon < 1$ twice. If we now assume that w.l.o.g. we have $\widetilde{H}_\alpha(A|B)_\rho \geq \widetilde{H}_\alpha(A|B)_\sigma$, we can deduce the following from the bound above:

$$\begin{aligned} |\widetilde{H}_\alpha(A|B)_\rho - \widetilde{H}_\alpha(A|B)_\sigma| &= \widetilde{H}_\alpha(A|B)_\rho - \widetilde{H}_\alpha(A|B)_\sigma \\ &= \frac{1}{\alpha-1}\log\frac{\widetilde{Q}_\alpha(A|B)_\rho}{\widetilde{Q}_\alpha(A|B)_\sigma} \\ &= \frac{1}{\alpha-1}\log\left(\frac{\widetilde{Q}_\alpha(A|B)_\rho - \widetilde{Q}_\alpha(A|B)_\sigma}{\widetilde{Q}_\alpha(A|B)_\sigma} + 1\right) \\ &\leq c(\alpha,m,d_A,d_{AB})\sqrt{\varepsilon} \end{aligned}$$

where we employed $\log(x+1) \leq x$ for $x \geq 0$ and finally $1 \leq \widetilde{Q}_\alpha(A|B)_\sigma$. This concludes the claim.

It is straightforward to derive a continuity bound for the non-variational sandwiched Rényi conditional mutual information as a consequence of the previous result.

Corollary 6.4 *Let* $\mathcal{H}_{ABC} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, $d_{ABC}^{-1} > m > 0$, $\rho,\sigma \in \mathcal{S}(\mathcal{H}_{ABC})$ *with* $\rho,\sigma \geq m\,\mathbb{1}$. *If* $\frac{1}{2}\|\rho-\sigma\|_1 \leq \varepsilon$, *then for* $\alpha \in [1/2,1) \cup (1,\infty)$

$$|\widetilde{I}_\alpha(A:C|B)_\rho - \widetilde{I}_\alpha(A:C|B)_\sigma| \leq 2c(\alpha,m,d_C,d_{ABC})\sqrt{\varepsilon}\,.$$

with $c(\alpha,m,d_C,d_{ABC})$ *from Corollary 6.3.*

[3] One can for example use $\frac{\mathbb{1}}{d_{AB}}$ to perturb.

Proof This is a direct consequence of Corollary 6.3, since we can write

$$|\widetilde{I}_\alpha(A:C|B)_\rho - \widetilde{I}_\alpha(A:C|B)_\sigma| \le |\widetilde{H}_\alpha(C|B)_\rho - \widetilde{H}_\alpha(C|B)_\sigma| \\ +|\widetilde{H}_\alpha(C|AB)_\rho - \widetilde{H}_\alpha(C|AB)_\sigma|\,.$$

and $\frac{1}{2}\|\rho_{ABC} - \sigma_{ABC}\|_1 \le \varepsilon$ implies $\frac{1}{2}\|\rho_{BC} - \sigma_{BC}\|_1 \le \varepsilon$ by DPI of $\|\cdot\|_1$. Similarly $\rho_{ABC}, \sigma_{ABC} \ge m\,\mathbb{1}$ implies $\rho_{BC}, \sigma_{BC} \ge m\,\mathbb{1}$.

6.2. Application: Approximate Quantum Markov Chains

In this section, we use the continuity bound for the (non-variational) sandwiched Rényi conditional mutual information to derive a stability result for approximate quantum Markov chains. Consider a tripartite Hilbert space $\mathcal{H}_{ABC} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ and $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_{ABC})$. Let us denote by $\mathcal{R}^{\rho}_{B\to BC}(\cdot)$ the Petz recovery map for the partial trace in B, given for $X \in \mathcal{B}(\mathcal{H}_{BC})$ by

$$\mathcal{R}^{\rho}_{B\to BC}(X) := \rho_{BC}^{1/2}(\rho_B^{-1/2}\mathrm{tr}_C[X]\rho_B^{-1/2} \otimes \mathbb{1}_C)\rho_{BC}^{1/2}\,.$$

This can be lifted to a map on $\mathcal{B}(\mathcal{H}_{ABC})$ as $\mathrm{id}_A \otimes \mathcal{R}^{\rho}_{B\to BC}$. To enhance readability, we will omit id_A, $\mathbb{1}_C$, and other identity operators whenever their inclusion is clear from the context. We remind the reader, however, that all matrices in each product act on the same space and are implicitly extended with identities as needed. It is well known [19,29] that

$$I(A:C|B)_\rho = 0 \;\Leftrightarrow\; \rho_{ABC} = \mathcal{R}^{\rho}_{B\to BC}(\rho_{AB}) \;\Leftrightarrow\; \widetilde{I}_\alpha(A:C|B)_\rho = 0 \\ \text{for any } \alpha \in (1/2,1)\cup(1,\infty), \tag{23}$$

where the last equivalence can be found, e.g. in [16, Corollary 4.23]. We can further replace the Petz recovery map in the previous equivalences by the *universal Petz recovery map* [21,38] $\mathcal{R}^{\rho,u}_{B\to BC}(\cdot)$, given by

$$\mathcal{R}^{\rho,u}_{B\to BC}(X) := \int_{\mathbb{R}} \mathcal{R}^{\rho,t/2}_{B\to BC}(X)\,\beta_0(t)dt\,, \quad \text{with} \quad \beta_0(t) = \frac{\pi}{2}(\cosh(\pi t)+1)^{-1}\,,$$

where $\mathcal{R}^{\rho,t}_{B\to BC}(\cdot)$ is the *rotated Petz recovery map*, namely

$$\mathcal{R}^{\rho,t}_{B\to BC}(X) := \rho_{BC}^{\frac{1}{2}+it}\rho_B^{-\frac{1}{2}-it}\mathrm{tr}_C[X]\rho_B^{-\frac{1}{2}-it}\rho_{BC}^{\frac{1}{2}+it}\,.$$

A state ρ_{ABC} satisfying Equation (23) is called a *quantum Markov chain*. This notion can be extended to an approximate version in the following way: Given a small $\varepsilon > 0$, a state $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ is an *approximate quantum Markov chain* [37] if, and only if, $I_\rho(A:C|B) < \varepsilon$. In an analogous way, for $\alpha \in (1/2,1)\cup(1,\infty)$, we can say that $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_{ABC})$ is an α*-approximate quantum Markov chain* whenever $\widetilde{I}_\alpha(A:C|B)_\rho < \varepsilon$.

In [8, Section 7.3], some of the authors of the current manuscript proved that a state $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ is an approximate quantum Markov chain if, and only if, it is close to its reconstructed state under the Petz recovery map. As a consequence of our new continuity bounds for sandwiched Rényi divergences, we can extend now that result to the case of α-approximate quantum Markov chains in the following way.

Proposition 6.5 *Let $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ be positive definite. Given $\alpha \in (1/2, 1) \cup (1, \infty)$, ρ_{ABC} is an α-approximate quantum Markov chain if, and only if, it is close to its (rotated, universal) Petz recovery. More specifically, we have for $\alpha \in (1/2, 1)$*

$$\frac{\alpha}{1-\alpha} \log\left(1 + \left(\mathcal{K}\,\|\rho_{ABC} - \rho_{BC}^{\frac{1}{2}+it}\rho_B^{-\frac{1}{2}-it}\rho_{AB}\rho_B^{-\frac{1}{2}-it}\rho_{BC}^{\frac{1}{2}+it}\|_1\right)^{\frac{1}{1-\frac{1}{2\alpha}-\varepsilon}}\right)$$
$$\leq \widetilde{I}_\alpha(A:C|B)_\rho$$
$$\leq c\left(\alpha, \|\rho_{ABC}^{-1}\|_\infty^{-1}, d_C, d_{ABC}\right)$$
$$\|\rho_{ABC} - \rho_{BC}^{1/2}\rho_B^{-1/2}\rho_{AB}\rho_B^{-1/2}\rho_{BC}^{1/2}\|_1^{1/2},$$

for any $\varepsilon \in \left(0, 1-\frac{1}{2\alpha}\right)$, with

$$\mathcal{K} = \left(\left(\frac{\pi}{e\varepsilon \sin(\pi\frac{1-\alpha}{\alpha})}\right)^{1/2} + 8\right)\frac{\pi}{2\cosh(\pi t)},$$

and for $\alpha \in (1, \infty)$,

$$\frac{\alpha}{\alpha-1} \log\left(1 + \left(\mathcal{K}'\,\|\rho_{ABC} - \rho_{BC}^{\frac{1}{2}+it}\rho_B^{-\frac{1}{2}-it}\rho_{AB}\rho_B^{-\frac{1}{2}-it}\rho_{BC}^{\frac{1}{2}+it}\|_1\right)^{\frac{1}{\frac{1}{2\alpha}-\varepsilon}}\right)$$
$$\leq \widetilde{I}_\alpha(A:C|B)_\rho$$
$$\leq c\left(\alpha, \|\rho_{ABC}^{-1}\|_\infty^{-1}, d_C, d_{ABC}\right)$$
$$\|\rho_{ABC} - \rho_{BC}^{1/2}\rho_B^{-1/2}\rho_{AB}\rho_B^{-1/2}\rho_{BC}^{1/2}\|_1^{1/2},$$

for any $\varepsilon \in \left(0, \frac{1}{2\alpha}\right)$, with

$$\mathcal{K}' = d_C^{\frac{2(1-\alpha)}{\alpha}}\left(\left(\frac{\pi}{e\varepsilon \sin(\pi\frac{\alpha-1}{\alpha})}\right)^{1/2} + 8\right)\frac{\pi}{2\cosh(\pi t)},$$

and $c(\alpha, \cdot, \cdot, \cdot)$ the function from Corollary 6.3.

Proof The lower bounds appear in [16, Corollary 4.21], where the only difference is a term[4] $\widetilde{Q}_\infty(\rho_{ABC}\|\rho_{AB} \otimes \mathbb{1}_C/d_C)$ which we lower bounded by one in $\mathcal{K}$ and $\mathcal{K}'$ as well as upper bounded it by d_C^2 in $\mathcal{K}'$. For the RHS we first note that

$$\widetilde{I}_\alpha(A:C|B)_\rho = \widetilde{H}_\alpha(C|B)_\rho - \widetilde{H}_\alpha(C|AB)_\rho \leq \widetilde{H}_\alpha(C|AB)_{\mathcal{R}^\rho_{B\to BC}(\rho)} - \widetilde{H}_\alpha(C|AB)_\rho$$

by the data processing inequality. An application of Corollary 6.3 proves the claim.

6.3. General Continuity Bounds Via the ALAFF Method

We conclude this section by deriving some continuity bounds for sandwiched Rényi divergences for both inputs. To that end, we first prove a continuity bound for $\widetilde{Q}_\alpha(\cdot\|\cdot)$ which we will subsequently use to obtain one for $\widetilde{D}_\alpha(\cdot\|\cdot)$. For that, we consider a perturbed Δ-invariant set $\mathcal{S}_0$ which is a modification of the aforementioned $\mathcal{S}_{\text{ker}}$.

[4] Note that the following is the notation from [16] where $\widetilde{Q}_\infty(\rho\|\sigma) := \|\rho^{1/2}\sigma^{-1}\rho^{1/2}\|_\infty$.

Theorem 6.6 *Let* $1 > 2m > 0$ *and*

$$\mathcal{S}_0 := \{(\rho,\sigma) \in \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \ : \ \ker\sigma \subseteq \ker\rho \ , \ 2m \leq m_\sigma\} \,,$$

where m_σ *is the minimal eigenvalue of* σ. *Then,* $\widetilde{Q}_\alpha(\cdot\|\cdot)$ *is uniformly continuous on* $\mathcal{S}_0$. *For* $(\rho_1,\sigma_1),(\rho_2,\sigma_2) \in \mathcal{S}_0$ *with* $\frac{1}{2}\|\rho_1-\rho_2\|_1 \leq \varepsilon \leq 1$ *and* $\frac{1}{2}\|\sigma_1-\sigma_2\|_1 \leq \delta \leq 1$, *we have for* $\alpha \in [1/2,1)$

$$|\widetilde{Q}_\alpha(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_2\|\sigma_2)| \leq (1+\sqrt{2})\sqrt{\varepsilon} + 2c(\alpha,m,d_{\mathcal{H}})\sqrt{\delta} \,, \tag{24}$$

and for $\alpha \in (1,\infty)$

$$|\widetilde{Q}_\alpha(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_2\|\sigma_2)| \leq (1+\sqrt{2})m^{1-\alpha}\sqrt{\varepsilon} + 2c(\alpha,m,d_{\mathcal{H}})\sqrt{\delta} \,. \tag{25}$$

with

$$c(\alpha,m,d_{\mathcal{H}}) = \begin{cases} \frac{1+\sqrt{2}(1-\alpha)(\log(m^{-1})+m^{-1}+2)m^{\alpha-1}}{1-md_{\mathcal{H}}} & \alpha \in [1/2,1) \\ m^{1-\alpha}\frac{1+\sqrt{2}(\alpha-1)(\log(m^{-1})+m^{-1}+2)}{1-md_{\mathcal{H}}} & \alpha \in (1,\infty) \end{cases}$$

Proof Let $(\rho_1,\sigma_1),(\rho_2,\sigma_2) \in \mathcal{S}_0$ with $\frac{1}{2}\|\rho_1-\rho_2\| \leq \varepsilon \leq 1$ and $\frac{1}{2}\|\sigma_1-\sigma_2\| \leq \delta \leq 1$. We define

$$\overline{\sigma} = \frac{1}{2}\sigma_1 + \frac{1}{2}\sigma_2 \,, \tag{26}$$

and obtain

$$\frac{1}{2}\|\overline{\sigma}-\sigma_1\|_1 = \frac{1}{4}\|\sigma_1-\sigma_2\|_1 \leq \frac{\delta}{2} \leq 1 \,,$$
$$\frac{1}{2}\|\overline{\sigma}-\sigma_2\|_1 = \frac{1}{4}\|\sigma_1-\sigma_2\|_1 \leq \frac{\delta}{2} \leq 1 \,.$$

Using this, the triangle inequality shows

$$\begin{aligned} |\widetilde{Q}_\alpha &(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_2\|\sigma_2)| \\ &\leq \underbrace{|\widetilde{Q}_\alpha(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_1\|\overline{\sigma})|}_{(I)} + \underbrace{|\widetilde{Q}_\alpha(\rho_1\|\overline{\sigma}) - \widetilde{Q}_\alpha(\rho_2\|\overline{\sigma})|}_{(II)} \\ &\quad + \underbrace{|\widetilde{Q}_\alpha(\rho_2\|\overline{\sigma}) - \widetilde{Q}_\alpha(\rho_2\|\sigma_2)|}_{(III)} \,. \end{aligned} \tag{27}$$

In the following, we bound each of these terms separately for the two cases $\alpha \in [1/2,1)$ and $\alpha \in (1,\infty)$. Let us begin with the case $\alpha \in [1/2,1)$:

- For (II), we require a continuity bound for $\widetilde{Q}_\alpha(\cdot\|\cdot)$ in the first argument: Note that $\rho \mapsto \widetilde{Q}_\alpha(\cdot\|\overline{\sigma})$ as a map from the 0-perturbed Δ-invariant set $\mathcal{S}(\mathcal{H})$ to the reals is ALAFF using Lemma 6.1 and Remark 6.2 with $a = 0$ and $b = \big(-1+p^\alpha+(1-p)^\alpha\big)$. Moreover, we have that

$$\sup_{\substack{\rho_1,\rho_2\in\mathcal{S}(\mathcal{H}) \\ \frac{1}{2}\|\rho_1-\rho_2\|=1}} |\widetilde{Q}_\alpha(\rho_1\|\overline{\sigma}) - \widetilde{Q}_\alpha(\rho_2\|\overline{\sigma})| \leq 1,$$

where we used that $0 \le \widetilde{Q}_\alpha(\rho_1\|\overline{\sigma}) \le 1$. Employing [8, Theorem 4.6], we conclude:

$$|\widetilde{Q}_\alpha(\rho_1\|\overline{\sigma}) - \widetilde{Q}_\alpha(\rho_2\|\overline{\sigma})| \le \varepsilon + (1+\varepsilon)\left(-1 + \left(\frac{\varepsilon}{1+\varepsilon}\right)^\alpha + \left(\frac{1}{1+\varepsilon}\right)^\alpha\right). \tag{28}$$

- For (I) and (III), we need a continuity bound for $\widetilde{Q}_\alpha(\cdot\|\cdot)$ in the second argument. We argue for (I) and (III) completely analogous:

  By Lemma 6.1, we find that $\sigma \mapsto \widetilde{Q}_\alpha(\rho_1\|\cdot)$ on $\mathcal{S}_{\ge m}(\mathcal{H}) = \{\rho \in \mathcal{S}(\mathcal{H}) \,:\, \sigma \ge m\,\mathbb{1}\}$ is ALAFF with $b = \xi(\alpha, p, m\,\mathbb{1}, m\,\mathbb{1})$ and $a = 0$. Moreover, we have that

$$\sup_{\substack{\sigma_1,\sigma_2\in\mathcal{S}_{\ge m}(\mathcal{H})\\ \frac{1}{2}\|\sigma_1-\sigma_2\|_1 = 1-m}} |\widetilde{Q}_\alpha(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_1\|\sigma_2)| \le 1.$$

  and that $\mathcal{S}_{\ge m}(\mathcal{H})$ is $md_{\mathcal{H}}$-perturbed Δ-invariant. Hence [8, Theorem 4.6] and the fact that $\sigma_1, \overline{\sigma} \in \mathcal{S}_{\ge m}(\mathcal{H})$ allows us to conclude:

$$\begin{aligned}
|\widetilde{Q}_\alpha(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_1\|\overline{\sigma})| &\le \frac{\delta}{1-md_{\mathcal{H}}} + \frac{1-md_{\mathcal{H}}+\delta}{1-md_{\mathcal{H}}}\xi\left(\alpha, \frac{\delta}{1-md_{\mathcal{H}}+\delta}, m\,\mathbb{1}, m\,\mathbb{1}\right)\\
&\le \frac{\delta}{1-md_{\mathcal{H}}} + \frac{\sqrt{1-md_{\mathcal{H}}+\delta}}{1-md_{\mathcal{H}}}(1-\alpha)(\log(m^{-1}) + m^{-1} + 2)m^{\alpha-1}\sqrt{\delta}\\
&\le \underbrace{\frac{1+\sqrt{2}(1-\alpha)(\log(m^{-1}) + m^{-1} + 2)m^{\alpha-1}}{1-md_{\mathcal{H}}}}_{c(\alpha,m,d_{\mathcal{H}})}\sqrt{\delta}\,,
\end{aligned} \tag{29}$$

  where we are using $\delta/2 \le \delta \le \sqrt{\delta}$ for $\delta \in [0,1]$.

Merging these bounds, we find

$$\begin{aligned}
|\widetilde{Q}_\alpha(\rho_1\|\sigma_1) - \widetilde{Q}_\alpha(\rho_2\|\sigma_2)| &\le \varepsilon + (1+\varepsilon)\left(-1 + \left(\frac{\varepsilon}{1+\varepsilon}\right)^\alpha + \left(\frac{1}{1+\varepsilon}\right)^\alpha\right)\\
&\quad + 2c(\alpha, m, d_{\mathcal{H}})\sqrt{\delta}\\
&\le (1+\sqrt{2})\sqrt{\varepsilon} + 2c(\alpha, m, d_{\mathcal{H}})\sqrt{\delta}\,,
\end{aligned}$$

where we are using the following inequality, proven with elementary calculus:

$$(1+\varepsilon)\left(-1 + \left(\frac{\varepsilon}{1+\varepsilon}\right)^\alpha + \left(\frac{1}{1+\varepsilon}\right)^\alpha\right) \le \sqrt{2\varepsilon}\,.$$

This concludes the case $\alpha \in [1/2, 1)$. For $\alpha \in (1, \infty)$ the bound is obtained similarly with the only difference that the uniform bounds in (I), (II), (III) are given by $m^{\alpha-1}$ instead of 1 (e.g. for (I) we have $\sup_{\rho_1,\rho_2\in\mathcal{S}(\mathcal{H})} |\widetilde{Q}_\alpha(\rho_1\|\overline{\sigma}) - \widetilde{Q}_\alpha(\rho_2\|\overline{\sigma})| \le m^{\alpha-1}$).

From this result, we can derive the following continuity bound for sandwiched Rényi divergences with respect to the first and second input.

Corollary 6.7 *Let* $1 > 2m > 0$ *and*

$$\mathcal{S}_0 := \{(\rho, \sigma) \in \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \,:\, 2m \le m_\sigma\}\,,$$

where m_σ is the minimal eigenvalue of σ. Then, $\widetilde{Q}_\alpha(\cdot\|\cdot)$ is uniformly continuous on $\mathcal{S}_0$. For $(\rho_1,\sigma_1),(\rho_2,\sigma_2) \in \mathcal{S}_0$ with $\frac{1}{2}\|\rho_1-\rho_2\|_1 \le \varepsilon \le 1$ and $\frac{1}{2}\|\sigma_1-\sigma_2\|_1 \le \delta \le 1$, we have for $\alpha \in [1/2,1)$

$$|\widetilde{D}_\alpha(\rho_1\|\sigma_1) - \widetilde{D}_\alpha(\rho_2\|\sigma_2)| \le \frac{m^{\alpha-1}}{1-\alpha}\left[(1+\sqrt{2})\sqrt{\varepsilon} + 2c(\alpha,m,d_{\mathcal{H}})\sqrt{\delta}\right], \quad (30)$$

and for $\alpha \in (1,\infty)$

$$|\widetilde{D}_\alpha(\rho_1\|\sigma_1) - \widetilde{D}_\alpha(\rho_2\|\sigma_2)| \le \frac{1}{\alpha-1}\left[(1+\sqrt{2})m^{1-\alpha}\sqrt{\varepsilon} + 2c(\alpha,m,d_{\mathcal{H}})\sqrt{\delta}\right], \quad (31)$$

with $c(\alpha,m,d_{\mathcal{H}})$ as in Theorem 6.6.

Proof For $\alpha \in (1,\infty)$, note that

$$\begin{aligned}\widetilde{D}_\alpha(\rho_1\|\sigma_1) - \widetilde{D}_\alpha(\rho_2\|\sigma_2) &= \frac{1}{\alpha-1}\log\left(\frac{\widetilde{Q}_\alpha(\rho_1\|\sigma_1)}{\widetilde{Q}_\alpha(\rho_2\|\sigma_2)}\right)\\ &= \frac{1}{\alpha-1}\log\left(\frac{\widetilde{Q}_\alpha(\rho_1\|\sigma_1)-\widetilde{Q}_\alpha(\rho_2\|\sigma_2)}{\widetilde{Q}_\alpha(\rho_2\|\sigma_2)}+1\right)\\ &\le \frac{1}{\alpha-1}\frac{\widetilde{Q}_\alpha(\rho_1\|\sigma_1)-\widetilde{Q}_\alpha(\rho_2\|\sigma_2)}{\widetilde{Q}_\alpha(\rho_2\|\sigma_2)}\\ &\le \frac{1}{\alpha-1}\left(\widetilde{Q}_\alpha(\rho_1\|\sigma_1)-\widetilde{Q}_\alpha(\rho_2\|\sigma_2)\right),\end{aligned}$$

where we are using $\widetilde{Q}_\alpha(\rho_2\|\sigma_2) \ge 1$ and $\log(x+1) \le x$ for every $x > -1$. Exchanging the roles of ρ_1,σ_1 with ρ_2,σ_2, we conclude

$$\begin{aligned}|\widetilde{D}_\alpha(\rho_1\|\sigma_1) - \widetilde{D}_\alpha(\rho_2\|\sigma_2)| &\le \frac{1}{\alpha-1}\left|\widetilde{Q}_\alpha(\rho_1\|\sigma_1)-\widetilde{Q}_\alpha(\rho_2\|\sigma_2)\right|\\ &\le \frac{1}{\alpha-1}\left[(1+\sqrt{2})m^{1-\alpha}\sqrt{\varepsilon} + 2c(\alpha,m,d_{\mathcal{H}})\sqrt{\delta}\right].\end{aligned}$$

For $\alpha \in [1/2,1)$ the proof follows the same lines.

7. Discussion

In this paper, we presented a framework for proving uniform continuity bounds for sandwiched Rényi divergencies and presented a comprehensive analysis of the continuity properties of the sandwiched Rényi conditional entropy, the sandwiched Rényi mutual information and the sandwiched Rényi divergence with fixed second argument. While our almost additive approach drew inspiration from [25] and our operator space approach from [5], we further developed and extended the methodologies introduced in both papers. This extension enabled us to enhance the bounds for the sandwiched Rényi conditional entropy and broaden the applicability of these methods to encompass other entropic measures, which could find practical use in resource theories, for example. Combining the two other approaches, the mixed approach yields bounds that perform well in all regimes and are optimal for large d_A and small α. Additionally, we explored the ALAFF method ([8]), devised by some of the authors. However, it yielded bounds that underperformed in comparison to those presented here. Comparing the bounds obtained by the operator space, the almost additive and mixed approaches we highlight their strengths and weaknesses. We find that the operator-space approach gives the best bound for large α, while the mixed and almost additive methods are optimal for low α. To the best of our knowledge, we provide the tightest bounds known for the sandwiched Rényi conditional entropy in the regime $\alpha \in (1, \infty)$, the sandwiched Rényi mutual information in the range $\alpha \in [1/2, 1) \cup (1, \infty)$ and the sandwiched Rényi divergence with fixed second argument in the range $\alpha \in [1/2, 1) \cup (1, \infty)$. Finally, we provide an application of continuity bounds to quantum Markov states. Here, we resort to the ALAFF method, since the almost additive, operator space and mixed approaches rely on the optimization in the second argument and are hence not applicable in this context.

Acknowledgements

We thank Li Gao and Marius Lemm for the valuable feedback and fruitful discussions. A.B. acknowledges the support of the French National Research Agency in the framework of the "France 2030" program (ANR-11-LABX-0025-01) for the LabEx PERSYVAL, A.C., P.G. and T.M. that of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 470903074 - TRR 352, and T.M. that of the Munich Center for Quantum Sciences and Technology.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Alicki, R., Fannes, M.: Continuity of quantum conditional information. J. Phys. A Math. Gen. **37**(5), L55–L57 (2004)

[2] Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis-A Hitchhiker's Guide. Springer (2006)

[3] Anshu, A., Hsieh, M.-H., Jain, R.: Quantifying resources in general resource theory with catalysts. Phys. Rev. Lett. **121**(19), 190504 (2018)

[4] Audenaert, K.M.R.: A sharp continuity estimate for the von Neumann entropy. J. Phys. A: Math. Theor. **40**(28), 8127 (2007)

[5] Beigi, S., Goodarzi, M.M.: Operator-valued Schatten spaces and quantum entropies. Lett. Math. Phys. **113**(5), 1–54 (2023)

[6] Belavkin, V.P., Staszewski, P.: c^*-algebraic generalization of relative entropy and entropy. Ann. l'inst. Henri Poincaré Sect. A Phy. Théor. **37**(1), 51–58 (1982)

[7] Bhatia, R.: Matrix Analysis, volume 169 of Graduate Texts in Mathematics. Springer (1997)

[8] Bluhm, A., Capel, A., Gondolf, P., Pérez-Hernández, A.: Continuity bounds for quantum entropic quantities via almost convexity. IEEE Trans. Inf. Theory **69**(9), 5869–5901 (2023)

[9] Bluhm, A., Capel, A., Gondolf, P., Pérez-Hernández, A.: General continuity bounds for quantum relative entropies. In: 2023 IEEE International Symposium on Information Theory (ISIT), pp. 162–167 (2023)

[10] Bourin, J.-C., Uchiyama, M.: A matrix subadditivity inequality for $f(a+b)$ and $f(a) + f(b)$. Linear Algebra Appl. **423**(2–3), 512–518 (2007)

[11] Cheng, H.C., Gao, L.: Tight analysis of convex splitting with applications in quantum information theory. In: 2023 IEEE International Symposium on Information Theory (ISIT), pp. 192–197 (2023)

[12] Cheng, H.C., Gao, L., Berta, M.: Quantum broadcast channel simulation via multipartite convex splitting. arXiv preprint arXiv:2304.12056, 2023

[13] Donald, M.J., Horodecki, M.: Continuity of relative entropy of entanglement. Phys. Lett. A **264**, 257–260 (1999)

[14] Dunford, N., Schwartz, J.T.: Linear Operators, Part 1: General Theory, volume 7 of Pure and Applied Mathematics. John Wiley & Sons (1958)

[15] Fannes, M.: A continuity property of the entropy density for spin lattice systems. Commun. Math. Phys. **31**(4), 291–294 (1973)

[16] Gao, L., Wilde, M.M.: Recoverability for optimized quantum f-divergences. J. Phys. A Math. Theor. **54**(38), 385302 (2021)

[17] Gour, G.: Resources of the quantum world. arXiv preprint arXiv:2402.05474, 2024

[18] Hayashi, M., Tomamichel, M.: Correlation detection and an operational interpretation of the Rényi mutual information. J. Math. Phys. **57**, 102201 (2016)

[19] Hayden, P., Jozsa, R., Petz, D., Winter, A.: Structure of states which satisfy strong subadditivity of quantum entropy with equality. Commun. Math. Phys. **246**(2), 359–374 (2004)

[20] Hiai, F., Mosonyi, M.: Quantum Rényi divergences and the strong converse exponent of state discrimination in operator algebras. Ann. Henri Poincaré **24**(5), 1681–1724 (2022)

[21] Junge, M., Renner, R., Sutter, D., Wilde, M.M., Winter, A.: Universal recovery map for approximate Markov chains. Ann. Henri Poincaré **19**, 2955–2978 (2018)

[22] Khatri, S., Wilde, M.M.: Principles of quantum communication theory: a modern approach. arXiv preprint arXiv:2011.04672, 2020

[23] Lami, L., Shirokov, M.E.: Attainability and lower semi-continuity of the relative entropy of entanglement and variations on the theme. Ann. Henri Poincaré **24**(12), 4069–4137 (2023)

[24] Lieb, E.H.: Convex trace functions and the Wigner-Yanase-Dyson conjecture. Adv. Math. **11**(3), 267–288 (1973)

[25] Marwah, A., Dupuis, F.: Uniform continuity bound for sandwiched Rényi conditional entropy. J. Math. Phys. **63**(5), 052201 (2022)

[26] McCarthy, C.A.: c_p. Israel J. Math. **5**(4), 249–271 (1967)

[27] Mosonyi, M., Hiai, F.: On the quantum Rényi relative entropies and related capacity formulas. IEEE Trans. Inf. Theory **57**(4), 2474–2487 (2011)

[28] Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., Tomamichel, M.: On quantum Rényi entropies: a new generalization and some properties. J. Math. Phys. **54**(12), 122203 (2013)

[29] Petz, D.: Monotonicity of quantum relative entropy revisited. Rev. Math. Phys. **15**(01), 79–91 (2003)

[30] Petz, D.: Quantum Information Theory and Quantum Statistics. Springer (2008)

[31] Rubboli, R., Tomamichel, M.: Fundamental limits on correlated catalytic state transformations. Phys. Rev. Lett. **129**(12), 120404 (2022)

[32] Rubboli, R., Tomamichel, M.: New additivity properties of the relative entropy of entanglement and its generalizations. arXiv preprint arXiv:2211.12804, 2022

[33] Rudin, W.: Principles of Mathematical Analysis. International series in pure and applied mathematics. 3rd edn, McGraw-Hill (1976)

[34] Shirokov, M.E.: Tight uniform continuity bounds for the quantum conditional mutual information, for the Holevo quantity, and for capacities of quantum channels. J. Math. Phys. **58**(10), 102202 (2017)

[35] Shirokov, M.E.: Advanced Alicki-Fannes-Winter method for energy-constrained quantum systems and its use. Quantum Inf. Process. **19**(5), 164 (2020)

[36] Shirokov, M.E.: Quantifying continuity of characteristics of composite quantum systems. Phys. Scr. **98**(4), 042002 (2023)

[37] Sutter, D.: Approximate Quantum Markov Chains. Springer International Publishing (2018)

[38] Sutter, D., Fawzi, O., Renner, R.: Universal recovery map for approximate Markov chains. Proc. R. Soc. Lond. A Math. Phys. Eng. Sci. **472**(2186), 20150623 (2016)

[39] Synak-Radtke, B., Horodecki, M.: On asymptotic continuity of functions of quantum states. J. Phys. A Math. Gen. **39**(26), L423 (2006)

[40] Tomamichel, M.: Quantum Information Processing with Finite Resources. Springer International Publishing (2016)

[41] Wilde, M.M., Winter, A., Yang, D.: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. Commun. Math. Phys. **331**(2), 593–622 (2014)

[42] Winter, A.: Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. Commun. Math. Phys. **347**(1), 291–313 (2016)

[43] Zhu, H., Hayashi, M., Chen, L.: Coherence and entanglement measures based on Rényi relative entropies. J. Phys. A Math. Theor. **50**(47), 475303 (2017)

Andreas Bluhm
Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
38000 Grenoble
France
e-mail: andreas.bluhm@univ-grenoble-alpes.fr

Ángela Capel, Paul Gondolf and Tim Möbus
Fachbereich Mathematik, Universität Tübingen
72076 Tübingen
Germany
e-mail: angela.capel@uni-tuebingen.de;
paul.gondolf@uni-tuebingen.de;
moebustim@gmail.com

Ángela Capel
Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Cambridge CB3 0WA
UK

Tim Möbus
Department of Mathematics
Technische Universität München
80333 München
Germany

Communicated by David Pérez-García.

Received: September 25, 2023.

Accepted: November 19, 2024.

Energy preserving evolutions over Bosonic systems

Paul Gondolf[1], Tim Möbus[2,3,1], and Cambyse Rouzé[2,3,4]

[1]Department of Mathematics, Eberhard Karls University Tübingen, 72074 Tübingen, Germany
[2]Department of Mathematics, Technical University of Munich, 80333 München, Germany
[3]Munich Center for Quantum Science and Technology (MCQST), 80799 München, Germany
[4]Inria, Télécom Paris - LTCI, Institut Polytechnique de Paris, 91120 Palaiseau, France

The exponential convergence to invariant subspaces of quantum Markov semigroups plays a crucial role in quantum information theory. One such example is in bosonic error correction schemes, where dissipation is used to drive states back to the code-space — an invariant subspace protected against certain types of errors. In this paper, we investigate perturbations of quantum dynamical semigroups that operate on continuous variable (CV) systems and admit an invariant subspace. First, we prove a generation theorem for quantum Markov semigroups on CV systems under the physical assumptions that (i) the generator is in GKSL form with corresponding jump operators defined as polynomials of annihilation and creation operators; and (ii) the (possibly unbounded) generator increases all moments in a controlled manner. Additionally, we show that the level sets of operators with bounded first moments are admissible subspaces of the evolution, providing the foundations for a perturbative analysis. Our results also extend to time-dependent semigroups and multi-mode systems. We apply our general framework to two settings of interest in continuous variable quantum information processing. First, we provide a new scheme for deriving continuity bounds on the energy-constrained capacities of Markovian perturbations of quantum dynamical semigroups. Second, we provide quantitative perturbation bounds for the steady state of the quantum Ornstein-Uhlenbeck semigroup and the invariant subspace of the photon dissipation used in bosonic error correction.

Paul Gondolf: paul.gondolf@uni-tuebingen.de
Tim Möbus: moebustim@gmail.com
Cambyse Rouzé: rouzecambyse@gmail.com

Accepted in Quantum 2024-30-11, click title to verify.

Table of Contents

1 Introduction **3**

2 Preliminaries **6**
- 2.1 Basic Banach space theory . . . 7
- 2.2 Weighted norms and compact embeddings . . . 8
- 2.3 Strongly continuous semigroups and evolution systems . . . 9
- 2.4 Continuous variable quantum systems . . . 13

3 Sobolev preserving quantum Markov semigroups **16**
- 3.1 Strongly continuous bosonic semigroups . . . 16
- 3.2 Bosonic evolution systems . . . 23
- 3.3 Multi-mode extension . . . 24

4 Examples of Sobolev preserving semigroups **27**
- 4.1 Quantum Ornstein Uhlenbeck semigroup . . . 28
- 4.2 Photon-dissipation and CAT qubits . . . 29

5 Perturbation bounds **41**
- 5.1 Gaussian perturbations of the quantum Ornstein Uhlenbeck semigroup . . . 42
- 5.2 Photon-dissipation and CAT qubits . . . 44
- 5.3 Application: entropic and capacity continuity bounds . . . 45

6 Discussion and open questions **47**

7 Declarations **47**

8 References **47**

Appendices **53**

Acknowledgments: We would like to thank Yu-Jie Liu for the fruitful discussions that initiated this project. Moreover, we would like to thank Marius Lemm, Angela Capel-Cuevas, Simone Warzel, Michael M. Wolf, Libor Caha, Vjosa Blakaj, Shin Ho Choe, Robert Salzmann, Simon Becker, Lauritz van Luijk, Niklas Galke for the valuable feedback and discussions on the topic. Moreover, we would like to thank Jochen Schmid for his help with questions on evolution systems. T.M. and C.R. acknowledge the support of the Munich Center for Quantum Sciences and Technology, C.R. that of the Humboldt Foundation, and P.G., T.M., and C.R. that of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 470903074 – TRR 352.

1 Introduction

Quantum processes are often described in physics by their infinitesimal action during short time intervals. When the process at a given time t can be assumed to be independent of previous times, an assumption often referred to as the memorylessness condition, the resulting dynamics can be formally described via a so-called master equation. The latter is an initial value problem of the form

$$\frac{\partial}{\partial t}\rho(t) = \mathcal{L}(\rho(t)), \qquad \rho(0) = \rho_0 .$$

In the case of uniformly continuous dual dynamics over a von Neumann algebra, i.e. a $*$-subalgebra of $\mathcal{B}(\mathcal{H})$ closed under involution (see [9, Sec. 2.4.2] for more details), the seminal results by Lindblad and Gorini, Kossakowski, and Sudarshan [46, 30] classify the generators of a quantum dynamical semigroup in terms of the following so-called GKSL form: $\mathcal{L}$ is a bounded operator on the space $\mathcal{T}_1(\mathcal{H})$ of trace-class operators that satisfies

$$\mathcal{L}(\rho) = -i[H,\rho] + \sum_{j=1}^{K} L_j\rho L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, \rho\} \tag{1}$$

for a Hamiltonian $H = H^\dagger \in \mathcal{B}(\mathcal{H})$ and so-called Lindblad operators $L_j \in \mathcal{B}(\mathcal{H})$, where $\{A,B\} \coloneqq AB + BA$ denotes the anticommutator of two bounded operators $A, B \in \mathcal{B}(\mathcal{H})$. In other words, the state at time $t \geq 0$ is described as $e^{t\mathcal{L}}(\rho)$, where the exponential can be defined e.g. in terms of its converging Taylor series. In that case, the set $(e^{t\mathcal{L}})_{t\geq 0}$ defines a quantum Markov semigroup (QMS), which is a time-continuous family of completely positive, trace-preserving maps.

Since their introduction, QMSs have become a standard tool and have been extensively studied in various areas of mathematical physics and quantum information processing. Unfortunately, an extension of the GKSL form (1) is known to fail for strongly continuous evolutions in general and thus requires additional assumptions. Conversely, unbounded operators satisfying an equation like (1) on a suitable domain can fail at generating quantum Markovian dynamics. Simple counterexamples can be constructed in the context of continuous variable (CV) quantum systems over $\mathcal{H} = L^2(\mathcal{H})$ as follows: denoting the creation and annihilation operators associated with a harmonic oscillator by $a^\dagger$ and a, respectively, the 2-photon pure birth process formally defined as in (1) with $K = 1$, $H = 0$ and $L_1 = (a^\dagger)^2$ leads to a semigroup satisfying the master equation but failing to preserve the trace [20, Example 3.3.]. Similar problems were encountered later on by Fagnola et al. [17, 25], who considered the problem in the Heisenberg instead of the Schrödinger picture. They solved the appearing issues by imposing additional technical conditions on the generators in question. A thorough analysis of semigroups that have GKSL form but fail to be trace-preserving, can be found in [63], where the authors also discuss the possibility of generators deviating from the GKSL form. Beyond generation theory, a priori estimates on quantum Markov semigroups are key for perturbation theory of C_0-semigroups, which have been considered for example in [16, 17].

In contrast, recent years have seen remarkable progress in the use of CV quantum systems. These systems encompass a wide range of applications in various areas of quantum information, including quantum communication [10, 36, 70, 66, 55, 67, 59, 44], sensing [1, 71, 48, 47] or simulation [27], enabled by advancements in non-classical radiation sources [53, 42, 37, 58, 22, 72]. Given the technological and experimental relevance of CV systems, there is a pressing need for a rigorous mathematical theory of quantum dynamical semigroups over such systems.

One specific area where CV systems governed by a Lindblad master equation like (1) have recently gained significant attention on theoretical as well as experimental grounds is in the field of bosonic quantum error correcting codes [31, 50, 33, 39, 15, 52, 49, 45, 60, 11, 7]. In particular, a certain class of CV codes known as CAT qubit codes has focused the attention of the community for their property of dynamically preserving quantum information through the action of a class of suitably engineered QMS which, loosely referred to as dissipative CAT-qubit dynamics in the present introduction. However, a mathematically rigorous analysis of these codes has only gotten little attention to the best of our knowledge, with the notable exception of [5].

Much theoretical work has focused on the more tractable generators of Gaussian dynamics semigroups, where the generator $\mathcal{L}$ is expressed as a quadratic form in the creation and annihilation operators [38, 26, 18, 2, 34]. For those generators, the Feller property as well as properties of the spectrum and convergence results are known [18, 12, 13, 14, 21]. Since generators of dissipative CAT-qubit dynamics typically involve higher order monomials in a and $a^\dagger$, the establishment of a more general theory of CV quantum Markov semigroups including them is timely.

1.1 Framework

Note that the technical details are introduced in Section 2. Here, we just give an overview of the framework and subsequently the results: In this paper, we consider an operator $\mathcal{L}$ on the space $\mathcal{T}_f$ of finite linear combinations of rank-one operators of the form $|k\rangle\langle l|$, where given $k \in \mathbb{N}$, $|k\rangle \in L^2(\mathbb{R})$ denotes the k-photon Fock state. We further assume that $\mathcal{L}$ satisfies the following two conditions: (i) $\mathcal{L}$ has a GKSL structure (1), where the Hamiltonian H as well as the jump operators L_j are polynomials of the annihilation and creation operators; (ii) the following condition is satisfied: for a divergent sequence $\{k_r\}_{r\in\mathbb{N}}$ in $\mathbb{R}_+$, there exist real coefficients $\{w_{k_r}\}_{r\in\mathbb{N}}$ such that for all states $\rho \in \mathcal{T}_f$:

$$\operatorname{tr}\Big[\mathcal{L}(\rho)(N+\mathbb{1})^{k_r/2}\Big] \le w_{k_r} \operatorname{tr}\Big[\rho(N+\mathbb{1})^{k_r/2}\Big] .$$

Above, $N := a^\dagger a = \sum_{n\in\mathbb{N}} n\,|n\rangle\langle n|$ denotes the photon number operator. This assumption implies not only that $\mathcal{L}$ defines a quantum dynamical semigroup, but also a quasi-contractive semigroup on the weighted Banach spaces $(\mathcal{D}(\mathcal{W}^k), \|\mathcal{W}^k(\cdot)\|_1)$ defined through the operator

$$\mathcal{W}(\cdot) := (N+\mathbb{1})^{1/4}\,(\cdot)\,(N+\mathbb{1})^{1/4} .$$

Here $\mathcal{D}(\mathcal{W}^k)$ denotes the domain of the operator $\mathcal{W}^k$. In the latter, we refer to these spaces as Sobolev spaces and denote them by $W^{k,1}$ in analogy with their classical analogues (see also [6]). Note that all mentioned definitions and results are extended to multi-mode systems later on. Next, we call operators $\mathcal{L}$ that satisfy both conditions (i) and (ii) generators of Sobolev preserving quantum dynamical semigroups. Indeed, in our first main result, we show that such operators generate QMSs with the extra property that the latter preserves Sobolev spaces. More precisely:

Theorem (Generation of bosonic semigroups, see Theorem 3.1) *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ be an operator defined on the Banach space $\mathcal{T}_{1,\mathrm{sa}}$ of self-adjoint, trace-class operators. If $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ satisfies conditions (i) and (ii) above, then its closure $\overline{\mathcal{L}}$ generates a strongly continuous, positivity preserving semigroup $(\mathcal{P}_t)_{t\ge0}$ on $W^{k,1}$ for all $k \in \mathbb{R}_+$ with*

$$\|\mathcal{P}_t\|_{W^{k,1}\to W^{k,1}} \le e^{\omega_k t} \quad \forall t \ge 0 .$$

where $\omega_k = \frac{k_{r_1}-k}{k_{r_1}-k_{r_0}}\omega_{k_{r_0}} + \frac{k-k_{r_0}}{k_{r_1}-k_{r_0}}\omega_{k_{r_1}}$ for an r such that $k_{r_0} \le k < k_{r_1}$. Finally, for $k = 0$, the semigroup is contractive and trace-preserving.

Additionally, our setup is directly suited to the establishment of a perturbation analysis akin to the result reported in [65] in the finite dimensional setting. Moreover, in some cases, our analysis allows us to conclude the existence of adherence points for the dynamics in the large time limit. We manage to prove the requirements (i)-(ii) of the generation theorems as well as rigorous perturbation analysis for several examples including dissipative CAT-qubit dynamics as well as Gaussian and quantum Ornstein Uhlenbeck generators. For the latter, we show for instance the following perturbation bound for all $t \geq 0$ (see Proposition 5.1 and Corollary 5.6):

Proposition *Let $(\mathcal{L}_{\text{qOU}}, \mathcal{T}_f)$ be the generator of the quantum Ornstein Uhlenbeck semigroup with jump operators λa and $\mu a^\dagger$, $\lambda > \mu \geq 0$ and $(\mathcal{L}_G, \mathcal{T}_f)$ a Gaussian perturbation with unique jump $\gamma a + \eta a^\dagger$ with $\gamma, \eta \in \mathbb{R}$, and $\varepsilon > 0$. Then, assuming $\lambda^2 - \mu^2 + |\gamma|^2 - |\eta|^2 > 0$, $\mathcal{L}_{\text{qOU}} + \varepsilon \mathcal{L}_G$ generates a positivity and Sobolev preserving semigroup on $W^{k,1}$ for $k \geq 1$, and there exist uniformly bounded functions $C(\varepsilon), D(\varepsilon)$ depending on $\lambda, \mu, |\eta|, |\gamma|$ such that, for all $t \geq 0$ and all state $\rho \in W^{k,1}$,*

$$\left\| \left(e^{t\mathcal{L}_{\text{qOU}}} - e^{t(\mathcal{L}_{\text{qOU}} + \varepsilon \mathcal{L}_G)} \right)(\rho) \right\|_1 \leq \varepsilon\, C(\varepsilon) \max\left\{ \|\rho\|_{W^{2,1}}, D(\varepsilon) \right\}. \tag{2}$$

In particular, for all $t \geq 0$

$$\left\| e^{t\mathcal{L}_{\text{qOU}}} - e^{t(\mathcal{L}_{\text{qOU}} + \varepsilon \mathcal{L}_G)} \right\|_\diamond^E \leq (1 + E)\varepsilon\, C(\varepsilon) \max\left\{ 1, D(\varepsilon) \right\}, \tag{3}$$

where $\|.\|_\diamond^E$ denotes the energy-constrained diamond norm defined in Equation (84).

We also note that our theory extends to the case of a time-dependent generator as well as to the multi-mode setting $\mathcal{H} = L^2(\mathbb{R}^m)$, $m \geq 1$, see Section 3.2 and 3.3. To prove our generation theorems, the compactly embedded Sobolev spaces play a crucial role and provide an interesting proof strategy, which follows the original method of Davies [20] by an explicit reduction to the seminal theorems by Hille, Yosida [35] and Feller, Myadera, Lumer and Phillips [23, Thm. II.3.8].

1.2 Dissipative CAT-qubit dynamics

As mentioned before, the interest in continuous variable QMS has been reignited by the modelling capabilities of a certain class of error-corrected universal quantum computing architectures. In [4] and later in [5], Azouit, Sarlette, and Rouchon prove the well-posedness of the dynamics that stabilises an l dimensional code-space, with a generator given for a fixed $\alpha \in \mathbb{R}$ by

$$\mathcal{L}_l(\rho) = L_l \rho L_l^\dagger - \frac{1}{2}\{L_l^\dagger L_l, \rho\} \qquad \text{with} \qquad L_l := a^l - \alpha^l \mathbb{1}\,. \tag{4}$$

In addition, they identified invariant operators of the dynamic and further showed that the semigroup exponentially drives states towards the code-space spanned by these invariants. By constructing a Banach space from composites of the generator, i.e. $L_l = a^l - \alpha^l \mathbb{1}$, compactly embedded in the self-adjoint trace class operators, they judiciously circumvented the problems previously encountered when trying to take limits of the minimal semigroups. This procedure was very much tailored towards the simple structure of the generator and also relied on a favourable commutation relation of $a^l - \alpha^l \mathbb{1}$ and $(a^l - \alpha^l \mathbb{1})^\dagger$, which one cannot hope for in general. In contrast, here we do not use parts of the generator to create our compactly embedded spaces, but instead use the most natural candidate at hand, namely the number operator N. Generalising the idea of Azouit et al. we take limits of sequences of semigroups for which our CV Sobolev spaces are admissible subspaces to prove our generation theorems. Combining their exponential dynamical convergence, stated as

$$\operatorname{tr}\Big[L_l\left(e^{t\mathcal{L}_l}(\rho)-\overline{\rho}\right)L_l^\dagger\Big]\leq e^{-l!t}\operatorname{tr}\Big[L_l|\rho-\overline{\rho}|L_l^\dagger\Big]\,,$$

where $\overline{\rho}$ is a ρ-dependent state in the code-space, with our generation and perturbation theory, we can, for example, show that any l-photon dissipation perturbed by a Hamiltonian admits the following large-time perturbation bounds (see Theorem 5.2):

Theorem *Let $\mathcal{L}_l$ be the l-photon dissipation defined in Equation (4) and $p_H\in\mathbb{C}[X,Y]$ with $\deg(p_H)=d_H\leq 2(l-1)$ such that $H=p_H(a,a^\dagger)$ is a symmetric operator. Then, there exist constants $c,\gamma>0$ depending on α and l such that for $\varepsilon\geq 0$ and all states $\rho\in W^{6l-4,1}$*

$$\Big|\operatorname{tr}\Big[L\left(e^{t\mathcal{L}_l}(\rho)-e^{t(\mathcal{L}_l+\varepsilon\mathcal{H}[H])}(\rho)\right)L^\dagger\Big]\Big|\leq\varepsilon c\left(1-e^{-l!t}\right)\max\{\gamma,\|\rho\|_{W^{6l-4,1}}\}$$

where $\mathcal{H}[H](\rho)\coloneqq -i[H,\rho]$.

The same idea can be extended to more general setups and thereby extends the result by Szehr and Wolf from finite dimensions to the case of strongly continuous semigroups over infinite dimensional systems.

1.3 Outline of the paper

In Section 2, we begin with an introduction to basic Banach space and operator theory, followed by a short overview of this theory in the context of Hilbert spaces and their associated bounded, compact and trace-class operator spaces. Building upon that we then introduce in Section 2.2 the notion of compact embeddings and weighted Banach spaces, followed by basic semigroup theory in Section 2.3. More specific to our application, we then briefly recapitulate Bosonic Hilbert spaces, and relevant operators thereon, and introduce our Bosonic Sobolev spaces. We prove that they are compactly embedded into one another and provide an interpolation theorem in the spirit of the Stein-Weiss theorem for weighted L_p spaces. In Section 3, we begin by showing the generation theorem in the time-independent case and then employ this theorem in Section 3.2 to prove a generation theorem for generators composed of polynomials in a and $a^\dagger$ with coefficients that are continuous functions of time. We extend our analysis to the multi-mode setting in Section 3.3 where we lift the generation theorems from the chapter before.

Section 4 begins with a short proposition making better use of tighter input-output moments of the generator and showing the existence of adherence points in the asymptotic time regime for semigroups that admit such bounds. We then proceed to prove the generation theorem for the quantum Ornstein Uhlenbeck generator as well as for a family of dissipative CAT-qubit dynamics in Section 4.2. This section is then followed by large time perturbation bounds for both the quantum Ornstein Uhlenbeck semigroup as well as the dissipative CAT-qubit dynamics in Section 5.

2 Preliminaries

We begin with a short review of valuable tools from Banach space theory in Section 2.1 and build upon them to prove a compact embedding theorem for a class of weighted spaces in Section 2.2. We then recall standard results from the theory of strongly continuous semigroups as well as evolution systems in Section 2.3. These will play an essential role in Section 3. Finally, we introduce continuous variable quantum systems, provide some valuable properties of polynomials of annihilation and creation operators, and introduce the notion of a *Sobolev preserving semigroup*, which are the main objects of study in the remainder of the paper.

2.1 Basic Banach space theory

We start with a brief recap on notions from the theory of Banach spaces that will be needed in this paper, and refer to [40, Chap. III, 19, Chap. IV, 64, Chap. 2-3, 35, Chap. 2-3] for more details. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a Banach space. We denote the space of bounded operators between two Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ by $\mathcal{B}(\mathcal{X},\mathcal{Y})$, with $\mathcal{B}(\mathcal{X},\mathcal{X}) = \mathcal{B}(\mathcal{X})$. The identity map in $\mathcal{B}(\mathcal{X})$ is denoted by $\mathbb{1}_{\mathcal{X}}$, or simply $\mathbb{1}$ when the underlying space is clear from the context. The operator norm is denoted by

$$\|A\|_{\mathcal{X}\to\mathcal{Y}} = \|A : \mathcal{X} \to \mathcal{Y}\| \coloneqq \sup_{\|x\|_{\mathcal{X}}=1} \|A(x)\|_{\mathcal{Y}}. \tag{5}$$

We recall that the linear space $\mathcal{B}(\mathcal{X},\mathcal{Y})$ equipped with the operator norm is a Banach space since $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is a Banach space. An operator $A : \mathcal{X} \to \mathcal{Y}$ is compact if the image sequence $\{Ax_n\}_{n\in\mathbb{N}} \subset \mathcal{Y}$ of any bounded sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ has a converging subsequence. In particular, every operator which can be approximated by a sequence of finite rank operators is compact [64, Thm. 3.1.9].

More generally, an unbounded operator A is a linear map $A : \mathcal{D}(A) \subset \mathcal{X} \to \mathcal{Y}$ defined on its domain $\mathcal{D}(A) \subset \mathcal{X}$. If the domain is dense in $\mathcal{X}$, the operator is said to be densely defined. In this paper, all unbounded operators are densely defined. Note that the addition and concatenation of two unbounded operators $(A, \mathcal{D}(A))$ and $(B, \mathcal{D}(B))$ is defined on $\mathcal{D}(A+B) = \mathcal{D}(A)\cap\mathcal{D}(B)$ and $\mathcal{D}(AB) = B^{-1}(\mathcal{D}(A))$ (cf. [40, Sec. III§5.1]). An operator $(A, \mathcal{D}(A))$ is closed iff its graph $\{(x, A(x)) : x \in \mathcal{D}(A)\}$ is a closed set in the product space $\mathcal{X} \times \mathcal{Y}$. A bounded operator is closed iff its domain is closed. By convention, we extend all densely defined and bounded operators by the bounded linear extension theorem to bounded operators on $\mathcal{X}$ [41, Thm. 2.7-11]. An operator is called closable if there exists a closed extension, where $\overline{A}$ is an extension of A if $\mathcal{D}(A) \subset \mathcal{D}(\overline{A})$ and $Ax = \overline{A}x$ for all $x \in \mathcal{D}(A)$. The closure of A is denoted by $\overline{A}$ [64, Sec. 7.1]. We also recall that for an unbounded operator $(A, \mathcal{D}(A))$ on $\mathcal{X}$, a core for A is a subset $\mathcal{D}_0 \subset \mathcal{D}(A)$ which is dense in $\mathcal{D}(A)$ w.r.t. the graph norm $\|\cdot\|_A \coloneqq \|A\cdot\|_{\mathcal{X}} + \|\cdot\|_{\mathcal{X}}$ of A (cf. [64, Def. 1.6]). Given two linear operators $(A, \mathcal{D}(A)))$ and $(B, \mathcal{D}(B))$ on $\mathcal{X}$, the operator $(B, \mathcal{D}(B))$ is relatively A-bounded if $\mathcal{D}(B) \subseteq \mathcal{D}(A)$ and there are $a, b \geq 0$ for all $x \in \mathcal{D}(\mathcal{L})$ such that

$$\|B(x)\|_{\mathcal{X}} \leq a\|A(x)\|_{\mathcal{X}} + b\|x\|_{\mathcal{X}}. \tag{6}$$

For a closed linear operator $(A, \mathcal{D}(A))$ on a Banach space $\mathcal{X}$ we call

$$\rho(A) \coloneqq \{\lambda \in \mathbb{C} : \lambda - A : \mathcal{D}(A) \to \mathcal{X} \text{ is bijective}\}$$

the resolvent set of $(A, \mathcal{D}(A))$. For $\lambda \in \rho(A)$ we call the inverse

$$R(\lambda, A) \coloneqq (\lambda - A)^{-1}$$

the resolvent, which is, by the closed graph theorem, a bounded operator on $\mathcal{X}$.

Besides the convergence w.r.t. the operator norm (i.e. uniform convergence), a sequence of operators $\{A_k\}_{k\in\mathbb{N}}$ defined on a common domain $\mathcal{D}(A)$ converges strongly if $\lim_{k\to\infty} \|A_k x - Ax\|_{\mathcal{Y}} = 0$ for all $x \in \mathcal{D}(A)$. Based on the underlying topologies associated with these two convergences, one can define the Bochner integral of vector and operator-valued maps on a compact interval equipped with the Lebesgue measure, e.g. $f : [a, b] \to \mathcal{X}$ and $F : [a, b] \to \mathcal{B}(\mathcal{X})$ with $a < b$. Under the assumption that the function f or F can be approximated by a step function and that the real-valued integral

$$\int_a^b \|f(s)\|_{\mathcal{X}}\, ds \qquad \text{or} \qquad \int_a^b \|F(s)\|_{\mathcal{X}\to\mathcal{X}}\, ds$$

is bounded, the Bochner integrals are defined by standard approximation with step functions. Since all the vector-valued maps considered in this work are continuous, the Bochner integral is always well-defined and coincides with the Riemann and Pettis integrals (more details can be found in [29]). Similar to the real-valued case, the integral satisfies the triangle inequality w.r.t. the norm, is invariant under closed linear transformations, and satisfies the fundamental theorem of calculus on continuous functions [35, Sec. 3.7-8].

Two special cases of Banach spaces that we will consider are Hilbert spaces and bounded operators defined on Hilbert spaces. We denote the latter by $\mathcal{B}(\mathcal{H})$ and use $\|\cdot\|_\infty$ for their norm. Given a separable Hilbert space $\mathcal{H}$ and $A \in \mathcal{B}(\mathcal{H})$, its adjoint $A^\dagger$ is uniquely defined by

$$\langle A\phi\,, \varphi\rangle = \langle \phi\,, A^\dagger \varphi\rangle \tag{7}$$

for all $\phi, \varphi \in \mathcal{H}$. The space of all bounded, self-adjoint operators, i.e. $A = A^\dagger$, is denoted by $\mathcal{B}_{\mathrm{sa}}(\mathcal{H})$. A special case of self-adjoint operators is those with finite support for a fixed orthonormal basis $\{|n\rangle\}_n$, whose set we denote by $\mathcal{T}_f \equiv \mathcal{T}_f(\mathcal{H}) \coloneqq \{A \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}) : \exists M \in \mathbb{N} : A = \sum_{n,m}^M a_{nm} |n\rangle\langle m|\}$. By a slight abuse of notations, we denote the formal adjoint of an unbounded operator $(A, \mathcal{D}(A))$ on $\mathcal{H}$ as $A^\dagger : \mathcal{D}(A^\dagger) \to \mathcal{H}$, where the latter satisfies Equation (7) for all $\phi \in \mathcal{D}(A)$ and for all φ in the maximally defined domain

$$\mathcal{D}(A^\dagger) \coloneqq \{|\varphi\rangle \in \mathcal{H} : |\phi\rangle \mapsto \langle A\phi\,, \varphi\rangle \text{ is bounded}\}.$$

The operator A is called symmetric if for all $|\phi\rangle\,, |\varphi\rangle \in \mathcal{D}(A)$, $\langle A\phi\,, \varphi\rangle = \langle \phi\,, A\varphi\rangle$. A is called self-adjoint if $\mathcal{D}(A) = \mathcal{D}(A^\dagger)$ and $A^\dagger = A$. An operator $(A, \mathcal{D}(A))$ is positive if $\langle A\phi, \phi\rangle \geq 0$ for all $\phi \in \mathcal{D}(A)$. In this case, we write $A \geq 0$. More generally we write $A \geq B$ if $A - B \geq 0$ with $\mathcal{D}(A) \subset \mathcal{D}(B)$. We defined the domain of the subtraction by $\mathcal{D}(A - B) = \mathcal{D}(A) \cap \mathcal{D}(B)$. The trace of a positive operator A is defined by $\mathrm{tr}[A] = \sum_{n\in\mathbb{N}} \langle n| A |n\rangle$. When $A \in \mathcal{B}(\mathcal{H})$, its trace-norm is defined by $\|A\|_1 \coloneqq \mathrm{tr}\,[|A|]$, where $|A| \coloneqq \sqrt{A^\dagger A}$ [64, Thm. 2.4.4]. Bounded operators with finite trace-norm are called trace-class and their class we denote by $\mathcal{T}_1 \equiv \mathcal{T}_1(\mathcal{H})$. We will most often consider the Banach space of self-adjoint trace-class operators denoted by $\mathcal{T}_{1,\mathrm{sa}} \equiv \mathcal{T}_{1,\mathrm{sa}}(\mathcal{H}) \coloneqq \{A \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}) : \|A\|_1 < \infty\}$. We define the set of density operators by $\mathcal{S} \equiv \mathcal{S}(\mathcal{H}) \coloneqq \{\rho : \rho \geq 0 \text{ and } \mathrm{tr}[\rho] = 1\}$.

2.2 Weighted norms and compact embeddings

For a Banach space $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ and an invertible operator $(\mathcal{W}, \mathcal{D}(\mathcal{W}))$ on $\mathcal{X}$, a natural way of defining a new norm out of $\|\cdot\|_\mathcal{X}$ is via the following procedure:

Definition 2.1 (Weighted normed space) Let $(\mathcal{W}, \mathcal{D}(\mathcal{W}))$ be an invertible linear operator. Then, $\|X\|_\mathcal{W} \coloneqq \|\mathcal{W}(X)\|_\mathcal{X}$ for $X \in \mathcal{D}(\mathcal{W})$ defines a norm on $\mathcal{D}(\mathcal{W})$. In the following, we denote by $\|P\|_{\mathcal{W}\to\mathcal{W}} \coloneqq \sup_{\|X\|_\mathcal{W}\leq 1} \|P(X)\|_\mathcal{W}$.

In the next lemma, we prove that completeness of the space $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ is preserved by the closedness of $(\mathcal{W}, \mathcal{D}(\mathcal{W}))$:

Lemma 2.2 *Let $(\mathcal{W}, \mathcal{D}(\mathcal{W}))$ be an invertible linear operator on the Banach space $\mathcal{X}$. Then, the weighted normed space $(\mathcal{D}(\mathcal{W}), \|\cdot\|_\mathcal{W})$ is a Banach space and the norm $\|\cdot\|_\mathcal{W}$ is equivalent to the graph norm of $\mathcal{W}$.*

Proof. First, it is clear that $\|\cdot\|_\mathcal{W}$ defines a norm because the linearity of $\mathcal{W}$ directly implies homogeneity and the triangle inequality, while the injectivity of $\mathcal{W}$ gives positive definiteness. By the closed graph theorem, we can conclude that $\mathcal{W}^{-1}$ is bounded, and therefore $\mathcal{W}$ is closed (see [40, Sec. III.2] and [35, Thm. 2.11.5]). Moreover,

$$\|X\|_\mathcal{W} = \|\mathcal{W}(X)\|_\mathcal{X} \leq \|X\|_\mathcal{X} + \|\mathcal{W}(X)\|_\mathcal{X} \leq (\|\mathcal{W}^{-1}\|_{\mathcal{X}\to\mathcal{X}} + 1)\|X\|_\mathcal{W}$$

shows that the graph norm of $\mathcal{W}$ is equivalent to $\|\cdot\|_{\mathcal{W}}$. By definition, the graph is a closed operator on the Banach space $\mathcal{X} \times \mathcal{X}$ such that the vector space $\mathcal{D}(\mathcal{X})$ equipped with the graph norm is a Banach space. The statement hence follows. □

Next, we introduce a compact embedding [24, Sec. 5.7] for weighted normed spaces, i.e. we want to reduce the compact embedding to a relation between the defining operators:

Definition 2.3 Let $(\mathcal{X}_1, \|\cdot\|_{\mathcal{X}_1})$ and $(\mathcal{X}_2, \|\cdot\|_{\mathcal{X}_2})$ be Banach spaces such that $\mathcal{X}_1 \subset \mathcal{X}_2$. We say that $\mathcal{X}_1$ is compactly embedded in $\mathcal{X}_2$, and denote this condition as $\mathcal{X}_1 \Subset \mathcal{X}_2$ iff

- $\exists c \geq 0$ such that $\|\cdot\|_{\mathcal{X}_2} \leq c\|\cdot\|_{\mathcal{X}_1}$ and
- any bounded sequence in $\mathcal{X}_1$ has a converging subsequence in $\mathcal{X}_2$ (i.e. is precompact).

Lemma 2.4 (Compact embedding) *Let $(\mathcal{W}_1, \mathcal{D}(\mathcal{W}_1))$ and $(\mathcal{W}_2, \mathcal{D}(\mathcal{W}_2))$ be invertible linear operators on $\mathcal{X}$ with bounded inverses and $\mathcal{D}(\mathcal{W}_1) \subset \mathcal{D}(\mathcal{W}_2)$. Then $(\mathcal{D}(\mathcal{W}_1), \|\cdot\|_{\mathcal{W}_1})$ is compactly embedded in $(\mathcal{D}(\mathcal{W}_2), \|\cdot\|_{\mathcal{W}_2})$ iff the extension of $\mathcal{W}_2\mathcal{W}_1^{-1}$ is a compact operator on $\mathcal{X}$.*

Proof. First, we prove that the first condition in Definition 2.3 is equivalent to the boundedness of $\mathcal{W}_2\mathcal{W}_1^{-1}$, which is defined on $\mathcal{X}$ by the bounded linear extension theorem [41, Thm. 2.7-11]. By assuming that $\mathcal{W}_2\mathcal{W}_1^{-1}$ is a bounded operator, which is implied by compactness,

$$\|X\|_{\mathcal{W}_2} = \|\mathcal{W}_2\mathcal{W}_1^{-1}\mathcal{W}_1(X)\|_{\mathcal{X}} \leq \|\mathcal{W}_2\mathcal{W}_1^{-1}\|_{\mathcal{X}\to\mathcal{X}}\|X\|_{\mathcal{W}_1}.$$

Now, if there is a $c \geq 0$ such that $\|X\|_{\mathcal{W}_2} \leq c\|X\|_{\mathcal{W}_1}$ then boundedness is given by

$$\|\mathcal{W}_2\mathcal{W}_1^{-1}(X)\|_{\mathcal{X}} \leq c\|\mathcal{W}_1\mathcal{W}_1^{-1}(X)\|_{\mathcal{X}} = c\|X\|_{\mathcal{X}}.$$

Let $\{X_k\}_{k\in\mathbb{N}} \subset \mathcal{X}$ be a bounded sequence and assume the second condition in Definition 2.3. Then, $\mathcal{W}_1^{-1}(X_k)$ is a bounded sequence in $(\mathcal{D}(\mathcal{W}_1), \|\cdot\|_{\mathcal{W}_1})$ which admits a converging subsequence in $(\mathcal{D}(\mathcal{W}_2), \|\cdot\|_{\mathcal{W}_2})$ by assumption. Therefore, $\mathcal{W}_2\mathcal{W}_1^{-1}(X_k)$ has a converging subsequence in $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$. Conversely, assume $\mathcal{W}_2\mathcal{W}_1^{-1}$ is compact and X_k a bounded sequence in $(\mathcal{D}(\mathcal{W}_1), \|\cdot\|_{\mathcal{W}_1})$. By definition $\mathcal{W}_1(X_k)$ is a bounded sequence in $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ so that $\mathcal{W}_2\mathcal{W}_1^{-1}(\mathcal{W}_1X_k) = \mathcal{W}_2(X_k)$ has a converging subsequence. □

Remark 1. A simple example of compact embedding is provided by classical Sobolev spaces $W^{k,p}(\mathbb{R})$, $k \in \mathbb{N}$ and $1 \leq p \leq \infty$, with the norm $\|f\|_{k,p} := \left(\sum_{i=0}^k \|f^{(i)}\|_p^p\right)^{1/p}$, where $\|f\|_p$ denotes the L^p norm of f with respect to the Lebesgue measure. Quantum extensions of these spaces recently appeared in [43, 6]. Here we will use the latter extension which we recall in Section 2.4.

2.3 Strongly continuous semigroups and evolution systems

The evolution of a quantum system is often described by a formal differential equation called the master equation. To rigorously study solutions of a master equation, the theory of C_0-semigroups constitutes an essential toolbox. While detailed expositions to this theory can be found e.g. in the books [23, Chap. II][40, Chap. 9] or [35, Chap. X-XIII], here we provide a short overview and introduce concepts that are relevant for the present paper. A family of operators $(P_t)_{t\geq 0} \subset \mathcal{B}(\mathcal{X})$ is called a C_0-semigroup if it satisfies the following properties:

- $P_tP_s = P_{t+s}$ for all $t, s \geq 0$;
- $P_0 = \mathbb{1}$, the identity map on $\mathcal{X}$; and

– $t \mapsto P_t$ is strongly continuous at 0.

To every C_0-semigroup one can associate a linear operator that is in the most general case unbounded, densely defined and closed. This operator determines the semigroup uniquely and is called its generator [23, Thm. II.1.4]. We will typically denote it by $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, where

$$\mathcal{D}(\mathcal{L}) = \{x \in \mathcal{X} : t \mapsto P_t(x) \text{ differentiable on } \mathbb{R}_+\} \tag{8}$$

and for $x \in \mathcal{D}(\mathcal{L})$

$$\mathcal{L}(x) = \lim_{t \to 0^+} \frac{1}{t}(P_t(x) - x)\,, \tag{9}$$

where the limit is with respect to the topology induced by $\mathcal{X}$. The semigroup leaves the domain of its generator invariant and further commutes with it on its domain allowing us to justify the well-posedness of the following differential equation on $\mathcal{X}$

$$\frac{d}{dt}x(t) = \mathcal{L}(x(t)) \quad x(0) \in \mathcal{D}(\mathcal{L}) \quad \text{and} \quad t \geq 0\,. \tag{10}$$

From the above considerations, this equation has a strongly continuous solution given by the semigroup, i.e. $P_t(x(0)) = x(t)$. Indeed the semigroup is the unique solution (asking for a continuously differentiable map $t \mapsto x(t)$) to this so-called master equation [23, Prop. II.6.2]. The formulation as a master equation or initial value problem also reveals the origin of the term "generator", since for bounded linear operators the solution to these problems is just given by the semigroup involving the exponential of the generator (i.e. $(e^{t\mathcal{L}})_{t\geq 0}$). When the operator $\mathcal{L}$ is bounded, the conditions of existence and uniqueness are immediately satisfied using the series expansion of the exponential. For unbounded operators, the answer is no longer straightforward and requires a different representation of the exponential. One possible choice involves the resolvent of the generator.

The well-known generation theorems by Lumer and Phillips, Hille and Yosida and Feller, Miyadera and Phillips all rely on the resolvent satisfying specific bounds, either directly, or indirectly e.g. by dissipativity of the underlying operator $\mathcal{L}$. Below we recall the theorems by Hille and Yosida and Lumer and Phillips, as they are going to be used frequently throughout this paper. It is noteworthy that the first two theorems are generalized by the third with the last allowing for the generation of semigroups that satisfy the bound

$$\left\| e^{t\mathcal{L}} \right\|_{\mathcal{X}\to\mathcal{X}} \leq c\, e^{\omega t}$$

with $\omega \in \mathbb{R}$ and $c \geq 0$. If $c = 1$, we call the semigroup ω-quasi contractive, and if further $\omega \leq 0$ we call it contractive. We start with the generation theorem by Hille and Yosida that gives necessary and sufficient conditions on an operator to be the generator of a contractive C_0-semigroups by imposing constraints on its resolvent.

Theorem 2.5 (Hille-Yosida) *A linear operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ on $\mathcal{X}$ generates a strongly continuous ω-quasi contraction semigroup iff $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is closed, densely defined, the resolvent set contains (ω, ∞) and for all $\lambda \in (\omega, \infty)$ one has*

$$\|R(\lambda, \mathcal{L})\|_{\mathcal{X}\to\mathcal{X}} \leq \frac{1}{\lambda - \omega}\,.$$

The other seminal result, which we use in the present paper is a modified formulation of Theorem 2.5 due to Lumer and Phillips [23, Thm. II.3.15] which, instead of asking for a certain bound on the resolvent, requires certain properties for $\mathcal{L}$ among which ω-dissipativity:

Definition 2.6 For $\omega \geq 0$, an operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ on $\mathcal{X}$ is ω-quasi dissipative if for all $x \in \mathcal{D}(\mathcal{L})$ and $\lambda > 0$

$$\|(\lambda - (\mathcal{L} - \omega))x\|_{\mathcal{X}} \geq \lambda \|x\|_{\mathcal{X}}.$$

If $\omega = 0$, we call the operator dissipative.

In what follows, the notation rg stands for the range of an operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, defined as $\mathrm{rg}(\mathcal{L}) = \{\mathcal{L}(x) : x \in \mathcal{D}(\mathcal{L})\}$. Then we can state the theorem in the following way:

Theorem 2.7 (Lumer-Phillips) *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ be a densely defined, ω-dissipative operator on $\mathcal{X}$. Then, the closure $\overline{\mathcal{L}}$ generates a ω-quasi contraction semigroup iff there exists a $\lambda > 0$ such that $\mathrm{rg}(\lambda - (\mathcal{L} - \omega))$ is dense in $\mathcal{X}$.*

Remark 2. When $\mathcal{L}$ is dissipative and there is a $\lambda > 0$ such that $\mathrm{rg}(\lambda - \mathcal{L})$ is dense in $\mathcal{X}$, then this holds for all $\lambda > 0$ [23, Prop. II.3.14].

In the case of a C_0-semigroup on a Hilbert space $\mathcal{H}$, the following result proves useful:

Proposition 2.8 *Let $(G, \mathcal{D}(G))$ be a densely defined linear operator on $\mathcal{H}$ and assume that G and $G^\dagger$ are ω-quasi dissipative. Then, $\overline{G}$ generates a ω-quasi contraction C_0-semigroup on $\mathcal{H}$. Moreover, if $(G, \mathcal{D}(G))$ generates a ω-quasi contraction semigroup, G and $G^\dagger$ are ω-quasi dissipative.*

In this paper, we are also concerned with extensions of the above theory to the setting of time-dependent C_0-semigroups. In this case, we refer to the family as a C_0-*evolution system*: A two-parameter family of bounded operators $(P_{t,s})_{0 \leq s \leq t}$ is called an *evolution system* if

- $P_{t,t} = \mathbb{1}$,
- $P_{t,r} P_{r,s} = P_{t,s}$ for all $0 \leq s \leq r \leq t$, and
- $(t,s) \mapsto P_{t,s}$ is strongly continuous.

A well-known class of evolution systems is given by C_0-semigroup after imposing $P_{t,s} = P_{t-s}$. A subtle difference between semigroups and evolution systems is that the latter are not necessarily differentiable for any $x \neq 0$ [23, p. 478]. Here, we recall sufficient conditions under which the following master equation admits a unique *solution operator*:

$$\frac{\partial}{\partial t} x(t) = \mathcal{L}_t(x(t)) \quad \text{and} \quad x(s) = x_s \quad \text{for } 0 \leq s \leq t\,. \tag{11}$$

We start with a set of assumptions often referred to as being of *hyperbolic type* [54, Chap. 5, 56, pp. 127 ff.] and which allow for the generation of an evolution system starting from a C_0-semigroup. Here, the existence of a so-called admissible subspace plays an important role:

Definition 2.9 (Admissible subspaces) For a C_0-semigroup $(P_t)_{t \geq 0} \subset \mathcal{B}(\mathcal{X})$ with generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, a subspace $(\mathcal{Y} \subset \mathcal{X}, \|\cdot\|_{\mathcal{Y}})$ is called admissible for $(P_t)_{t \geq 0}$, or simply $\mathcal{L}$-admissible if $\mathcal{Y}$ is an invariant closed subspace of the semigroup, i.e. $e^{t\mathcal{L}}\mathcal{Y} \subset \mathcal{Y}$, and $e^{t\mathcal{L}}|_{\mathcal{Y}}$ defines a C_0-semigroup on $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$. Similarly, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is an admissible subspace of an evolution system $P_{0 \leq s \leq t}$ if it is an invariant closed subspace of the evolution system and $(P_{0 \leq s \leq t}|_{\mathcal{Y}})_{0 \leq s \leq t}$ defines an evolution system on $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$.

Our first basic assumption is that for every fixed s the operator $(\mathcal{L}_s, \mathcal{D}(\mathcal{L}_s))$ generates a C_0-semigroup. Moreover,

(1) $(\mathcal{L}_s)_{s \geq 0}$ is a *stable* family, i.e. there is $c \geq 0$ and $\omega \in \mathbb{R}$ such that $\|e^{t\mathcal{L}_s}\|_{\mathcal{X} \to \mathcal{X}} \leq c\, e^{\omega t}$ for all $s \geq 0$;

(2) There exists a subspace $\mathcal{Y} \subset \mathcal{X}$ and a norm $\|\cdot\|_{\mathcal{Y}}$ on $\mathcal{Y}$ endowing $\mathcal{Y}$ with a Banach space structure, such that for all $s \geq 0$, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is $\mathcal{L}_s$-*admissible* and $(\mathcal{L}_s)_{s\geq 0}$ is stable on $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$;

(3) Finally, the map $s \mapsto \mathcal{L}_s \in \mathcal{B}((\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}), (\mathcal{X}, \|.\|_{\mathcal{X}}))$ is uniformly continuous.

Under these assumptions, [54, Thm. 3.1] shows the existence of a unique evolution system $(P_{t,s})_{0\leq s\leq t}$. If one further requests this evolution system to have the following properties

(4) $P_{t,s}\mathcal{Y} \subseteq \mathcal{Y}$ for $0 \leq s \leq t$;

(5) $(s,t) \mapsto P_{s,t}$ is strongly continuous on $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$;

one obtains the following theorem:

Theorem 2.10 (Time-dependent semigroups [54, Thm. 3.1, 4.3]) *Let $\{(\mathcal{L}_s, \mathcal{D}(\mathcal{L}_s)\}_{s\geq 0}$ be a family of generators of C_0-semigroups, which satisfy assumption $(1-3)$. Then, there exists a unique evolution system which satisfies*

- $\|P_{t,s}\|_{\mathcal{X}\to\mathcal{X}} \leq c\, e^{(t-s)\omega}$ *for all* $0 \leq s \leq t$;
- $\lim_{t\downarrow s} \frac{1}{t-s}(P_{t,s}x - x) = \mathcal{L}_s x$ *for all* $x \in \mathcal{Y}$; *and*
- $\frac{\partial}{\partial s} P_{t,s}x = -P_{t,s}\mathcal{L}_s x$ *for all* $x \in \mathcal{Y}$ *and* $0 \leq s \leq t$.

The two limits above are both with respect to the topology induced by $\|.\|_{\mathcal{X}}$. If further (4) and (5) hold then for every $v \in \mathcal{Y}$, $P_{t,s}v$ is a unique solution for the initial value problem in Equation (11) in $(\mathcal{Y}, \|.\|_{\mathcal{X}})$.

Remark (Kato's C^1-condition). For a time-independent domain $\mathcal{D}$, the above conditions follow if $\{(\mathcal{L}_s, \mathcal{D})\}_{s\geq 0}$ is a stable family of generators and if $s \mapsto \mathcal{L}_s$ is strongly continuously differentiable w.r.t. $\|.\|_{\mathcal{X}}$ [54, Chap. 5, Thm. 4.8].

Finally, we discuss the general idea of how to control perturbed semigroups through certain admissible subsets associated with the domain of an invertible operator. More explicit variants are postponed to Section 5.

Theorem 2.11 *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ and $(\mathcal{L}+\mathcal{K}, \mathcal{D}(\mathcal{L}+\mathcal{K}))$ be two generators of C_0-semigroups on $\mathcal{X}$, for an operator $(\mathcal{K}, \mathcal{D}(\mathcal{K}))$. Moreover, let $(\mathcal{W}, \mathcal{D}(\mathcal{W}))$ be an invertible operator on $\mathcal{X}$ with bounded inverse, such that $\mathcal{D}(\mathcal{W})$ is an $\mathcal{L}+\mathcal{K}$-admissible subspace (see Definition 2.9) and such that $\mathcal{K}\mathcal{W}^{-1}$ is bounded. Then, for all $t \geq 0$,*

$$\|e^{t\mathcal{L}} - e^{t(\mathcal{L}+\mathcal{K})} : \mathcal{W} \to \mathcal{X}\| \leq t\, \|\mathcal{K}\mathcal{W}^{-1}\|_{\mathcal{X}\to\mathcal{X}} \int_0^1 \|e^{(1-s)t\mathcal{L}}\|_{\mathcal{X}\to\mathcal{X}}\, \|e^{st(\mathcal{L}+\mathcal{K})}\|_{\mathcal{W}\to\mathcal{W}}\, ds\,.$$

In particular, for all $t \geq 0$ and $x \in \mathcal{D}(\mathcal{W})$, the following equation holds in the Bochner sense:

$$(e^{t\mathcal{L}} - e^{t(\mathcal{L}+\mathcal{K})})x = t \int_0^1 e^{(1-s)t\mathcal{L}}\mathcal{K}e^{st(\mathcal{L}+\mathcal{K})}x\, ds.$$

Proof. See Theorem A.1. □

Remark 3. In words, Theorem 2.11 shows that the integral equation for semigroups is well-defined by generalizing the standard method that requires the following relative boundedness condition (see e.g. [40, Chapter 2]): for all $x \in \mathcal{D}(\mathcal{L}+\mathcal{K})$, $x \in \mathcal{D}(\mathcal{K})$ and

$$\|\mathcal{K}x\|_{\mathcal{X}} \leq \|(\mathcal{L}+\mathcal{K})x\|_{\mathcal{X}}.$$

Indeed, the choice $\mathcal{W} \coloneqq \mathbb{1} - (\mathcal{L}+\mathcal{K})$ with the resolvent $R(1, \mathcal{L}+\mathcal{K})$ as its bounded inverse shows that our scheme is a generalization of the above. Clearly, $\mathcal{W}$ generates an admissible subspace and $\|\mathcal{K}\mathcal{W}^{-1}x\|_{\mathcal{X}} \leq \|x\|_{\mathcal{X}}$ shows the implication. Note also that the above bound can be extended to evolution systems,

2.4 Continuous variable quantum systems

An important feature in quantum physics is that of the indistinguishability of particles which results in the introduction of Bosonic and Fermionic particles [8, Chap. 5.2]. In the second quantization formalism, a Bosonic or continuous variable quantum system can be described by the Fock space $\mathcal{H} = L^2(\mathbb{R})$ endowed with an orthonormal (Fock) basis $\{|n\rangle\}_{n=0}^{\infty}$, where n labels the number of photons present in a given mode. The space of vectors with finite support is denoted by $\mathcal{H}_f = \{|\psi\rangle \in \mathcal{H} : \exists M \in \mathbb{N} : |\psi\rangle = \sum_{n=0}^{M} \langle n\,,\psi\rangle\,|n\rangle\}$, where $\langle \varphi, \psi\rangle$ denotes the standard inner product on $L^2(\mathbb{R})$. Next, we define the *annihilation* and *creation* operators through the following relations

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \quad a\,|0\rangle = 0, \quad a^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\ .$$

The operators a and $a^\dagger$ satisfy the *canonical commutation relation* (CCR), i.e. $[a, a^\dagger] = \mathbb{1}$ on $\mathcal{H}_f$. We can construct the number operator of the latter two as,

$$N = a^\dagger a = \sum_{n=0}^{\infty} n\,|n\rangle\langle n|\ . \tag{12}$$

It counts the number of photons in a mode. All of a, $a^\dagger$, and N, although linear, are unbounded operators, hence are only defined on a (dense) subset of $\mathcal{H}$, namely

$$\mathcal{D}(a^\dagger) = \mathcal{D}(a) = \{|\phi\rangle \in \mathcal{H} : \|a\,|\phi\rangle\,\| < \infty\} = \{|\phi\rangle = \sum_{n=0}^{\infty} \lambda_n\,|n\rangle : \sum_{n=0}^{\infty} n|\lambda_n|^2 < \infty\} = \mathcal{D}(N^{\frac{1}{2}})\,. \tag{13}$$

Note that the above domains are maximal, i.e. $\mathcal{D}(a) = \{|\phi\rangle \in \mathcal{H} : a\,|\phi\rangle \in \mathcal{H}\}$. In most parts of the paper, we consider operators constructed by polynomials $p \in \mathbb{C}[X, Y]$ of a, $a^\dagger$ where the variables X and Y are considered non-commuting, i.e. XY is a different polynomial then YX. Note the $\mathbb{C}[X, Y]$ is the polynomial ring in X, Y over the complex field. Using the CCR, we can always assume without loss of generality that the polynomial has the following normal form: there exist complex coefficients λ_{ij} and μ_{kl} such that

$$p(a\,, a^\dagger) = \sum_{i+2j \leq \deg(p)} \lambda_{ij}(a^\dagger)^i N^j + \sum_{k+2l \leq \deg(p)} \mu_{kl} N^l a^k\,. \tag{14}$$

One possible domain of these operators can be described by the degree d of p (see Section B):

$$\mathcal{D}(p(a, a^\dagger)) = \mathcal{D}(N^{d/2}). \tag{15}$$

Next, we add the number operator to the power of twice the leading order to the polynomial, i.e.

$$\tilde{p}(a, a^\dagger) := (N + \mathbb{1})^{2d} + p(a, a^\dagger) \tag{16}$$

which allows us to show that the domain is maximal in the sense that the operator is closed. Note that he choice of $(N + \mathbb{1})^{2d}$ is adapted to the proof of Lemma 3.2. Moreover, we prove that $\mathcal{H}_f$ is a core for the considered polynomial (cf. [64, Sec. 7.1]).

Lemma 2.12 (Adjoint and core of polynomials of $a, a^\dagger$) *Let $p \in \mathbb{C}[X, Y]$ be a polynomial on $\mathbb{C}$ and $(p(a, a^\dagger), \mathcal{D}(N^{d/2}))$ the unbounded operator in normal form (15). Then, $p(a, a^\dagger)$ is closable and there is a $c \geq 0$ such that for all $\phi \in \mathcal{D}(N^{d/2})$*

$$\|p(a, a^\dagger)\,|\phi\rangle\,\| \leq c\|(\mathbb{1} + N)^{d/2}\,|\phi\rangle\,\|\,.$$

The modification $\tilde{p}(a, a^\dagger) = (N+\mathbb{1})^{2d} + p(a, a^\dagger)$ is a closed operator with domain $\mathcal{D}(\tilde{p}(a, a^\dagger)) = \mathcal{D}(\tilde{p}(a, a^\dagger)^\dagger) = \mathcal{D}(N^{2d})$ and core $\mathcal{H}_f$.

Proof. See Lemma B.1. □

Remark. The above lemma will allow us to reduce the analysis of the unbounded operator $p(a, a^\dagger)$ in the strong topology to that on finite sums.

We end this preliminary section by introducing a family of weighted Banach spaces which we coin as *Bosonic Sobolev spaces* in analogy with classical harmonic analysis. The Bosonic Sobolev space of order $k \in \mathbb{R}_+$ is defined on

$$\mathcal{D}(\mathcal{W}^k) = \{(\mathcal{W}^k)^{-1}(x) \in \mathcal{T}_{1,\mathrm{sa}} \ : \ x \in \mathcal{T}_{1,\mathrm{sa}}\}$$

via

$$\mathcal{W}^k(x) \coloneqq (\mathbb{1} + N)^{k/4} x (\mathbb{1} + N)^{k/4} \,. \tag{17}$$

Since the inverse $(\mathcal{W}^k)^{-1}(x) = (\mathbb{1}+N)^{-k/4} x (\mathbb{1}+N)^{-k/4}$ is a bounded operator, $(\mathcal{D}(\mathcal{W}^k), \|\cdot\|_{\mathcal{W}^k})$ is a Banach space by Theorem 2.2. For the sake of notation, we define

$$W^{k,1} \coloneqq \mathcal{D}(\mathcal{W}^k) \quad \text{and} \quad \|\cdot\|_{W^{k,1}} \coloneqq \|\cdot\|_{\mathcal{W}^k}. \tag{18}$$

For $k = 0$, $\mathcal{W}^k = \mathbb{1}$ and $(\mathcal{D}(\mathcal{W}^0), \|\cdot\|_{W^{0,1}}) = (\mathcal{T}_{1,\mathrm{sa}}, \|\cdot\|_1)$.

Lemma 2.13 *Let $k < k' \in \mathbb{R}_+ \coloneqq [0, \infty)$. Then,*

$$W^{k',1} \Subset W^{k,1}.$$

Proof. The proof relies on the abstract Theorem 2.4, which is applied for different values $k \in \mathbb{R}_+$ of

$$\mathcal{W}^k(x) \coloneqq (\mathbb{1} + N)^{k/4} x (\mathbb{1} + N)^{k/4}.$$

For $k' > k$, the operator $\mathcal{W}^k \mathcal{W}^{-k'} = \mathcal{W}^{k-k'}$ is bounded. Next we show compactness by proving that $\mathcal{W}^{-l}$ with $-l = k - k'$ is approximated by a sequence of finite rank operators:

$$\mathcal{W}^{-l}_{f,M}(x) \coloneqq \sum_{n}^{M} (1+n)^{-l/4} \, |n\rangle\langle n| \, x \sum_{m}^{M} (1+m)^{-l/4} \, |m\rangle\langle m| \,,$$

which can be deduced through Hölder's inequality

$$\|\mathcal{W}^{-l}(x) - \mathcal{W}^{-l}_{f,M}(x)\|_1 = \| \sum_{m,n > M} (1+n)^{-l/4} \, |n\rangle\langle n| \, x (1+m)^{-l/4} \, |m\rangle\langle m| \, \|_1 \leq M^{-l/2} \|x\|_1.$$

Since finite rank operators are compact by the Bolzano-Weierstrass theorem and the limit is a compact operator again [35, Thm. 2.13.4], the operator $\mathcal{W}^{-l}$ is a compact operator on $\mathcal{T}_{1,\mathrm{sa}}$. Applying Theorem 2.4 shows that

$$W^{k',1} \Subset W^{k,1} \,.$$

□

The following theorem will become helpful later and has an analogue in the theory of commutative L_p spaces, which inspired its name. Although it can be proved by interpolation theory, we provide a more rudimentary approach using only Hadamard's three-line theorem.

Theorem 2.14 (Stein-Weiss theorem for Bosonic Sobolev spaces) *Let $k_0 < k_1 \in \mathbb{R}_+$ and $T : W^{k_j,1} \to W^{k_j,1}$, be a linear map with $\|T\|_{W^{k_j,1} \to W^{k_j,1}} \leq M_j$ for some $M_j \geq 0$, $j = 1, 2$. Then for $\theta \in [0, 1]$, $T : W^{k_\theta,1} \to W^{k_\theta,1}$ with $k_\theta = (1-\theta)k_0 + \theta k_1$ obtained by restriction of the input $T : W^{k_0,1} \to W^{k_0,1}$ to $W^{k_\theta,1} \cap W^{k_0,1}$ is a well defined bounded linear map with*

$$\|T\|_{W^{k_\theta,1} \to W^{k_\theta,1}} \leq M_0^{1-\theta} M_1^{\theta} \,. \tag{19}$$

Proof. We have that $k_0 < k_1$ and hence $W^{k_1,1} \Subset W^{k_0,1}$. Let $k_\theta = (1-\theta)k_0 + \theta k_1$ with $\theta \in (0,1)$ and $x \in \mathcal{T}_f \subseteq W^{k_1,1} \Subset W^{k_\theta,1} \Subset W^{k_0,1}$. We will show

$$\|T(x)\|_{W^{k_\theta,1}} \leq M_0^{1-\theta} M_1^\theta \|x\|_{W^{k_\theta,1}} \tag{20}$$

which proves that T can be uniquely extended to a bounded linear map on $W^{k_\theta,1}$ that agrees with the restriction of $T : W^{k_0,1} \to W^{k_0,1}$, to $W^{k_\theta,1} \cap W^{k_0,1}$. This agreement on intersections is due to the compact embeddings of the Bosonic Sobolev spaces into one another. For $x \in \mathcal{T}_f$ and $z \in S \coloneqq \{z \in \mathbb{C} \;:\; 0 \leq \mathrm{Re}(z) \leq 1\}$, we define $k(z) = (1-z)k_0 + zk_1$ so that for a fixed $Z \in \mathcal{B}(\mathcal{H})$ with $\|Z\|_\infty \leq 1$,

$$g : S \to \mathbb{C},\; g(z) = \mathrm{tr}\Big[(N+\mathbb{1})^{\frac{k(z)}{4}} T\Big((N+\mathbb{1})^{\frac{k_\theta - k(z)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k(z)}{4}}\Big)(N+\mathbb{1})^{\frac{k(z)}{4}} Z\Big] \tag{21}$$

is well-defined, uniformly bounded, and continuous on S (q.v. Lemma D.1) and further holomorphic on $\mathring{S} \coloneqq \{z \in \mathbb{C} \;:\; 0 < \mathrm{Re}(z) < 1\}$ (q.v. Lemma D.2). Note that for $\theta \in (0,1)$

$$|g(\theta)| = \left|\mathrm{tr}\Big[(N+\mathbb{1})^{\frac{k_\theta}{4}} T(x)(N+\mathbb{1})^{\frac{k_\theta}{4}} Z\Big]\right|, \tag{22}$$

for $t \in \mathbb{R}$

$$\begin{aligned}
|g(it)| &= \left|\mathrm{tr}\left[(N+\mathbb{1})^{\frac{k_0 + it(k_1-k_0)}{4}} T\Big((N+\mathbb{1})^{\frac{k_\theta - k(it)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k(it)}{4}}\Big)(N+\mathbb{1})^{\frac{k_0+it(k_1-k_0)}{4}} Z\right]\right| \\
&\leq \left\| T\Big((N+\mathbb{1})^{\frac{k_\theta - k_0 + it(k_0-k_1)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k_0 + it(k_0-k_1)}{4}}\Big)\right\|_{W^{k_0,1}} \|Z\|_\infty \\
&\leq \|T\|_{W^{k_0,1} \to W^{k_0,1}} \left\|(N+\mathbb{1})^{\frac{k_\theta - k_0}{4}} x (N+\mathbb{1})^{\frac{k_\theta-k_0}{4}}\right\|_{W^{k_0,1}} \\
&= M_0 \|x\|_{W^{k_\theta,1}},
\end{aligned} \tag{23}$$

and similarly

$$\begin{aligned}
|g(1+it)| &\leq \left\| T\Big((N+\mathbb{1})^{\frac{k_\theta - k_1 + it(k_0-k_1)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k_1 + it(k_0-k_1)}{4}}\Big)\right\|_{W^{k_1,1}} \|Z\|_\infty \\
&\leq \|T\|_{W^{k_1,1} \to W^{k_1,1}} \left\|(N+\mathbb{1})^{\frac{k_\theta - k_1 + it(k_0-k_1)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k_1 + it(k_0-k_1)}{4}}\right\|_{W^{k_1,1}} \\
&= M_1 \|x\|_{W^{k_\theta,1}}.
\end{aligned} \tag{24}$$

An application of Hadamard's three-lines theorem now gives us that for $G_0 \coloneqq \sup_{t\in\mathbb{R}} |g(it)|$ and $G_1 \coloneqq \sup_{t \in \mathbb{R}} |g(it+1)|$

$$|g(\theta)| \leq G_0^{1-\theta} G_1^\theta \leq M_0^{1-\theta} M_1^\theta \|x\|_{W^{k_\theta,1}}, \tag{25}$$

where the last inequality follows from the bounds in Equation (23) and Equation (24). Since Z was arbitrary, we can deduce that

$$\begin{aligned}
\|T(x)\|_{W^{k_\theta,1}} &= \sup\left\{\left|\mathrm{tr}\Big[(N+\mathbb{1})^{\frac{k_\theta}{4}} T(x)(N+\mathbb{1})^{\frac{k_\theta}{4}} Z\Big]\right| \;:\; \|Z\|_\infty \leq 1\right\} \\
&\leq M_0^{1-\theta} M_1^\theta \|x\|_{W^{k_\theta,1}}
\end{aligned} \tag{26}$$

where we used the dual characterisation of $\|\cdot\|_1$. This concludes the claim. ☐

With the Sobolev embedding at hand, we introduce the notion of a *Sobolev preserving semigroup* as a semigroup defined on a sequence of Bosonic Sobolev spaces.

Definition 2.15 (Sobolev preserving evolution system) Let $(\mathcal{P}_t)_{t\geq 0}$ be a C_0-semigroup on $\mathcal{T}_{1,\mathrm{sa}}$. We then call $(\mathcal{P}_t)_{t\geq 0}$ *Sobolev preserving* if there exists a divergent sequence $\{k_r\}_{r\in\mathbb{N}} \to \infty$, such that for all $r \in \mathbb{N}$, $W^{k_r,1}$ is an admissible subspace for $(\mathcal{P}_t)_{t\geq 0}$. Similarly for $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ an evolution system on $\mathcal{T}_{1,\mathrm{sa}}$, we call $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ *Sobolev preserving* if for all $r \in \mathbb{N}$, $W^{k_r,1}$ is admissible for $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ to $W^{k_r,1}$.

Note that with the Stein-Weiss theorem for Bosonic Sobolev spaces, Theorem 2.14, one can immediately interpolate a semigroup and evolution system defined on $W^{k_0,1}$ and $W^{k_1,1}$ to $W^{k_\theta,1}$ with $k_\theta = (1-\theta)k_0 + \theta k_1$, $\theta \in [0,1]$ (q.v. Lemma E.4). This means the above definition is equivalent to the definition, requiring that for all $k \in \mathbb{R}_+$, $W^{k,1}$ are admissible subspaces of the semigroup or evolution system respectively.

The following example shows that not every semigroup is Sobolev preserving:

Example 4. An example of a C_0-semigroup that is not Sobolev preserving is the depolarizing semigroup, i.e. $\mathcal{P}_t(\rho) = e^{-t}\rho + (1-e^{-t})\operatorname{tr}[\rho]\sigma$ where σ is a quantum state with $\operatorname{tr}\left[N^{\frac{1}{2}}\sigma N^{\frac{1}{2}}\right] = \infty$. Then for a quantum state $\rho \in W^{2,1}$ we find that

$$\|\mathcal{P}_t(\rho)\|_{W^{2,1}} = \infty \qquad \forall t > 0\,.$$

3 Sobolev preserving quantum Markov semigroups

A quantum evolution in bosonic systems is described by a master equation

$$\frac{d}{dt}x(t) = \mathcal{L}(x(t)) \quad x(0) \in \mathcal{D}(\mathcal{L}) \quad \text{and} \quad t \geq 0\,. \tag{27}$$

where $\mathcal{L}$ is potentially unbounded. In the following, we state two sufficient assumptions for the existence and uniqueness of an operator-valued solution to (27) in terms of a semigroup. In other words, we prove a generation theorem for bosonic quantum Markov semigroups. This is generalized in Section 3.2 to the case of time-dependent generators.

3.1 Strongly continuous bosonic semigroups

We start with the time-independent setting, for which we need two working assumptions. The first assumption is motivated by the so-called GKSL [46, 30] form that generators of quantum dynamical semigroups over finite-dimensional quantum systems take, as well as our natural choice to consider jump and Hamiltonian operators described by polynomials in the annihilation and creation operators:

Assumption 1 *The operator* $(\mathcal{L}, \mathcal{T}_f)$ *has* GKSL *form, i.e. for* $x \in \mathcal{T}_f$

$$\begin{aligned}\mathcal{L}: \mathcal{T}_f \to \mathcal{T}_f \quad x \mapsto \mathcal{L}(x) &= -i[H,x] + \sum_{j=1}^{K} L_j x L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, x\} \\ &\coloneqq Gx + xG^\dagger + \sum_{j=1}^{K} L_j x L_j^\dagger\,,\end{aligned} \tag{28}$$

for some $K \in \mathbb{N}$ *and with* $G = -iH - \frac{1}{2}\sum_{j=1}^{K} L_j^\dagger L_j$*, where* $\{A,B\} = AB + BA$ *denotes the anticommutator of two operators* A, B *on a suitable domain. For the above equation to make sense, the operators* H *and* L_j *are assumed to be polynomials of the creation and annihilation operators, i.e.* $H \coloneqq p_H(a,a^\dagger)$ *and* $L_j \coloneqq p_j(a,a^\dagger)$*, and* H *is assumed to be symmetric. This ensures that* $\mathcal{T}_f$ *is invariant under* $\mathcal{L}$*. We denote the degree of* p_H *by* $d_H \coloneqq \deg(p_H)$*, those of* p_j *by* $d_j \coloneqq \deg(p_j)$*, and* $d \coloneqq \max\{d_1, ..., d_K, d_H\}$*.*

The second assumption will lead to the semigroup being Sobolev preserving, which allows us not only to prove the existence and uniqueness of the evolution generated by (27) but further to conduct a perturbation analysis as well as to extend our results to the case of a time-dependent Master equation:

Assumption 2 *There exists a non-negative sequence $\{k_r\}_{r\in\mathbb{N}} \to \infty$ s.t. for all $r \in \mathbb{N}$ there exist $\omega_{k_r} \geq 0$ such that for all positive semi-definite $x \in \mathcal{T}_f$*

$$\operatorname{tr}\Big[\mathcal{L}(x)(N+\mathbb{1})^{k_r/2}\Big] \leq \omega_{k_r} \operatorname{tr}\Big[x(N+\mathbb{1})^{k_r/2}\Big]\,. \tag{29}$$

We are now ready to state and prove the main theorem of the section:

Theorem 3.1 (Generation of bosonic semigroups) *Let $(\mathcal{L},\mathcal{D}(\mathcal{L}))$ be an operator defined on $\mathcal{T}_{1,\mathrm{sa}}$. If $(\mathcal{L},\mathcal{D}(\mathcal{L}))$ satisfies Assumption 1 and Assumption 2, then the closure $\overline{\mathcal{L}}$ generates a strongly continuous, positivity preserving semigroup $(\mathcal{P}_t)_{t\geq 0}$ on $W^{k,1}$ for all $k \geq 0$ with*

$$\|\mathcal{P}_t\|_{W^{k,1}\to W^{k,1}} \leq e^{\omega_k t} \quad \forall t \geq 0\,. \tag{30}$$

where $\omega_k = \frac{k_{r_1}-k}{k_{r_1}-k_{r_0}}\omega_{k_{r_0}} + \frac{k-k_{r_0}}{k_{r_1}-k_{r_0}}\omega_{k_{r_1}}$ for an r such that $k_{r_0} \leq k < k_{r_1}$. Finally, for $k=0$, the semigroup is contractive and trace-preserving.

Remark 5. For the existence and well-posedness of a semigroup between k and 0, Assumption 2 with $\tilde{k} \geq k+4d$ would be sufficient. However, in all of the examples we found, Assumption 2 was either fully satisfied or completely violated under the precondition of Assumption 1. Since we intend to later perform perturbation analysis, it is more convenient to adopt the stricter condition, allowing us to compare semigroups with very different degrees. Similarly, one could also weaken the requirements for Theorem 3.5, Theorem 3.9, and Theorem 3.10.

Before proving Theorem 3.1, we provide an example for which Assumption 2 is not satisfied.

Example 6 (Pure birth process [20, Ex. 3.3.]). Let $L = (a^\dagger)^2$ and $G = -\frac{1}{2}a^2(a^\dagger)^2$, i.e.

$$\mathcal{L}(x) = Gx + xG^\dagger + LxL^\dagger\,.$$

By construction, this generator satisfies Assumption 1. However, one can show that it is not trace-preserving, and therefore it cannot satisfy Assumption 2.

Remark. Note that Assumption 2 implies the conditions in [16, 17] under the precondition that the closure of $(\mathcal{L},\mathcal{D}(\mathcal{L})$ is a generator of a C_0-semigroup. Then a similar bound to Equation (30) can be shown.

***Proof strategy*:** Our proof is partly inspired by [20], however, Assumption 2 allows us to go beyond minimal semigroups and obtain a trace-preserving evolution. We believe that a similar result can be obtained by using the tools in [25]. An important intermediate step is that the considered generators are also generators on $W^{k,1}$, which will allow us to provide simple perturbation analysis on specific examples in Section 5. Our proof starts with Lemma 3.2, where we show that $G_\varepsilon \coloneqq G - \varepsilon(N+\mathbb{1})^{4d}$ is a generator on the Fock space and the implemented semigroup $t \mapsto e^{tG_\varepsilon} \cdot e^{tG_\varepsilon^\dagger}$ admits $\mathcal{T}_f$ as a core. Then Lemma 3.3 extends the result to semigroups on $W^{k,1}$ for all $k \in \mathbb{R}_+$. By the compact embedding lemma 2.4 for $W^{k,1}$ in $\mathcal{T}_{1,\mathrm{sa}}$, we can transfer these properties to the unperturbed evolution. Next, we more closely follow the method introduced in [20]. In particular, we prove that a perturbed version of Equation (28) generates a Sobolev and positivity preserving C_0-semigroup.

Lemma 3.2 *For $\varepsilon > 0$, the closure of the operator*

$$\mathcal{G}_\varepsilon : \mathcal{T}_f \to \mathcal{T}_f, \qquad x \mapsto Gx + xG^\dagger - \varepsilon\{(N+\mathbb{1})^{4d}, x\}\,,$$

where G is defined in Equation (28), generates a strongly continuous, contractive, positivity preserving semigroup on $\mathcal{T}_{1,\mathrm{sa}}$.

Proof. The proof is structured in the following two steps:

1) The closure of $G_\varepsilon : \mathcal{H}_f \to \mathcal{H}$, $|\psi\rangle \mapsto G_\varepsilon |\psi\rangle \coloneqq (-\varepsilon(N+\mathbb{1})^{4d} + G)|\psi\rangle$ generates a strongly continuous contractive semigroup on $\mathcal{H}$, which we denote by $(P_t^\varepsilon)_{t\geq 0}$.

2) The family of maps $(\mathcal{P}_t^\varepsilon \coloneqq P_t^\varepsilon \cdot (P_t^\varepsilon)^\dagger : \mathcal{T}_{1,\mathrm{sa}} \to \mathcal{T}_{1,\mathrm{sa}})_{t\geq 0}$, with $(P_t^\varepsilon)_{t\geq 0}$ from step 1, defines a strongly continuous, contractive, positivity preserving semigroup on $\mathcal{T}_{1,\mathrm{sa}}$ generated by the closure of $\mathcal{G}_\varepsilon$.

Step 1) By Assumption 1 there exists $p_\varepsilon \in \mathbb{C}[X,Y]$ such that $G_\varepsilon = p_\varepsilon(a, a^\dagger)$ which shows by Lemma 2.12 that G_ε is closed with domain $\mathcal{D}(N^{4d})$. We will now show dissipativity for G_ε and $G_\varepsilon^\dagger$ to conclude the claim using Corollary 2.8. It suffices also to consider $G_\varepsilon^\dagger$ on $\mathcal{H}_f$, as it is a core by Lemma 2.12 ($\deg(G) = 2d$) and therefore dissipativity of $G_\varepsilon^\dagger$ on $\mathcal{H}_f$ directly implies dissipativity of $G_\varepsilon^\dagger$ on all of its domain. We only show the dissipativity of G_ε, since the proof for $G_\varepsilon^\dagger$ is completely analogous. Let $|\psi\rangle \in \mathcal{H}_f$, then for any $\lambda > 0$

$$\begin{aligned}
&\|(\lambda - G_\varepsilon)|\psi\rangle\|^2 \\
&\quad= \lambda^2 \langle\psi,\psi\rangle + \langle G_\varepsilon\psi, G_\varepsilon\psi\rangle - \lambda(\langle G_\varepsilon\psi,\psi\rangle + \langle\psi, G_\varepsilon\psi\rangle) \\
&\quad\geq \lambda^2 \langle\psi,\psi\rangle - \lambda(\langle G_\varepsilon\psi,\psi\rangle + \langle\psi, G_\varepsilon\psi\rangle) \\
&\quad\geq \lambda^2 \langle\psi,\psi\rangle + \varepsilon\lambda(\langle(N+\mathbb{1})^{4d}\psi,\psi\rangle + \langle(N+\mathbb{1})^{4d}\psi,\psi\rangle) - \lambda(\langle G\psi,\psi\rangle + \langle\psi, G\psi\rangle) \\
&\quad\geq \lambda^2 \langle\psi,\psi\rangle - \lambda(\langle G\psi,\psi\rangle + \langle\psi, G\psi\rangle)\,.
\end{aligned}$$

By the requirement of Assumption 1, it is clear that $\mathrm{tr}[\mathcal{L}(x)] = 0$ for $x \in \mathcal{T}_f$, using the cyclicity of the trace. For the explicit case of a pure state $x = |\psi\rangle\!\langle\psi| \in \mathcal{T}_f$ where the last inclusion holds due to $|\psi\rangle \in \mathcal{H}_f$, we get

$$\lambda\,\langle\psi, -(G^\dagger + G)\psi\rangle = \lambda \sum_{j=1}^{K} \langle L_j\psi, L_j\psi\rangle \geq 0\,.$$

Hence, we conclude

$$\|(\lambda - G_\varepsilon)|\psi\rangle\|^2 \geq \lambda^2 \langle\psi,\psi\rangle = \lambda^2 \||\psi\rangle\|^2\,.$$

Taking the square root on both sides proves the claim. Note that for $G_\varepsilon^\dagger$ all steps are similar due to the simple observation that $\langle G\psi,\psi\rangle + \langle\psi, G\psi\rangle = \langle G^\dagger\psi,\psi\rangle + \langle\psi, G^\dagger\psi\rangle$ for $|\psi\rangle \in \mathcal{H}_f$.

Step 2) That the implemented semigroup[1] $(\mathcal{P}_t)_{t\geq 0}$ is a strongly continuous, positivity-preserving contractive semigroup that can be easily checked. We further get from [20, Prop. 2.1] that it is generated by the closure of the operator

$$\widetilde{\mathcal{G}}_\varepsilon : \mathcal{D}(\widetilde{\mathcal{G}}_\varepsilon) = \{R(1,\overline{G}_\varepsilon)xR(1,\overline{G}_\varepsilon)^\dagger \;:\; x \in \mathcal{T}_{1,\mathrm{sa}}\} \to \mathcal{T}_{1,\mathrm{sa}}, \quad x \mapsto \overline{G}_\varepsilon x + x\overline{G}_\varepsilon^\dagger$$

where $\overline{G}_\varepsilon$ is the closure of G_ε, $\overline{G}_\varepsilon^\dagger$ its adjoint, and $R(1,\overline{G}_\varepsilon)$ its resolvent on $\mathcal{H}$, respectively. Since $\mathcal{H}_f$ is a core for the generator $\overline{G}_\varepsilon$, the set $O \coloneqq (\mathbb{1} - \overline{G}_\varepsilon)\mathcal{H}_f = (\mathbb{1} - G_\varepsilon)\mathcal{H}_f$ is dense in $\mathcal{H}$, which in turn means $\mathcal{O} = \mathrm{span}\{|\psi\rangle\!\langle\varphi| \;:\; |\psi\rangle, |\varphi\rangle \in O\}$ is dense in $\mathcal{T}_{1,\mathrm{sa}}$. A simple calculation

[1] A discussion on implemented semigroups can be found in [3].

further shows that $R(1,\overline{G}_\varepsilon)\mathcal{O}R(1,\overline{G}_\varepsilon)^\dagger = \mathcal{T}_f$. Hence for $y \in \mathcal{D}(\widetilde{\mathcal{G}}_\varepsilon)$, we find $x \in \mathcal{T}_{1,\mathrm{sa}}$ and a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathcal{O}$ with $x_n \to x$ for $n \to \infty$ such that

$$\begin{aligned}\widetilde{\mathcal{G}}_\varepsilon(y) &= \widetilde{\mathcal{G}}_\varepsilon(R(1,\overline{G}_\varepsilon)xR(1,\overline{G}_\varepsilon)) = \widetilde{\mathcal{G}}_\varepsilon(R(1,\overline{G}_\varepsilon)\lim_{n\to\infty} x_nR(1,\overline{G}_\varepsilon)^\dagger)\\ &= \lim_{n\to\infty}\widetilde{\mathcal{G}}_\varepsilon(R(1,\overline{G}_\varepsilon)x_nR(1,\overline{G}_\varepsilon)^\dagger)\\ &= \lim_{n\to\infty}\mathcal{G}_\varepsilon(y_n)\end{aligned}$$

where we used the continuity of the map $\widetilde{\mathcal{G}}_\varepsilon(R(1,\overline{G}_\varepsilon)\cdot R(1,\overline{G}_\varepsilon)^\dagger)$ and the set defined by $y_n = R(1,\overline{G}_\varepsilon)x_nR(1,\overline{G}_\varepsilon)^\dagger$. Note that $\{y_n\}_{n\in\mathbb{N}}$ is by construction a convergent sequence on $\mathcal{T}_f$. In the last equality, we used that $\widetilde{\mathcal{G}}_\varepsilon$ and $\mathcal{G}_\varepsilon$ agree on $\mathcal{T}_f$. This shows not only that $\mathcal{G}_\varepsilon$ is closable but also that its closure is the closure of $\widetilde{\mathcal{G}}_\varepsilon$. Hence the closure of $\mathcal{G}_\varepsilon$ is the generator of $(\mathcal{P}_t^\varepsilon)_{t\geq 0}$. □

Using the above lemma, we are now able to prove the following.

Lemma 3.3 *For all $k \geq 0$ and $\varepsilon > 0$, the closure of the operator $\mathcal{G}_\varepsilon$ from Lemma 3.2 generates a strongly continuous, positivity preserving semigroup $(\mathcal{P}_t^\varepsilon)_{t\geq 0}$ on $W^{k,1}$ such that, for all $t \geq 0$,*

$$\|\mathcal{P}_t^\varepsilon\|_{W^{k,1}\to W^{k,1}} \leq e^{\omega_k t}$$

where $\omega_k = \frac{k_{r_1}-k}{k_{r_1}-k_{r_0}}\omega_{k_{r_0}} + \frac{k-k_{r_0}}{k_{r_1}-k_{r_0}}\omega_{k_{r_1}}$ for an r such that $k_{r_0} \leq k < k_{r_1}$. Finally, for $k = 0$, the semigroup is contractive.

Proof. Without loss of generality, we can restrict to $k \in \{k_r\}_{r\in\mathbb{N}}$ as for k inbetween, we can interpolate between the $\{k_r\}_{r\in\mathbb{N}}$ (shown below) and $k = 0$ (shown in Lemma 3.2) using Lemma E.4. Let $\varepsilon > 0$. In the following proof, the closure, domain and boundedness of an operator are always with respect to the Banach space $W^{k,1}$ if not stated otherwise. We will show the claim, by first arguing that $\mathcal{G}_\varepsilon$ is closable, that all $\lambda > \omega_k$ are in the resolvent set of the closure $\overline{\mathcal{G}}_\varepsilon$ and further that $\left\|R(\lambda,\overline{\mathcal{G}}_\varepsilon)\right\|_{W^{k,1}\to W^{k,1}} \leq \frac{1}{\lambda-\omega_k}$. By Theorem 2.5, the above immediately gives the existence of the semigroup on $W^{k,1}$ and provides us with the claimed bound. The property of positivity preservation traces back to the representation of the semigroup via the Euler approximation and therefore the positivity of the resolvent: for any $x \in \mathcal{T}_{1,\mathrm{sa}}$:

$$\mathcal{P}_t^\varepsilon(x) = \lim_{n\to\infty}\left(\frac{n}{t}\right)^n R(n/t,\overline{\mathcal{G}}_\varepsilon)^n(x)\,. \tag{31}$$

The claims are proven in three steps.

Step 1. Show that $\mathcal{G}_\varepsilon : \mathcal{T}_f \to \mathcal{T}_f$ is closable and there exists a $\lambda > \omega_k$ such that $\lambda - \overline{\mathcal{G}}_\varepsilon : \mathcal{D}(\overline{\mathcal{G}}_\varepsilon) \to W^{k,1}$ is bijective.

Step 2. Using Assumption 2 and Lemma 3.2 we prove that if $\lambda > \omega_k$ is in the resolvent set of $\overline{\mathcal{G}}_\varepsilon$, we not only have that the resolvent is positivity preserving but further

$$\left\|R(\lambda,\overline{\mathcal{G}}_\varepsilon)\right\|_{W^{k,1}\to W^{k,1}} \leq \frac{1}{\lambda-\omega_k}\,.$$

Step 3. The surjectivity of $\lambda - \overline{\mathcal{G}}_\varepsilon$ for a specific $\lambda > \omega_k$ from step 1. and the bound on the resolvent from step 2. allow us to successively use the series expansion of the resolvent as it is done in [23, Prop. IV.1.3] to get that (ω_k,∞) is in the resolvent set of $\overline{\mathcal{G}}_\varepsilon$, and therefore conclude the proof.

Proof of step 1. We introduce the map

$$\mathcal{I}_{d,\varepsilon} : \mathcal{T}_f \to \mathcal{T}_f, \qquad x \mapsto \mathcal{I}_{d,\varepsilon}(x) := -\varepsilon\{(N+\mathbb{1})^{4d}, x\}\,.$$

For $\lambda \geq 0$, $x \in \mathcal{T}_f$, we can use Lemma E.1 to write

$$(\lambda - \mathcal{G}_\varepsilon)(x) = (\mathbb{1} - \mathcal{G}_0 \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}) \circ (\lambda - \mathcal{I}_{d,\varepsilon})(x)$$

where $\circ$ is the function composition and $\lambda - \mathcal{I}_{d,\varepsilon} : \mathcal{T}_f \to \mathcal{T}_f$ is a bijection, with bounded inverse (see Lemma E.1) between dense subspaces of $W^{k,1}$. This means in particular that it is closable and that its closure has a bounded inverse. We will hence focus on the map $\mathbb{1} - \mathcal{G}_0 \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1} : \mathcal{T}_f \to \mathcal{T}_f$ and show that it is bounded (on the dense subset $\mathcal{T}_f$ of $W^{k,1}$) and hence uniquely extendable to all of $W^{k,1}$. This then immediately gives us that $\lambda - \mathcal{G}_0 : \mathcal{T}_f \to \mathcal{T}_f$ is closable as it is the composition of a map with a dense range succeeded by a bounded map. Note first that we can apply Lemma E.3 to $\mathcal{G}_0$ to get that there exists $C_k \geq 0$ such that for all $\kappa > 0$ and $x \in \mathcal{T}_f$

$$\|\mathcal{G}_0(x)\|_{W^{k,1}} \leq \kappa \|\mathcal{I}_{d,\varepsilon}(x)\|_{W^{k,1}} + \frac{C_k}{\kappa\varepsilon}\|x\|_{W^{k,1}}\,.$$

Using the bijectivity of $\lambda - \mathcal{I}_{d,\varepsilon} : \mathcal{T}_f \to \mathcal{T}_f$ we get for $x \in \mathcal{T}_f$

$$\begin{aligned}\left\|\mathcal{G}_0 \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1} x\right\|_{W^{k,1}} &\leq \kappa \left\|\mathcal{I}_{d,\varepsilon} \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x)\right\|_{W^{k,1}} + \frac{C_k}{\varepsilon\kappa}\left\|(\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x)\right\|_{W^{k,1}} \\ &\leq (2\kappa + \frac{C_k}{\kappa\varepsilon}\frac{1}{\lambda + 2\varepsilon})\|x\|_{W^{k,1}} =: f_k(\lambda,\kappa)\|x\|_{W^{k,1}}\end{aligned}$$

where we used properties of $\mathcal{I}_{d,\varepsilon}$ derived in Lemma E.1. This gives us not only that $\mathcal{G}_0 \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1} : \mathcal{T}_f \to \mathcal{T}_f$ is bounded, hence uniquely extendable to a bounded map on $W^{k,1}$ but for a fixed $\kappa < \frac{1}{2}$ and $\lambda > \lambda_\kappa$ where λ_κ is chosen s.t. $f_k(\lambda_\kappa, \kappa) < 1$, we get that its closure is a strict contraction on $W^{k,1}$. As a direct consequence, we find that again for $\lambda > \lambda_\kappa$ the closure of $\mathbb{1} - \mathcal{G}_0 \circ (\lambda - \mathcal{I}_{d,\varepsilon}) : \mathcal{T}_f \to \mathcal{T}_f$ is invertible with bounded inverse, and that its inverse function is just given by the geometric series of the closure of $\mathcal{G}_0 \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}$. To conclude, we can set $\lambda = 0$ in the above result and get that $-\mathcal{G}_\varepsilon$ and hence $\mathcal{G}_\varepsilon$ is closable and further that for $\kappa < \frac{1}{2}$ all λ with $\lambda > \lambda_\kappa$ are in the resolvent set of $\overline{\mathcal{G}}_\varepsilon$.

Proof of step 2. Let $\lambda > \omega_k$ be in the resolvent set of $\overline{\mathcal{G}}_\varepsilon$. From the compact embedding of $W^{k,1}$ in $\mathcal{T}_{1,\text{sa}}$, we immediately get that $R(\lambda, \overline{\mathcal{G}}_\varepsilon) : W^{k,1} \to W^{k,1}$ agrees with the respective restricted resolvent of the closure $\widehat{\mathcal{G}}_\varepsilon$ of $\mathcal{G}_\varepsilon$ on $\mathcal{T}_{1,\text{sa}}$ that we obtained in Lemma 3.2. We know that the latter resolvent is positivity preserving, as the semigroup is. This is due to the following integral representation for strongly continuous semigroups [23, Thm. II.1.10 (i)]: for all $x \in x(\widehat{\mathcal{G}}_\varepsilon)$,

$$R(\lambda, \widehat{\mathcal{G}}_\varepsilon)(x) = \int_0^\infty e^{-\lambda s}\, e^{s\widehat{\mathcal{G}}_\varepsilon}(x)\, ds\,. \tag{32}$$

Hence $R(\lambda, \overline{\mathcal{G}}_\varepsilon)$ is positivity preserving as well. Using Assumption 2, we have that for $x \in \mathcal{T}_f$, x positive semi-definite,

$$\operatorname{tr}\left[\mathcal{L}(x)(N+\mathbb{1})^{k/2}\right] \leq \omega_k \operatorname{tr}\left[x(N+\mathbb{1})^{k/2}\right].$$

Adding non-negative terms, using the cyclicity of the trace and splitting up $\mathcal{L}$ gives us

$$\begin{aligned}\sum_{j=1}^{K} \operatorname{tr}\left[(N+\mathbb{1})^{k/4} L_j x L_j^\dagger (N+\mathbb{1})^{k/4}\right] + (\lambda - \omega_k)\operatorname{tr}\left[(N+\mathbb{1})^{k/4} x (N+\mathbb{1})^{k/4}\right] \\ \leq \operatorname{tr}\left[(N+\mathbb{1})^{k/4}(\lambda - \mathcal{G}_\varepsilon)(x)(N+\mathbb{1})^{k/4}\right],\end{aligned}$$

and therefore

$$(\lambda - \omega_k)\|x\|_{W^{k,1}} \le \|(\lambda - \mathcal{G}_\varepsilon)(x)\|_{W^{k,1}}$$

where we have just dropped non-negative terms and used $\mathrm{tr}[\cdot] \le \|\cdot\|_1$ with equality if the argument is positive semi-definite. Since $\overline{\mathcal{G}}_\varepsilon$ is the closure of $\mathcal{G}_\varepsilon$, the above inequality extends to $x \in \mathcal{D}(\overline{\mathcal{G}}_\varepsilon)$, x positive semi-definite and $\overline{\mathcal{G}}_\varepsilon$ instead of $\mathcal{G}_\varepsilon$. Together with the positivity preserving property of the resolvent, this gives us that for all $x \in W^{k,1}$, x positive semi-definite

$$\left\|R(\lambda, \overline{\mathcal{G}}_\varepsilon)x\right\|_{W^{k,1}} \le \frac{1}{\lambda - \omega_k}\|x\|_{W^{k,1}}\,. \tag{33}$$

For a general $x \in W^{k,1}$, we set $x_\pm = \frac{1}{(N+\mathbb{1})^{k/4}}\left[(N+1)^{k/4}x(N+\mathbb{1})^{k/4}\right]_\pm \frac{1}{(N+\mathbb{1})^{k/4}} \in W^{k,1}$, where $[\cdot]_\pm$ denotes the positive, resp. the negative part of a self-adjoint trace-class operator. We have that $x = x_+ - x_-$ and further that x_+, x_- are positive semi-definite by construction. Hence

$$\begin{aligned}\left\|R(\lambda, \overline{\mathcal{G}}_\varepsilon)x\right\|_{W^{k,1}} &\le \left\|R(\lambda, \overline{\mathcal{G}}_\varepsilon)x_+\right\|_{W^{k,1}} + \left\|R(\lambda, \overline{\mathcal{G}}_\varepsilon)x_-\right\|_{W^{k,1}} \\ &\le \frac{1}{\lambda - \omega_k}(\|x_+\|_{W^{k,1}} + \|x_-\|_{W^{k,1}}) = \frac{1}{\lambda - \omega_k}\|x_+ - x_-\|_{W^{k,1}} \\ &= \frac{1}{\lambda - \omega_k}\|x\|_{W^{k,1}}\end{aligned}$$

where we used Equation (33) and the construction of x_+ and x_-, which concludes step 2.

Proof of step 3. From step 1. we get that there exists a $\lambda > \omega_k$ in the resolvent set of $\overline{\mathcal{G}}_\varepsilon$ whereas step 2. tells us that, for this λ, the resolvent is bounded by $\frac{1}{\lambda - \omega_k}$. We can use the same proof strategy as in [23, Prop. II.3.14 (ii)] where the authors employ the series expansion of the resolvent and its explicit bound to make conclusions about the resolvent set. Following their steps we first get that $(\omega_k, 2\lambda - \omega_k)$ is part of the resolvent set, and then using step 2. again, we obtain the positivity preservation property as well as the explicit bound for all of those resolvents. This allows us to successively use these arguments and conclude that the resolvent set contains (ω_k, ∞). □

Putting together the results from Lemma 3.2 and Lemma 3.3, we are now able to get rid of the perturbation $\mathcal{I}_{d,\varepsilon}$.

Lemma 3.4 *The closure of*

$$\mathcal{G} : \mathcal{T}_f \to \mathcal{T}_f, \quad x \mapsto \mathcal{G}(x) = Gx + xG^\dagger\,,$$

where G is defined in Equation (28), generates a strongly continuous, positivity preserving semigroup $(\mathcal{P}_t)_{t\ge 0}$ on $W^{k,1}$ for all $k \in \mathbb{N}$ with

$$\|\mathcal{P}_t\|_{W^{k,1}\to W^{k,1}} \le e^{\omega_k t}\,.$$

where $\omega_k = \frac{k_{r_1} - k}{k_{r_1} - k_{r_0}}\omega_{k_{r_0}} + \frac{k - k_{r_0}}{k_{r_1} - k_{r_0}}\omega_{k_{r_1}}$ for an r such that $k_{r_0} \le k < k_{r_1}$. Finally, for $k = 0$, the semigroup is contractive.

Proof. The proof is a direct application of Lemma E.5 to the semigroups we obtained in Lemma 3.3 taking $\varepsilon \to 0$. Since the semigroups in Lemma 3.3 were positivity preserving, so is the obtained semigroup in the limit $\varepsilon \to 0$ (c.f. Lemma E.5). □

We are now ready to prove the main Theorem of the section.

Proof of Theorem 3.1. The proof strategy is inspired by [20, Thm. 2.5]. It however makes use of Lemma E.5 to avoid the issues discussed in [20, §3]. Let $k \in \{k_r\}_{r\in\mathbb{N}}$ or $k = 0$ for the moment. Note that from Assumption 1, we can conclude $\omega_k = 0$ for $k = 0$ in Equation (29). We first define for $\delta \in (0, 1)$ the map

$$\mathcal{L}_\delta : \mathcal{T}_f \to \mathcal{T}_f, \quad x \mapsto \mathcal{L}_\delta(x) = Gx + xG^\dagger + \delta \sum_{j=1}^{K} L_j x L_j^\dagger =: \mathcal{G}(x) + \delta\Sigma(x),$$

and show that its closure defines a strongly continuous, positivity preserving semigroup $(\mathcal{P}_t^\delta)_{t\geq 0}$ on $W^{k,1}$ which further satisfies

$$\left\|\mathcal{P}_t^\delta\right\|_{W^{k,1}\to W^{k,1}} \leq e^{\omega_k t}.$$

We first note that for $\widetilde{\lambda} > 0$, a rearrangement of Equation (29) using cyclicity of the trace and that $\mathrm{tr}[\cdot] \leq \|\cdot\|_1$ with equality if the argument is positive semi-definite gives

$$\|\Sigma(x)\|_{W^{k,1}} \leq \left\|(\widetilde{\lambda} + \omega_k - \mathcal{G})(x)\right\|_{W^{k,1}}$$

for $x \in \mathcal{T}_f$ and $x \geq 0$. Now using that $\mathcal{G}$ is closable (Lemma 3.4) and its resolvent positivity preserving we can conclude for $\lambda := \widetilde{\lambda} + \omega_k > \omega_k$, $x \in (\lambda - \mathcal{G})\mathcal{T}_f$ and $x \geq 0$,

$$\left\|\Sigma \circ R(\lambda, \overline{\mathcal{G}})(x)\right\|_{W^{k,1}} \leq \|x\|_{W^{k,1}}.$$

Applying similar methods as in step 2. of the proof of Lemma 3.3, we can extend the above inequality to general $x \in (\lambda - \mathcal{G})\mathcal{T}_f$. Hence, $\Sigma \circ R(\lambda, \overline{\mathcal{G}})$ is contractive on the dense set $(\lambda - \mathcal{G})\mathcal{T}_f$ and positivity preserving, since both Σ and $R(\lambda, \overline{\mathcal{G}})$ are. It can therefore be uniquely extended to a positivity preserving contractive map on all of $W^{k,1}$ which we will call $\mathcal{A}_\lambda$ in the following. As a consequence $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$ is closable and $\lambda > \omega_k$ in the resolvent set of the closure. Both facts follow from the representation

$$(\lambda - \mathcal{L}_\delta) = (\mathbb{1} - \delta\Sigma \circ R(\lambda, \overline{\mathcal{G}})) \circ (\lambda - \mathcal{G})$$

which decomposes $\lambda - \mathcal{L}_\delta$ into a composition of a closable map with a dense range and a map that is bounded on that range. We further get for the resolvent of the closure

$$R(\lambda, \overline{\mathcal{L}}_\delta) = R(\lambda, \overline{\mathcal{G}}) \sum_{n=0}^{\infty} \delta^n \mathcal{A}_\lambda^n,$$

which immediately lets us conclude that the resolvent is positivity preserving as $\mathcal{A}_\lambda$ and $R(\lambda, \overline{\mathcal{G}})$ are. Lastly, we will show that for $\lambda > \omega_k$

$$\left\|R(\lambda, \overline{\mathcal{L}}_\delta)\right\|_{W^{k,1}\to W^{k,1}} \leq \frac{1}{\lambda - \omega_k}. \tag{34}$$

To obtain this inequality we again rearrange Assumption 2, add non-negative terms, use cyclicity of the trace and that $\mathrm{tr}[\cdot] \leq \|\cdot\|_1$ with equality if the argument is positive semi-definite, to conclude that for $x \in (\lambda - \mathcal{L}_r)\mathcal{T}_f$, x positive semi-definite,

$$\left\|R(\lambda, \overline{\mathcal{L}}_\delta)x\right\|_{W^{k,1}} \leq \frac{1}{\lambda - \omega_k}\|x\|_{W^{k,1}}.$$

We again extend the above bound to all $x \in (\lambda - \mathcal{L}_\delta)\mathcal{T}_f$ analogously to step 2 in the proof of Lemma 3.4. Using that $(\lambda - \mathcal{L}_\delta)\mathcal{T}_f$ is dense then gives Equation (34). Employing Theorem 2.7, we get that indeed for all $\delta \in (0,1)$ the closure of $(\mathcal{L}_\delta, \mathcal{D}(\mathcal{L}_\delta))$ generates a strongly continuous semigroup which is positivity preserving since the resolvent is and satisfies the claimed bound. To now fill the gap between 0 and the $\{k_r\}_{r\in\mathbb{N}}$ respectively, we interpolate between the semigroups (q.v. Lemma E.4), obtaining $e^{t\omega_k}$ where $\omega_k = \frac{k_{r_1}-k}{k_{r_1}-k_{r_0}}\omega_{k_{r_0}} + \frac{k-k_{r_0}}{k_{r_1}-k_{r_0}}\omega_{k_{r_1}}$ for an r such that $k_{r_0} \leq k < k_{r_1}$, as the bound of the interpolated semigroups. Now that we have the result for all $k \geq 0$ we can employ Lemma E.5 and take the limit $\delta \to 1$ to obtain the assertion. The contractivity and trace-preserving property of the semigroup in the case $k = 0$ just follows from the GKSL form of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, i.e. $\operatorname{tr}[\mathcal{L}(x)] = 0$ for $x \in \mathcal{T}_f$, or put differently Assumption 1. □

3.2 Bosonic evolution systems

Next, we consider time-dependent generators in GKSL form. For this, we modify Assumptions 1 and 2 in the following way:

Assumption 3 *The operator $(\mathcal{L}_s, \mathcal{T}_f)$ has* GKSL *form, i.e. for $x \in \mathcal{T}_f$ and $s \in [0,\infty)$*

$$\begin{aligned}\mathcal{L}_s : \mathcal{T}_f \to \mathcal{T}_f \quad x \mapsto \mathcal{L}_s(x) &= -i[H(s), x] + \sum_{j=1}^K L_j(s) x L_j^\dagger(s) - \frac{1}{2}\{L_j^\dagger(s) L_j(s), x\} \\ &\coloneqq G(s)x + xG^\dagger(s) + \sum_{j=1}^K L_j(s) x L_j^\dagger(s)\,,\end{aligned} \tag{35}$$

where $K \in \mathbb{N}$, $G(s) = -iH(s) - \frac{1}{2}\sum_{j=1}^K L_j^\dagger(s) L_j(s)$, and $H(s) \coloneqq p_{H(s)}(a, a^\dagger), L_j(s) \coloneqq p_{j,s}(a, a^\dagger)$ are polynomials of the creation and annihilation operators with time-dependent, continuous coefficients. Again, $d_H \coloneqq \sup_{s\geq 0} \deg(p_{H(s)}) < \infty$, $d_j \coloneqq \sup_{s\geq 0}\deg(p_{j,s}) < \infty$, and $d \coloneqq \max\{d_1, ..., d_K, d_H\}$.

The next assumption will lead to the evolution system being Sobolev preserving, which allows us not only to prove the existence and uniqueness of the evolution generated by (27) but further to conduct a perturbation analysis as well as to extend our results to the case of a time-dependent Master equation:

Assumption 4 *There exists a non-negative sequence $\{k_r\}_{r\in\mathbb{N}} \to \infty$ s.t. for all $r \in \mathbb{N}$ there exist $\omega_{k_r} \geq 0$ such that for all $s \in \mathbb{R}_+$ and $x \in \mathcal{T}_f$ positive semi-definite,*

$$\operatorname{tr}\Big[\mathcal{L}_s(x)(N+\mathbb{1})^{k/2}\Big] \leq \omega_{k_r} \operatorname{tr}\Big[x(N+\mathbb{1})^{k/2}\Big]\,. \tag{36}$$

Note that the coefficients ω_{k_r} are independent of s.

Under the above assumptions we can state the generation theorem for evolution systems as follows:

Theorem 3.5 (Generation of bosonic evolution systems) *Let $(\mathcal{L}_s, \mathcal{D}(\mathcal{L}_s))_{s\in[0,\infty)}$ be a family of operators that fulfill Assumption 3 and Assumption 4. Then $(\overline{\mathcal{L}}_s, \mathcal{D}(\overline{\mathcal{L}}_s))_{s\in\mathbb{R}_+}$ gives rise to a unique evolution system $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ on $W^{k,1}$ for all $k \geq 0$ with the following properties*

1. *$\mathcal{P}_{t,s}(W^{k+4d,1}) \subseteq W^{k+4d,1}$ for all $0 \leq s \leq t$*
2. *For any $x \in W^{k+4d,1}$, the family $(\mathcal{P}_{t,s}(x))_{0\leq s\leq t}$ is the unique solution to the initial value problem*
$$\frac{d}{dt}x(t) = \overline{\mathcal{L}}_t(x(t)) \qquad t \in [s,\infty),\ x(s) = x\,. \tag{37}$$

For $k = 0$, the evolution system is contractive and trace-preserving.

Proof. We assume w.l.o.g. that $s \in [0,1]$ is fixed since the same argument works for all compact intervals. Theorem 3.1 shows that $(\overline{\mathcal{L}}_s, \mathcal{D}(\overline{\mathcal{L}}_s))$ generates an ω_k-quasi-contractive semigroup $(\mathcal{P}_t^s)_{t\geq 0}$ on $W^{k,1}$. Next, we realize that $W^{k+4d,1}$ are $\mathcal{L}_s$-admissible subspaces, where we recall that d denotes the degree of $\mathcal{L}_s$. This already proves assumptions (1) and (2) in Theorem 2.10. Since the coefficients of the polynomials p_H and p_j are continuous and operators of the form

$$(N+\mathbb{1})^{k/4} a^j (a^\dagger)^l (N+\mathbb{1})^{-(k/4+d)}\,,$$

for $j + l \leq d$, are bounded (see Lemma E.2) w.r.t. the operator norm, we have by Hölder inequality that

$$s \mapsto (N+\mathbb{1})^{k/4} \mathcal{L}_s((N+\mathbb{1})^{-k/4+d}(\cdot)(N+\mathbb{1})^{-k/4+d})(N+\mathbb{1})^{k/4} =: \mathcal{A}(s)$$

is a bounded and uniformly continuous family of operators. Therefore,

$$s \mapsto \mathcal{L}_s \in \mathcal{B}(W^{k,1}, W^{k+4d,1})$$

is uniformly continuous, which proves condition (3) in Theorem 2.10. Hence Theorem 2.10 provides the existence and uniqueness of an evolution system on $W^{k,1}$. By repeating the above arguments on $\mathcal{Y} := W^{k+4d}$, i.e. by choosing our $\mathcal{L}_s$-admissible subspace as $W^{k+8d,1}$, Theorem 2.10 provides existence and uniqueness of a solution on $\mathcal{Y} = W^{k+4d,1}$ which agrees with the former one on $W^{k,1}$ by the compact embedding of $W^{k+4d,1}$ into $W^{k,1}$. Therefore, conditions (4) and (5) are satisfied for the evolution system on $W^{k,1}$ and the admissible subspace $\mathcal{Y} = W^{k+4d,1}$, which through Theorem 2.10 proves the claim. Moreover, the evolution system is positivity preserving because it can be constructed by a concatenation of time-independent positivity preserving semigroups (see Theorem 3.1 and [54, Eq. 5.3.5]). Contractivity and the property of trace preservation are a consequence of the fact that ω_0 can be chosen to be 0. □

3.3 Multi-mode extension

This section discusses the extension of Section 3 to the multi-mode setting. Since the details are almost completely analogous to the single-mode situation, we choose to elaborate only at places where some ambiguities might remain. Let us first fix the notations for this setting. We consider the Hilbert space of an m-mode system, $m \in \mathbb{N}$, whose Hilbert space we conveniently denote by $\mathcal{H}_m = L^2(\mathbb{R}^m)$. We further use $\mathcal{B}(\mathcal{H}_m)$ for the bounded, $\mathcal{T}_1$ for the trace class, and $\mathcal{T}_{1,\mathrm{sa}}$ for the self-adjoint trace class operators. Now we define $\mathcal{T}_f$ to be

$$\mathcal{T}_f := \{x = \sum_{\text{finite}} f_{\mathbf{n},\mathbf{p}} \,|n_1\rangle\langle p_1| \otimes \ldots \otimes |n_m\rangle\langle p_m| \;:\; f_{\mathbf{n},\mathbf{p}} \in \mathbb{C},\; x = x^\dagger\}\,,$$

where $\mathbf{n} = (n_1, \ldots, n_m) \in \mathbb{N}^m$ and $\mathbf{p}$ analogously function as an index in $f_{\mathbf{n},\mathbf{p}}$. For $\mathbf{k} = (k_1, \ldots, k_m) \in \mathbb{R}_+^m$, we define $\mathbf{k} \prec \mathbf{k}'$ if $k_j < k_j'$ for all $j = 1, \ldots, m$. Analogously, we define $\preceq$. We set for $\mathbf{k} \in \mathbb{R}_+^m$

$$(N+\mathbb{1})^{\mathbf{k}} := (N_1+\mathbb{1})^{k_1} \otimes \ldots \otimes (N_m+\mathbb{1})^{k_m},$$

and with this define $W^{\mathbf{k},1}$, $\|\cdot\|_{W^{\mathbf{k},1}}$. Remark that the latter spaces are Banach spaces and in correspondence to Lemma 2.13 we find:

Lemma 3.6 *Let $\boldsymbol{k}, \boldsymbol{k}' \in \mathbb{N}^m$ with $\boldsymbol{k} \prec \boldsymbol{k}'$, then*

$$W^{\boldsymbol{k}',1} \Subset W^{\boldsymbol{k},1}\,. \tag{38}$$

The strategy to prove the above claims is analogous to the single-mode case. Next, we slightly generalize the single-mode results Theorem 2.14 and Definition 2.15 to the multi-mode setting.

Theorem 3.7 (Stein-Weiss theorem for multi-mode Bosonic Sobolev spaces) *Let $\boldsymbol{k}_0, \boldsymbol{k}_1 \in \mathbb{R}_+^m$, $\boldsymbol{k}_0 \prec \boldsymbol{k}_1$ and $T : W^{\boldsymbol{k}_j,1} \to W^{\boldsymbol{k}_j,1}$ a linear map with $\|T\|_{W^{\boldsymbol{k}_j,1}\to W^{\boldsymbol{k}_j,1}} \leq M_j$, bounded by $M_j \geq 0$ for $j = 0, 1$ respectively. Then for $\theta \in [0, 1]$, $T : W^{\boldsymbol{k}_\theta,1} \to W^{\boldsymbol{k}_\theta,1}$ with $\boldsymbol{k}_\theta = (1-\theta)\boldsymbol{k}_0 + \theta\boldsymbol{k}_1$ obtained by restriction of the input of $T : W^{\boldsymbol{k}_0,1} \to W^{\boldsymbol{k}_0,1}$ to $W^{\boldsymbol{k}_\theta,1} \cap W^{\boldsymbol{k}_0,1}$, is a well defined bounded linear map with*

$$\|T\|_{W^{\boldsymbol{k}_\theta,1}\to W^{\boldsymbol{k}_\theta,1}} \leq M_0^{(1-\theta)} M_1^{\theta} \,. \tag{39}$$

In the multi-mode setting, we cannot interpolate between the elements of the divergent sequence to obtain $W^{\mathbf{k},1}$ as admissible subspace for all $0 \prec \mathbf{k}$ but only for elements in the convex hull of the divergent sequence. The property of being Sobolev preserving is again defined for a sequence $\{\mathbf{k}_r\}_{r\in\mathbb{N}}$ such that $\lim_{r\to\infty} \min_{j=1,\ldots,m} k_{j,r} = \infty$.

Definition 3.8 (Sobolev preserving semigroup/evolution system in multi-mode systems) Let $(\mathcal{P}_t)_{t\geq 0}$ be a C_0-semigroup on $\mathcal{T}_{1,\mathrm{sa}}$. We then call $(\mathcal{P}_t)_{t\geq 0}$ *Sobolev preserving* if there exists a divergent sequence $\{\mathbf{k}_r\}_{r\in\mathbb{N}} \subset \mathbb{R}_+^m$, in the sense that $\lim_{r\to\infty} \min_{j=1,\ldots,m} k_{j,r} = \infty$, s.t. for all $r \in \mathbb{N}$, $W^{\mathbf{k}_r,1}$ is an admissible subspace for $(\mathcal{P}_t)_{t\geq 0}$. Similarly for an evolution system $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ on $\mathcal{T}_{1,\mathrm{sa}}$, we call it $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ *Sobolev preserving* if for all $r \in \mathbb{N}$, $W^{\mathbf{k}_r,1}$ is admissible for $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ to $W^{\mathbf{k}_r,1}$.

With these preliminaries in place we can now lift Assumption 1, Assumption 2, Assumption 3, and Assumption 4.

Assumption 5 *The operator $(\mathcal{L}, \mathcal{T}_f)$ has* GKSL *form, i.e. for $x \in \mathcal{T}_f$,*

$$\begin{aligned}\mathcal{L} : \mathcal{T}_f \to \mathcal{T}_f \quad x \mapsto \mathcal{L}(x) &= -i[H, x] + \sum_{j=1}^{K} L_j x L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, x\} \\ &\coloneqq Gx + xG^\dagger + \sum_{j=1}^{K} L_j x L_j^\dagger \,,\end{aligned} \tag{40}$$

for some $K \in \mathbb{N}$ and with $G = -iH - \frac{1}{2}\sum_{j=1}^{K} L_j^\dagger L_j$, where $\{A, B\} = AB + BA$ denotes the anti-commutator of two operators A, B on a suitable domain. Further H and L_j are assumed to be polynomials of the creation and annihilation operators, i.e. $H \coloneqq p_H(a_1, a_1^\dagger, \ldots, a_m, a_m^\dagger)$, $L_j \coloneqq p_j(a_1, a_1^\dagger, \ldots, a_m, a_m^\dagger)$ and H symmetric. This ensures that $\mathcal{T}_f$ is invariant under $\mathcal{L}$. We denote the degree of p_H by $d_H \coloneqq \deg p_H$, those of p_j by $d_j \coloneqq \deg p_j$, and $d \coloneqq \max\{d_1, \ldots, d_K, d_H\}$.

The second assumption becomes:

Assumption 6 *There exists a non-negative sequence $\{\boldsymbol{k}_r\}_{r\in\mathbb{N}} \subset \mathbb{R}_+^m$, in the sense that $\lim_{r\to\infty} \min_{j=1,\ldots,m} k_{j,r} = \infty$, s.t. for every $r \in \mathbb{N}$, there exist $\omega_{\boldsymbol{k}_r} \geq 0$ such that for all positive semi-definite $x \in \mathcal{T}_f$*

$$\operatorname{tr}\Big[\mathcal{L}(x)(N + \mathbb{1})^{\boldsymbol{k}_r/2}\Big] \leq \omega_{\boldsymbol{k}_r} \operatorname{tr}\Big[x(N + \mathbb{1})^{\boldsymbol{k}_r/2}\Big] \,. \tag{41}$$

Then employing the single-mode strategy, we obtain the following theorem.

Theorem 3.9 (Generation of multi-mode bosonic semigroups) *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ be an operator defined on $\mathcal{T}_{1,\mathrm{sa}}$. If $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ satisfies Assumption 5 and Assumption 6, then the closure $\overline{\mathcal{L}}$ generates a strongly continuous, positivity preserving semigroup $(\mathcal{P}_t)_{t\geq 0}$ on $W^{\boldsymbol{k}_r,1}$ for all $\{\boldsymbol{k}_r\}_{r\in\mathbb{N}}$ from Assumption 6. We further find that the semigroup satisfies the bound*

$$\|\mathcal{P}_t\|_{W^{\boldsymbol{k}_r,1}\to W^{\boldsymbol{k}_r,1}} \leq e^{\omega_{\boldsymbol{k}_r} t} \quad \forall t \geq 0\,. \tag{42}$$

In the special case $\boldsymbol{k} = 0$, the semigroup is contractive and trace-preserving.

Note that we can extend the above semigroups to the convex hull of $\{\mathbf{k}_r\}_{r\in\mathbb{N}} \cup \{0\}$ using a generalisation of the interpolation lemma for single mode semigroups (q.v. Lemma E.4).

Lastly, we can also generalize the generation theorem for evolution systems modifying the assumptions accordingly.

Assumption 7 *The operator $(\mathcal{L}_s, \mathcal{T}_f)$ has* GKSL *form, i.e. for $x \in \mathcal{T}_f$ and $s \in [0,\infty)$*

$$\begin{aligned}\mathcal{L}_s : \mathcal{T}_f \to \mathcal{T}_f \quad x \mapsto \mathcal{L}_s(x) &= -i[H(s), x] + \sum_{j=1}^{K} L_j(s) x L_j^\dagger(s) - \frac{1}{2}\{L_j^\dagger(s) L_j(s), x\} \\ &\coloneqq G(s)x + xG^\dagger(s) + \sum_{j=1}^{K} L_j(s) x L_j^\dagger(s)\,,\end{aligned} \tag{43}$$

where $K \in \mathbb{N}$, $G(s) = -iH(s) - \frac{1}{2}\sum_{j=1}^{K} L_j^\dagger(s) L_j(s)$, and $H(s) \coloneqq p_{H(s)}(a_1, a_1^\dagger, \ldots, a_m, a_m^\dagger)$, $L_j(s) \coloneqq p_{j,s}(a_1, a_1^\dagger, \ldots, a_m, a_m^\dagger)$ are polynomials of the creation and annihilation operators with time-dependent, continuous coefficients. Again, $d_H \coloneqq \sup_{s\geq 0} \deg(p_{H(s)}) < \infty$, $d_j \coloneqq \sup_{s\geq 0} \deg(p_{j,s}) < \infty$, and $d \coloneqq \max\{d_1, ..., d_K, d_H\}$.

The second assumption in the time-dependent case generalizes to the following:

Assumption 8 *There is a divergent sequence $\{\boldsymbol{k}_r\}_{r\in\mathbb{N}} \subset \mathbb{R}_+^m$, meaning $\lim_{r\to\infty} \min_{j=1,\ldots,m} k_{j,r} = \infty$, s.t. for every $r \in \mathbb{N}$, there exist $\omega_{\boldsymbol{k}_r} \geq 0$ such that for all $s \in \mathbb{R}_+$ and $x \in \mathcal{T}_f$ positive semi-definite,*

$$\operatorname{tr}\Big[\mathcal{L}_s(x)(N+\mathbb{1})^{\boldsymbol{k}_r/2}\Big] \leq \omega_{\boldsymbol{k}_r} \operatorname{tr}\Big[x(N+\mathbb{1})^{\boldsymbol{k}_r/2}\Big]\,. \tag{44}$$

Note that the coefficients $\omega_{\boldsymbol{k}_r}$ are independent of s.

Under the above assumptions we can state the generation theorem for multi-mode evolution systems as follows:

Theorem 3.10 (Generation of multi-mode bosonic evolution systems) *Let $(\mathcal{L}_s, \mathcal{D}(\mathcal{L}_s))_{s\in[0,\infty)}$ be a family of operators that fulfills Assumption 7 and Assumption 8. Then $(\overline{\mathcal{L}}_s, \mathcal{D}(\overline{\mathcal{L}}_s))_{s\in\mathbb{R}_+}$ gives rise to a unique evolution system $(\mathcal{P}_{t,s})_{0\leq s\leq t}$ on $W^{\boldsymbol{k}_r,1}$ for all $r \in \mathbb{N}$ with the following properties: for $\boldsymbol{k}_{r'}$ with $\min_{j=1,\ldots,m} |k_{j,r} - k_{j,r'}| \geq d$*

1. *$\mathcal{P}_{t,s}(W^{\boldsymbol{k}_{r'},1}) \subseteq W^{\boldsymbol{k}_{r'},1}$ for all $0 \leq s \leq t$;*
2. *For any $x \in W^{\boldsymbol{k}_{r'},1}$, the family $(\mathcal{P}_{t,s}(x))_{0\leq s\leq t}$ is the unique solution to the initial value problem*

$$\frac{d}{dt}x(t) = \overline{\mathcal{L}}_t(x(t)) \qquad t \in [s,\infty),\ x(s) = x\,. \tag{45}$$

For $\boldsymbol{k} = 0$ as a special case, we get that the evolution system is contractive and trace-preserving.

4 Examples of Sobolev preserving semigroups

In this section, we consider two classes of examples of practical relevance in quantum information processing for which Assumption 1 (or 3,5,7) trivially holds and derive Assumption 2 (or 4, 6, 8). Particular care will be given to finding time-independent upper bounds on the $W^{k,1} \to W^{k,1}$ norm of the semigroup. For this, the overall strategy is as follows: given the generator $(\mathcal{L}, \mathcal{T}_f)$, we prove that there are coefficients $\mu_{k_r} \geq 0, c_{k_r} > 0$ for a divergent sequence $\{k_r\}_{r\in\mathbb{N}}$ such that for all state $\rho \in \mathcal{T}_f$

$$\begin{aligned} \operatorname{tr}\Big[\mathcal{L}(\rho)(N+\mathbb{1})^{k_r/2}\Big] &\leq -c_{k_r} \operatorname{tr}\Big[\rho(N+\mathbb{1})^{k_r/2}\Big] + \mu_{k_r} \\ &\leq (\mu_{k_r} - c_{k_r}) \operatorname{tr}\Big[\rho(N+\mathbb{1})^{k_r/2}\Big], \end{aligned} \tag{46}$$

where we have used $\operatorname{tr}\Big[\rho(N+\mathbb{1})^{k_r/2}\Big] \geq \operatorname{tr}[\rho] = 1$ in the second inequality. Then, Theorem 3.1 can be applied, which shows that for all $k \in \mathbb{R}_+$, the closure of $(\mathcal{L}, \mathcal{T}_f)$ generates a positivity preserving C_0-semigroup $(\mathcal{P}_t)_{t\geq 0}$ on $W^{k,1}$. In the case $k \in \{k_r\}_{r\in\mathbb{N}}$:

$$\|\mathcal{P}_t(x)\|_{W^{k,1}} \leq e^{|\mu_k - c_k|\, t} \|x\|_{W^{k,1}}\,. \tag{47}$$

for all $x \in W^{k,1}$. The bounds for the intermediate values of k can be obtained using Lemma E.4. One can strengthen the above bounds using Equation (46) as follows:

Proposition 4.1 *Let $(\mathcal{L}, \mathcal{T}_f)$ be an operator satisfying Assumption 1 and Equation (46). Then, for all $k \in \mathbb{N}$, the closure of $(\mathcal{L}, \mathcal{T}_f)$ generates a positivity preserving C_0-semigroup $(\mathcal{P}_t)_{t\geq 0}$ on $W^{k,1}$. For all $r \in \mathbb{N}$ and all states $\rho \in W^{k,1}$,*

$$\|\mathcal{P}_t(\rho)\|_{W^{k_r,1}} \leq \max\left\{\|\rho\|_{W^{k_r,1}}, \frac{\mu_{k_r}}{c_{k_r}}\right\}.$$

For a general $k \in \mathbb{R}_+$ and $x \in W^{k,1}$ one obtains

$$\|\mathcal{P}_t(x)\|_{W^{k_r,1}} \leq \gamma_k \|x\|_{W^{k_r,1}}\,, \tag{48}$$

where $\gamma_k = \max\{1, \frac{\mu_k}{c_k}\}$ for $k \in \{k_r\}_{r\in\mathbb{N}}$ and an interpolated time-independent constant in all other cases. Note that for $k > 0$ and $\rho \in W^{k,1}$ there exists a sequence $\{t_n\}_{n\in\mathbb{N}}$, such that

$$\lim_{t_n\to\infty} \mathcal{P}_{t_n}(\rho) = \overline{\rho}$$

for $\overline{\rho} \in W^{k,1}$. Similar conclusions hold in multi-mode as well as time-dependent settings.

Proof. By assumption, Theorem 3.1 shows that the closure of $(\mathcal{L}, \mathcal{T}_f)$ defines a positivity preserving, quasi-contractive semigroup $(\mathcal{P}_t)_{t\geq 0}$. Moreover, for $k \in \{k_r\}_{r\in\mathbb{N}}$, $\rho(t) \coloneqq \mathcal{P}_t(\rho)$

$$\begin{aligned} \frac{d}{dt}\|\rho(t)\|_{W^{k,1}} &= \operatorname{tr}\Big[\mathcal{L}(\rho(t))(N+\mathbb{1})^{k/2}\Big] \\ &\leq -c_k \operatorname{tr}\Big[\rho(t)(N+\mathbb{1})^{k/2}\Big] + \mu_k \\ &= -c_k\|\rho\|_{W^{k,1}} + \mu_k\,. \end{aligned}$$

Thus, for $\|\rho(t)\|_{W^{k,1}} \geq \frac{\mu_k}{c_k}$, we have $\frac{d}{dt}\|\rho(t)\|_{W^{k,1}} \leq 0$, which concludes the bound. Using the positivity preserving property of the semigroup and that $\|\cdot\|_1 \leq \|\cdot\|_{W^{k,1}}$ one can lift the bound to Equation (48) for general $x \in W^{k,1}$ and Theorem 2.14 allows us to conclude

$$\|\mathcal{P}_t(x)\|_{W^{k_r,1}} \leq \gamma_k \|x\|_{W^{k_r,1}}$$

extend to all $k \in \mathbb{R}_+$. Finally, for every $k > 0$, every sequence $n \to \mathcal{P}_{t_n}(\rho)$ is uniformly bounded in $W^{k,1}$ so that the compact embedding shows that there exists a converging subsequence in $W^{k-\varepsilon,1}$ for ε suitably chosen, which is also converging in $W^{k,1}$. This finishes the proof. □

To achieve the inequality stated in Equation (46), we will make heavy use of the following simple commutation relations: given a real-valued function $f : \mathbb{N} \to \mathbb{R}$,

$$\begin{aligned} af(N + j\mathbb{1}) &= f(N + (j+1)\mathbb{1})a, & a^\dagger 1_{>j} f(N - j\mathbb{1}) &= f(N - (j+1)\mathbb{1})a^\dagger 1_{>j}\,, \\ f(N - j\mathbb{1})a\, 1_{>j} &= af(N - (j+1)\mathbb{1})1_{>j}, & f(N + j\mathbb{1})a^\dagger &= a^\dagger f(N + (j+1)\mathbb{1})\,, \end{aligned} \tag{49}$$

where the operators above are defined e.g. on $\mathcal{H}_f$. We also use the canonical commutation relation to write $(a^\dagger)^l a^l$ as a function of N (see Lemma B.2):

$$\begin{aligned} (a^\dagger)^l a^l &= (N - (l-1)\mathbb{1})(N - (l-2)\mathbb{1}) \cdots (N - \mathbb{1})N \\ a^l (a^\dagger)^l &= (N + \mathbb{1})(N + 2\mathbb{1}) \cdots (N + (l-1)\mathbb{1})(N + l\mathbb{1})\,. \end{aligned}$$

In the following, we adopt the notations:

$$\mathcal{L}[L] \coloneqq L(\cdot)L^\dagger - \frac{1}{2}\{L^\dagger L, \cdot\} \qquad \text{and} \qquad \mathcal{H}[H] \coloneqq -i[H, \cdot]\,.$$

Although this notation collides with the one for the Hilbert space, the meaning can always be deduced from context.

4.1 Quantum Ornstein Uhlenbeck semigroup

We start with the generator of the quantum Ornstein Uhlenbeck semigroup [18, 13] defined by

$$\mathcal{L}_{\mathrm{qOU}} = \lambda^2 \mathcal{L}[a] + \mu^2 \mathcal{L}[a^\dagger] \tag{50}$$

for $\mu, \lambda \geq 0$. Given an suitably domain $\mathcal{D}(\mathcal{L}_{\mathrm{qOU}})$, the operator $(\overline{\mathcal{L}}_{\mathrm{qOU}}, \mathcal{D}(\mathcal{L}_{\mathrm{qOU}}))$ is known to generate a quantum dynamical semigroup $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$. Here, we further show that the quantum Ornstein Uhlenbeck semigroup defines a semigroup on all $W^{k,1}$. This is the topic of the following lemma:

Lemma 4.2 *Let $(\mathcal{L}_{\mathrm{qOU}}, \mathcal{T}_f)$ be the generator of the quantum Ornstein Uhlenbeck semigroup and $k \in \mathbb{N}$. Then, there exist constants μ_k explicated in (52) such that, for all states $\rho \in \mathcal{T}_f$,*

$$\operatorname{tr}\Big[\mathcal{L}_{\mathrm{qOU}}(\rho)(N + \mathbb{1})^{\frac{k}{2}}\Big] \leq \begin{cases} \frac{k}{4}(\mu^2 - \lambda^2) \operatorname{tr}\left[\rho(N+\mathbb{1})^{k/2}\right] + \mu_k & \lambda > \mu \\ \frac{k}{2}(2\mu^2 + k) \operatorname{tr}\left[\rho(N+\mathbb{1})^{k/2}\right] & \lambda \leq \mu \end{cases}.$$

Therefore, the semigroup $e^{t\mathcal{L}_{\mathrm{qOU}}}$ is a Sobolev and positivity preserving quantum Markov semigroup satisfying for all states $\rho \in W^{k,1}$

$$\|e^{t\mathcal{L}_{\mathrm{qOU}}}(\rho)\|_{W^{k,1}} \leq \begin{cases} \max\left\{\|\rho\|_{W^{k,1}},\, \frac{4\mu_k}{k(\mu^2 - \lambda^2)}\right\} & \lambda > \mu \\ e^{t\frac{k}{2}(2\mu^2+k)}\|\rho\|_{W^{k,1}} & \lambda \leq \mu \end{cases}.$$

Proof. We consider $\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N))$ where $f(x) = (x+1)^{k/2} 1_{x \geq -1}$. By Equation (49),

$$\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N)) = \lambda^2 N(f(N - \mathbb{1}) - f(N)) + \mu^2 (N + \mathbb{1})(f(N + \mathbb{1}) - f(N))\,.$$

Note that the case $k=0$ follows from the GKLS form and $k=2$ is by definition of f trivially given by $(\mu^2-\lambda^2)N+\mathbb{1}$. Next, we define an auxiliary function which will also prove useful in the following proofs:

$$g_l(x)=\begin{cases} f(x)-f(x-l) & x\geq l;\\ f(x) & l>x\geq 0;\\ 0 & 0>x\,. \end{cases} \tag{51}$$

It allows us to redefine $\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N))$ by

$$\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N))=-\lambda^2Ng_1(N)+\mu^2(N+\mathbb{1})g_1(N+\mathbb{1})\,.$$

Then, applying Lemma C.2 to the spectral decomposition of the polynomial in the number operator above, we get

$$\begin{aligned}&\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N))\\ &\leq\frac{k}{2}(\mu^2-\lambda^2)(N+\mathbb{1})^{k/2}+\lambda^2\frac{k}{2}(N+\mathbb{1})^{k/2-1}+1_{k\geq3}(N+\mathbb{1})^{k/2-2}\frac{k^2}{8}+\mu^2\frac{2-k}{2}\,|0\rangle\langle0|\\ &\leq\frac{k}{2}(\mu^2-\lambda^2)(N+\mathbb{1})^{k/2}+\frac{k}{2}\left(\lambda^2+\mu^2+k\right)(N+\mathbb{1})^{k/2-1}\end{aligned}$$

where we separated the vacuum state from the rest of the decomposition. Note that this bound can also be used when $k=1$ since $(N+\mathbb{1})^{-1/2}$ is then bounded by 1. Therefore, we assume $k\geq3$ in the following and start with the case $\lambda>\mu$ so that the leading order is negative. Then, we use half of the latter to bound the other terms by a constant. This is done by the following classical optimization

$$\sup_{x\geq0}\left(-x^\nu+cx^{\nu-1}\right)=c^\nu\left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right)$$

for $\nu\geq1$ and $c\geq0$ defined as

$$c=2\frac{\lambda^2+\mu^2+k}{\lambda^2-\mu^2}\qquad\text{and}\qquad\nu=\frac{k}{2}\,.$$

Then,

$$\begin{aligned}\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N))&\leq\frac{k}{4}(\mu^2-\lambda^2)(N+\mathbb{1})^{k/2}+c^\nu\left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right)\\ &=:\frac{k}{4}(\mu^2-\lambda^2)(N+\mathbb{1})^{k/2}+\mu_k^{\lambda>\mu}\end{aligned} \tag{52}$$

The second case is $\lambda\leq\mu$, which can be easily upper bounded by

$$\mathcal{L}^\dagger_{\mathrm{qOU}}(f(N))\leq\frac{k}{2}(2\mu^2+k)(N+\mathbb{1})^{k/2}\quad.$$

This completes the proof of the statement by Theorem 3.1 and Proposition 4.1. □

4.2 Photon-dissipation and CAT qubits

Next, we consider a family of Lindbladians that has been recently studied in the setting of error correction with continuous variable quantum systems. For an introduction to the field, we refer the interested reader to the following lecture notes [57, 32]. The abstract idea here is that the code-space is continuously protected by a dissipative evolution, i.e. an evolution which is exponentially converging for $t \to \infty$ to an invariant subspace — the code-space. This behavior is achieved through the so-called l-photon dissipation generated for $\kappa > 0$ and $\alpha \in \mathbb{C}$ by

$$\kappa\mathcal{L}[a^l - \alpha^l]\,, \tag{53}$$

where we sometimes omit the identity so that $\alpha^l := \alpha^l \mathbb{1}$ in what follows. The invariant subspace (code-space) to which the evolution is exponentially converging [5] is defined by

$$\mathcal{C}_l := \operatorname{span}\left\{|\alpha_1\rangle\!\langle\alpha_2| \,:\, \alpha_1, \alpha_2 \in \left\{\alpha e^{\frac{i2\pi j}{l}} \mid j \in \{0, ..., l-1\}\right\}\right\}\,, \tag{54}$$

where $|\alpha\rangle$ denotes the coherent state

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle\,.$$

and satisfies $a|\alpha\rangle = \alpha|\alpha\rangle$ by definition.

Besides the l-photon dissipation, we consider the CAT qubit error correction protocol introduced in [33] associated with the 2-photon dissipation and code-space $\mathcal{C}_2$ and with corresponding universal gate-set generated by the following generators: for some parameters $T, \kappa, \varepsilon > 0$,

Identity-gate:

$$\kappa\mathcal{L}[a^2 - \alpha^2] \tag{55}$$

$Z(\theta)$-gate:

$$\kappa\mathcal{L}[a^2 - \alpha^2] + \varepsilon\mathcal{H}[a + a^\dagger] \tag{56}$$

X-gate:

$$\kappa\mathcal{L}[a^2 - e^{2i\pi t/T}\alpha^2] \tag{57}$$

CNOT-*gate:*

$$\kappa\mathcal{L}[a^2 - \alpha^2] + \varepsilon\mathcal{L}[b^2 - \alpha^2 - \frac{\alpha}{2}(1 - e^{2i\pi t/T})(a - \alpha)] \tag{58}$$

Toffoli-gate:

$$\kappa\mathcal{L}[a^2 - \alpha^2] + \kappa\mathcal{L}[b^2 - \alpha^2] + \varepsilon\mathcal{L}[c^2 - \alpha^2 + \frac{1}{4}(1 - e^{2i\pi t/T})(ab - \alpha(a + b) + \alpha^2)]\,. \tag{59}$$

Note that the CNOT gate operates on two modes, while the Toffoli gate acts on three modes, with the annihilation and creation operators on the second mode denoted by b and $b^\dagger$, and on the third mode by c and $c^\dagger$. Furthermore, from a mathematical standpoint, the implementation remains somewhat unclear, as the sum gates are constructed using an adiabatic limit, for which, to our knowledge, rigorous error bounds have yet to be established.

In the following, we prove that the above operators generate Sobolev-preserving quantum dynamical semigroups, with the exception of the Toffoli gate. Due to its more complicated structure, we leave the analysis of the latter to future work. We start by proving that the l-photon dissipation satisfies Equation (46), and therefore that it generates a Sobolev preserving semigroup by Proposition 4.1.

Lemma 4.3 (l-photon dissipation) *For any $k \geq 1$, $l \geq 2$, $\alpha \in \mathbb{C}$ and any state $\rho \in \mathcal{T}_f$,*

$$\operatorname{tr}\left[\mathcal{L}[a^l - \alpha^l](\rho)(N + \mathbb{1})^{k/2}\right] \leq -\frac{l}{2}\operatorname{tr}\left[\rho\,(N + \mathbb{1})^{\nu}\right] + \frac{l}{2}\mu_k^{(l)} \leq -\frac{l}{2}\operatorname{tr}\left[\rho\,(N + \mathbb{1})^{k/2}\right] + \frac{l}{2}\mu_k^{(l)}\,,$$

where $\mu_k^{(l)} = \Delta_l^\nu \left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right)$ *with* $\nu = l + \frac{k}{2} - 1$ *and* $\Delta_l = (l+1)l + 2|\alpha|^l k l^{k/2-1}\sqrt{l!}$. *Therefore,* $\mathcal{L}_l := \mathcal{L}[a^l - \alpha^l]$ *generates a Sobolev and positivity preserving quantum Markov semigroup satisfying for all states* $\rho \in W^{k,1}$

$$\|e^{t\mathcal{L}_l}(\rho)\|_{W^{k,1}} \leq \max\left\{\|\rho\|_{W^{k,1}}, \mu_k^{(l)}\right\}.$$

Proof. By Equation (49), we have for $f(x) = (x+1)^{k/2}1_{x\geq -1}$:

$$\begin{aligned}\mathcal{L}[a_1^l - \alpha^l]^\dagger(f(N)) &= (a^\dagger)^l f(N) a^l - \frac{1}{2}\Big((a^\dagger)^l a^l f(N) + f(N)(a^\dagger)^l a^l\Big) \\ &\quad + \frac{1}{2}(\overline{\alpha}^l a^l f(N) - \overline{\alpha}^l f(N) a^l + \alpha^l f(N)(a^\dagger)^l - \alpha^l (a^\dagger)^l f(N)) \\ &= (a^\dagger)^l a^l \Big(f(N - l\mathbb{1}) - f(N)\Big) \\ &\quad + \frac{1}{2}\left[\overline{\alpha}^l a^l \Big(f(N) - f(N-l\mathbb{1})\Big) + \alpha^l \Big(f(N) - f(N - l\mathbb{1})\Big)(a^\dagger)^l\right].\end{aligned}$$

In what follows, we use the function defined in Equation (91)

$$g_l(x) = \begin{cases} f(x) - f(x-l) & x \geq l; \\ f(x) & l > x \geq 0; \\ 0 & 0 > x\,. \end{cases}$$

Using the canonical commutation relation to write $(a^\dagger)^l a^l$ as a function of N (cf. Lemma B.2) and with help of the notation

$$N_k[r:j] := (N_k + r\mathbb{1})\cdots(N_k + j\mathbb{1}) \tag{60}$$

with the convention $N_k[r:j] = \mathbb{1}$ whenever $r > j$, we thus have that

$$\operatorname{tr}\left[\rho\,\mathcal{L}[a^l - \alpha^l]^\dagger(f(N))\right] = -\operatorname{tr}\left[\rho\, N[-l+1:0] g_l(N)\right] + \frac{1}{2}\operatorname{tr}\left[\rho\,(\overline{\alpha}^l a^l g_l(N) + \alpha^l g_l(N)(a^\dagger)^l)\right],$$

Since g_l is positive and increasing, the last term above can be upper bounded by Lemma B.3,

$$\begin{aligned}\frac{1}{2}\operatorname{tr}\left[\rho(\overline{\alpha}^l a^l g_l(N) + \alpha^l g_l(N)(a^\dagger)^l)\right] &\leq |\alpha|^l \operatorname{tr}\Big[\rho\, g_l(N + l\mathbb{1})\sqrt{N[1:l]}\Big] \\ &\overset{(1)}{\leq} |\alpha|^l \frac{k l^{k/2}}{2} \operatorname{tr}\Big[\rho\,(N+\mathbb{1})^{k/2-1}\sqrt{N[1:l]}\Big] \\ &\leq |\alpha|^l k l^{k/2}\sqrt{l!}\,\operatorname{tr}\Big[\rho\,(N+\mathbb{1})^{k/2-1+\frac{l}{2}}\Big].\end{aligned}$$

In (1) above, we used Lemma C.2 for the bound

$$g_l(N + l\mathbb{1}) \leq \frac{kl}{2}(N + l\mathbb{1})^{k/2-1} \leq \frac{k l^{k/2}}{2}(N+\mathbb{1})^{k/2-1}.$$

Therefore, we have proven that

$$\begin{aligned}\operatorname{tr}\left[\rho\mathcal{L}[a^l - \alpha^l]^\dagger(f(N))\right] \leq &-\operatorname{tr}\left[\rho\, N[-l+1:0]\, g_l(N)\right] \\ &+ |\alpha|^l k l^{k/2}\sqrt{l!}\,\operatorname{tr}\Big[\rho\,(N+\mathbb{1})^{k/2-1+\frac{l}{2}}\Big].\end{aligned}$$

Next, we upper bound the first term above

$$\begin{aligned}\operatorname{tr}\left[\rho\, N[-l+1:0]\, g_l(N)\right] &\overset{(3)}{\geq} l \operatorname{tr}\left[\rho\, N[-l+1:0](N+\mathbb{1})^{k/2-1}\right] \\ &\overset{(4)}{\geq} l \operatorname{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-1}\right] - \frac{(l+1)l^2}{2}\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-2}\right].\end{aligned}$$

In (3), we used Lemma C.2 below with the fact that $N[-l+1:0]$ is supported on the Fock states $|n\rangle$ with $n \geq l-1$; in (4) we used that

$$\begin{aligned}N[-l+1:0] &= \sum_{n\geq 0}(n-l+1)\dots n\ |n\rangle\langle n| \\ &= \sum_{n\geq l}(n-l+1)\dots n\ |n\rangle\langle n| \\ &\overset{(5)}{\geq} \sum_{n\geq l}\left((n+1)^l - \frac{(l+1)l}{2}(n+1)^{l-1}\right)|n\rangle\langle n| \\ &\geq (N+\mathbb{1})^l - \frac{(l+1)l}{2}(N+\mathbb{1})^{l-1},\end{aligned}$$

where (5) comes from Lemma C.3 below, whereas the last inequality follows from the fact that $l \geq 2$. To sum up, we showed that

$$\begin{aligned}\operatorname{tr}\left[\mathcal{L}[a^l-\alpha^l](\rho)(f(N))\right] &\leq -l \operatorname{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-1}\right] + \frac{(l+1)l^2}{2}\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-2}\right] \\ &\quad + |\alpha|^l k l^{k/2}\sqrt{l!}\operatorname{tr}\Big[\rho\,(N+\mathbb{1})^{k/2-1+\frac{l}{2}}\Big] \\ &\leq -l \operatorname{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-1}\right] \\ &\quad + \frac{l}{2}\Big(\underbrace{(l+1)l + 2|\alpha|^l k l^{k/2-1}\sqrt{l!}}_{=:\Delta_l}\Big)\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-2}\right]\end{aligned} \tag{61}$$

where we used again that $l \geq 2$ in the last inequality. Half of the leading order term can be used to control the second term by a constant. For that, we use the spectral decomposition of the operator N so that the above problem can be reduced to the following simple optimization:

$$\sup_{x\geq 0}\left(-x^\nu + \Delta_l x^{\nu-1}\right) = \Delta_l^\nu\left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right) \tag{62}$$

for $\nu \geq 1$ defined as

$$\nu = l + \frac{k}{2} - 1\,.$$

The result follows after invoking Proposition 4.1. □

Remark 7. The single-mode bound proved above can be generalized to the multi-mode setting, with generated given for some $\alpha_j \in \mathbb{C}$, $j \in [m]$, by

$$\mathcal{L}_l^{(m)} := \sum_{j=1}^m \mathcal{L}[a_j^l - \alpha_j^l]\,.$$

Since all the bounds used in the proof of Lemma 4.3 were derived at the operator level, we directly get for $\mathbf{k} \in \mathbb{N}^m$

$$\operatorname{tr}\left[\mathcal{L}_l^{(m)}(\rho)(N+\mathbb{1})^{\mathbf{k}/2}\right] \leq \sum_{i=1}^m -\frac{l}{2}\operatorname{tr}\left[\rho\,(N_i+\mathbb{1})^{l-1}(N+\mathbb{1})^{\mathbf{k}/2}\right] + \mu_{k_i}^{(l)}\operatorname{tr}\left[\rho\prod_{j\neq i}(N_j+\mathbb{1})^{k_j/2}\right]$$

For later references, we single out the case $l = 2$.

Corollary 4.4 (2-photon dissipation) *For any integers $k \geq 1$, $\alpha \in \mathbb{C}$ and any state $\rho \in \mathcal{T}_f$,*

$$\operatorname{tr}\left[\mathcal{L}[a^2-\alpha^2](\rho)(N+\mathbb{1})^{k/2}\right] \leq -\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2}\right] + \mu_k^{(2)}$$

where $\mu_k^{(2)} = (\Delta_k^{(2)})^\nu \left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right)$ with $\nu = \frac{k}{2}+1$ and $\Delta_2 = 6 + 2\sqrt{2!}|\alpha|^2 k 2^{k/2-1}$. Therefore, $\mathcal{L}_2 \coloneqq \mathcal{L}[a^2-\alpha^2]$ generates a Sobolev and positivity preserving quantum Markov semigroup which satisfies for all states ρ in $W^{k,1}$

$$\|e^{t\mathcal{L}_2}(\rho)\|_{W^{k,1}} \leq \max\left\{\|\rho\|_{W^{k,1}}, \mu_k^{(2)}\right\}.$$

From the above bounds, we directly get the property of Sobolev preservation for the X-gate:

Corollary 4.5 (X-gate) *For any $T > 0$, $\alpha \in \mathbb{C}$, $k \in \mathbb{N}$ and all states $\rho \in \mathcal{T}_f$,*

$$\operatorname{tr}\Big[\mathcal{L}[a^2 - e^{2i\pi t/T}\alpha^2](\rho)(N+\mathbb{1})^{k/2}\Big] \leq -\operatorname{tr}\Big[\rho(N+\mathbb{1})^{k/2}\Big] + \mu_k^{(2)},$$

where $\mu_k^{(2)}$ is defined in Lemma 4.4. Therefore, $\mathcal{L}[a^2 - e^{2i\pi t/T}\alpha^2]$ generates a Sobolev and positivity preserving quantum evolution system $\mathcal{P}_{t,t_0}$ which satisfies for all states $\rho \in W^{k,1}$

$$\|\mathcal{P}_{t,t_0}(\rho)\|_{W^{k,1}} \leq \max\left\{\|\rho\|_{W^{k,1}}, \mu_k^{(2)}\right\}.$$

Proof. The statement directly follows from Lemma 4.4 and $|e^{2i\pi t/T}\alpha^2| = |\alpha^2|$ □

Since Equation (46), which implies Assumption 2 (or 8), is linear in the supposed generators and the leading order found in Equation (61) has the ability to suppress smaller terms, certain Hamiltonians can be regularized in the sense of Equation (46) by adding an l-photon dissipation. Especially, we consider a Hamiltonian of degree $d_H = 2(l-1)$ with the structure: for $\lambda_{i,j} \in \mathbb{C}$ with $\max_{i,j}|\lambda_{i,j}| = \Lambda$,

$$H = p(a, a^\dagger) = \sum_{\substack{i \leq j \\ i+j \leq d_H}} \lambda_{i,j} a^i (a^\dagger)^j + \overline{\lambda_{i,j}} a^j (a^\dagger)^i\,. \tag{63}$$

Note that any monomial in $a, a^\dagger$ of degree at most d_H can be achieved from the representation above thanks to the CCR.

Lemma 4.6 *Let $\mathcal{L}_l \coloneqq \mathcal{L}[a^l - \alpha^l]$, $\alpha \in \mathbb{C}$, be the l-photon dissipation and H as in (63). Then, for all states $\rho \in \mathcal{T}_f$*

$$\operatorname{tr}\Big[(\mathcal{L}_l + \mathcal{H}[H])(\rho)(N+\mathbb{1})^{k/2}\Big] \leq -\frac{l}{2}\operatorname{tr}\Big[\rho(N+\mathbb{1})^{k/2}\Big] + \frac{l}{2}\mu_k\,. \tag{64}$$

for $\mu_k, \nu \geq 1$ defined by

$$\mu_k = c^\nu \left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right) \quad \text{with} \quad c = (l+1)l + 2|\alpha|^l k l^{k/2-1}\sqrt{l!} + \Lambda(2l)^{k/2}\sqrt{(2l)!}\,, \quad \nu = l + \frac{k}{2} - 1\,.$$

Therefore, $\mathcal{L}_l + \mathcal{H}[H]$ generates a Sobolev and positivity preserving quantum Markov semigroup which satisfies for all states $\rho \in W^{k,1}$

$$\|e^{t(\mathcal{L}_l + \mathcal{H}[H])}(\rho)\|_{W^{k,1}} \leq \max\left\{\|\rho\|_{W^{k,1}}, \mu_k\right\}. \tag{65}$$

Proof. We reuse the bound given in Equation (61):

$$\mathrm{tr}\Big[\rho\mathcal{L}_l^\dagger(f(N))\Big] \leq -l\,\mathrm{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-1}\right] + \frac{l}{2}\Delta_l\;\mathrm{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-2}g_l(N)\right],$$

where $f(x) = (x+1)^{k/2}1_{x\geq -1}$ and $\Delta_l = (l+1)l + 2|\alpha|^l k l^{k/2-1}\sqrt{l!}$. To upper bound

$$\mathrm{tr}\Big[\mathcal{H}[H](\rho)(N+\mathbb{1})^{k/2}\Big] = i\,\mathrm{tr}\Big[\rho[(N+\mathbb{1})^{k/2}, H]\Big],$$

we define g_u similarly to Equation (51) by

$$g_u(x) = \begin{cases} f(x) - f(x-u) & x \geq u-1; \\ f(x) & u-1 > x \geq 0; \\ 0 & 0 > x\,. \end{cases}$$

For $d_H = 0$ the bound is trivial, so we assume $d_H \geq 1$. Then, we compute

$$\begin{aligned}
&i[f(N), H] \\
&= i \sum_{\substack{0\leq j<i\\0<i+j\leq d_H}} f(N)(\lambda_{i,j}(a^\dagger)^i a^j + \overline{\lambda_{i,j}}(a^\dagger)^j a^i) - (\lambda_{i,j}(a^\dagger)^i a^j + \overline{\lambda_{i,j}}(a^\dagger)^j a^i)f(N) \\
&= i \sum_{\substack{0\leq j<i\\0<i+j\leq d_H}} \lambda_{i,j} f(N)N[-i+1:-i+j](a^\dagger)^{i-j} + \overline{\lambda_{i,j}}a^{i-j}f(N-i+j)N[-i+1:-i+j] \\
&\qquad\qquad - \lambda_{i,j}N[-i+1:-i+j]f(N-i+j)(a^\dagger)^{i-j} - \overline{\lambda_{i,j}}a^{i-j}N[-i+1:-i+j]f(N) \\
&= i \sum_{\substack{0<r\leq i\\0<2i-r\leq d_H}} -\overline{\lambda_{i,i-r}}a^r N[-i+1:-r]g_r(N) + \lambda_{i,i-r}g_r(N)N[-i+1:-r](a^\dagger)^r \\
&\overset{(1)}{\leq} \sum_{\substack{0<r\leq i\\0<2i-r\leq d_H}} 2|\lambda_{i,i-r}|\sqrt{(N+\mathbb{1})\cdots(N+r\mathbb{1})}g_r(N+r\mathbb{1})N[r-i+1:0] \\
&\overset{(2)}{\leq} \sum_{\substack{0<r\leq i\\0<2i-r\leq d_H}} 2\sqrt{r!}|\lambda_{i,i-r}|g_r(N+r\mathbb{1})(N+\mathbb{1})^{i-r/2}
\end{aligned}$$

$$\begin{aligned}
i\,\mathrm{tr}\Big[[H,\rho](N+\mathbb{1})^{k/2}\Big] &\leq 2\Lambda\sum_{i=1}^{d_H}\sum_{r=1}^{i}\sqrt{r!}\,\mathrm{tr}\Big[\rho g_r(N+r)(N+\mathbb{1})^{i-r/2}\Big] \\
&\overset{(3)}{\leq} 2\Lambda\sum_{i=1}^{d_H}\sum_{r=1}^{i}\sqrt{r!}r^{k/2-1}\,\mathrm{tr}\Big[\rho(N+\mathbb{1})^{k/2+i-r/2-1}\Big] \\
&\leq \Lambda(d_H+1)d_H\sqrt{d_H!}d_H^{k/2-1}\,\mathrm{tr}\Big[\rho(N+\mathbb{1})^{k/2+d_H/2-1}\Big],
\end{aligned}$$

where we used Lemma C.2 in (3). As the above function is monotone in d_H we can w.l.o.g assume $d_H = 2(l-1)$ and conclude

$$\begin{aligned}
\mathrm{tr}[(\mathcal{L}_l + \mathcal{H}[H])(f(N))] \leq -l\,\mathrm{tr}\left[\rho\,(N+\mathbb{1})^{l+k/2-1}\right] \\
+ \frac{l}{2}\left(\Delta_l + \Lambda(2l)^{k/2}\sqrt{(2l)!}\right)\mathrm{tr}\Big[\rho(N)(N+\mathbb{1})^{l+k/2-2}\Big].
\end{aligned}$$

The same optimization as in Equation (62) provides inequality (64). Inequality (65) follows after invoking Proposition 4.1. □

The case $l = 2$ deals with the sum of a 2-photon dissipation with the displacement operator $i[a+a^\dagger, \cdot]$ used in the construction of the $Z(\theta)$-gate [50]. In this specific case, one can improve the error bound in the following way:

Lemma 4.7 ($Z(\theta)$-gate) *For any state $\rho \in \mathcal{T}_f$, $\alpha \in \mathbb{C}$, $\varepsilon > 0$ and $k \in \mathbb{N}$*

$$\operatorname{tr}\Big[(\varepsilon\mathcal{H}[a+a^\dagger] + \mathcal{L}[a^2+\alpha^2])(\rho)(N+\mathbb{1})^{k/2}\Big] \leq -\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2}\right] + \mu_k\,.$$

where $\mu_k \geq 0$ is defined by

$$\mu_k = (\Delta_2 + \varepsilon 4k)^\nu \left(\frac{(\nu-1)^{\nu-1}}{\nu^\nu}\right) \qquad \textit{with} \qquad \nu = \frac{k}{2}+1\,.$$

Therefore, $\varepsilon\mathcal{H}[a+a^\dagger] + \mathcal{L}[a^2+\alpha^2]$ generates a Sobolev and positivity preserving quantum Markov semigroup which satisfies for all states $\rho \in W^{k,1}$

$$\|e^{t(\varepsilon\mathcal{H}[a+a^\dagger]+\mathcal{L}[a^2+\alpha^2])}(\rho)\|_{W^{k,1}} \leq \max\Big\{\|\rho\|_{W^{k,1}}, \mu_k\Big\}\,. \tag{66}$$

Proof. By Equation (61) in Lemma 4.3,

$$\begin{aligned}\operatorname{tr}\left[\mathcal{L}[a^2-\alpha^2](\rho)(f(N))\right] \leq{}& -2\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2+1}\right] \\ &+ \Big(\underbrace{6+2|\alpha|^2k2^{k/2-1}\sqrt{2}}_{=:\Delta_2}\Big)\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2}\right]\end{aligned}$$

where $f(x) = (x+1)^{k/2}1_{x\geq -1}$. Next, by Equation (49), Lemma B.3 and Lemma C.2, we have that

$$\begin{aligned}\operatorname{tr}\Big[\mathcal{H}[a+a^\dagger](\rho)f(N)\Big] &= i\operatorname{tr}\Big[\rho\left(f(N)(a+a^\dagger) - (a+a^\dagger)f(N)\right)\Big] \\ &= \operatorname{tr}\Big[\rho\left(-iag_1(N) + ig_1(N)a^\dagger\right)\Big] \\ &\leq 2\operatorname{tr}\Big[\rho g_1(N+\mathbb{1})\sqrt{N+\mathbb{1}}\Big] \\ &\leq 2k\operatorname{tr}\Big[\rho(N+\mathbb{1})^{k/2-\frac{1}{2}}\Big]\,,\end{aligned} \tag{67}$$

where we recall that

$$g_1(x) = \begin{cases} f(x) - f(x-1) & x \geq 0; \\ 0 & 0 > x\,. \end{cases} \tag{68}$$

Thus,

$$\begin{aligned}&\operatorname{tr}\left[(\epsilon\mathcal{H}[a+a^\dagger] + \mathcal{L}[a^2-\alpha^2])(\rho)(f(N))\right] \\ &\qquad \leq -2\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2+1}\right] + (\Delta_2+\varepsilon 2k)\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2}\right].\end{aligned}$$

for $\nu \geq 1$ defined as

$$\nu = \frac{k}{2}+1$$

ends the proof of the differential upper bound, and (66) follows from Proposition 4.1. □

Interestingly, Equation (67) shows directly that the displacement operator satisfies Assumption 2 so that Theorem 3.1 can be applied.

Corollary 4.8 *For any state $\rho \in \mathcal{T}_f$, $\alpha \in \mathbb{C}$, $\varepsilon > 0$ and $k \in \mathbb{N}$*

$$\operatorname{tr}\Big[\mathcal{H}[a+a^\dagger](N+\mathbb{1})^{k/2}\Big] \leq 2k\operatorname{tr}\left[\rho\,(N+\mathbb{1})^{k/2}\right]$$

Therefore, $\mathcal{H}[a+a^\dagger]$ generates a Sobolev and positivity-preserving quantum Markov semigroup.

Nevertheless, the above result is not of the form given in Equation (47) so the improvement Equation (48) is not applicable. In the context of bosonic error correction, the projective Zeno effect wr.r.t. the semigroup above and the projection onto the code space (54) helps to understand the $Z(\theta)$-gate mathematically [50]. A quantitative convergence rate can be proven via [51].

Proposition 4.9 (CNOT-gate) *For all* $\boldsymbol{k} \coloneqq (k_1, k_2) \in \mathbb{N}^2$ *such that*

$$32|\alpha|k_1 2^{k_1/2-1/2} \leq k_2 \,,$$

there exists a constant $\mu_{\boldsymbol{k}}$ *such that for all states* $\rho \in \mathcal{T}_f$

$$\operatorname{tr}\Big[\Big(\mathcal{L}[a^2-\alpha^2] + \mathcal{L}[b^2-\alpha^2-\frac{\alpha}{2}(1-e^{2i\pi t/T})(a-\alpha)]\Big)(\rho)(N_1+\mathbb{1})^{k_1/2}(N_2+\mathbb{1})^{k_2/2}\Big] \\ \leq -\frac{1+k_2}{8}\operatorname{tr}\Big[\rho\Big((N_1+\mathbb{1})^{k_1/2}(N_2+\mathbb{1})^{k_2/2}\Big)\Big] + \mu_{\boldsymbol{k}}\,.$$

Therefore, the CNOT-gate generates a Sobolev and positivity preserving quantum Markov semigroup which satisfies for all states $\rho \in W^{\boldsymbol{k},1}$

$$\|\mathcal{P}^{\mathrm{CNOT}}_{t,t_0}(\rho)\|_{W^{\boldsymbol{k},1}} \leq \max\left\{\|\rho\|_{W^{\boldsymbol{k},1}}, \frac{8\mu_{\boldsymbol{k}}}{1+k_2}\right\}. \tag{69}$$

For a general $\boldsymbol{k} \in \mathbb{R}^2_+$ *and* $x \in W^{\boldsymbol{k},1}$ *one obtains*

$$\|\mathcal{P}^{\mathrm{CNOT}}_{t,t_0}(x)\|_{W^{\boldsymbol{k}_r,1}} \leq \gamma_k \|x\|_{W^{\boldsymbol{k}_r,1}}\,,$$

where $\gamma_{\boldsymbol{k}} = \max\{1, \frac{8\mu_k}{1+k_2}\}$ *for* $\boldsymbol{k} \in \{\boldsymbol{k}_r\}_{r\in\mathbb{N}}$ *and an interpolated constant in all other cases.*

Proof. We denote $f(x_1,x_2) = f_1(x_1)f(x_2)$ with $f_1(x_1) = (x_1+1)^{k_1/2}1_{x_1\geq -1}$, $f_2(x_2) = (x_2+1)^{k_2/2}1_{x_2\geq -1}$, and rewrite the CNOT-generator as

$$\mathcal{L} \coloneqq \mathcal{L}[a^2-\alpha^2] + \mathcal{L}[b^2-\alpha^2-\frac{\alpha}{2}(1-e^{2i\pi t/T})(a-\alpha)] = \mathcal{L}[a^2-\alpha^2] + \mathcal{L}[b^2+za+w]$$

where $z \coloneqq -\frac{\alpha}{2}(1-e^{2i\pi t/T})$ and $w \coloneqq -\alpha(z+\alpha)$. As in the previous proofs, we investigate the action of the adjoint on $f(N) \coloneqq f(N_1,N_2)$

$$\operatorname{tr}[\mathcal{L}(\rho)f(N)] = \operatorname{tr}\Big[\rho\mathcal{L}^\dagger(f(N))\Big]\,.$$

We first focus on the second Lindbladian $\mathcal{L}[b^2+za+w]$: we first consider, for $n \coloneqq (n_1,n_2) \in \mathbb{N}^2$,

$$\begin{aligned}
&\mathcal{L}[b^2+za+w]^\dagger(|n\rangle\!\langle n|) \\
&\quad = ((b^\dagger)^2+\overline{z}a^\dagger+\overline{w})\,|n\rangle\!\langle n|\,(b^2+za+w) - \frac{1}{2}\Big\{((b^\dagger)^2+\overline{z}a^\dagger+\overline{w})(b^2+za+w), |n\rangle\!\langle n|\Big\} \\
&\quad = F_1(n)\,|n\rangle\!\langle n| + F_2(n)\,|n_1,n_2+2\rangle\!\langle n_1,n_2+2| + F_3(n)\,|n_1+1,n_2\rangle\!\langle n_1+1,n_2| \\
&\qquad + (F_4(n)\,|n_1,n_2+2\rangle\!\langle n_1+1,n_2| + h.c.) + (F_5(n)\,|n_1+1,n_2-2\rangle\!\langle n| + h.c.) \\
&\qquad + (F_6(n)\,|n_1,n_2-2\rangle\!\langle n| + h.c.) + (F_7(n)\,|n_1-1,n_2\rangle\!\langle n| + h.c.) \\
&\qquad + (F_8(n)\,|n_1,n_2+2\rangle\!\langle n| + h.c.) + (F_9(n)\,|n_1+1,n_2\rangle\!\langle n| + h.c.) \\
&\qquad + (F_{10}(n)\,|n_1-1,n_2+2\rangle\!\langle n| + h.c.)\,,
\end{aligned}$$

where the notation $h.c.$ above stands for Hermitian conjugate, $|n\rangle = 0$ whenever $n \notin \mathbb{N}^2$ by convention, and where

$$
\begin{aligned}
F_1(n) &:= -n_2(n_2-1) - |z|^2 n_1 \\
F_2(n) &:= (n_2+1)(n_2+2) \\
F_3(n) &:= |z|^2(n_1+1) \\
F_4(n) &:= z\sqrt{(n_1+1)(n_2+1)(n_2+2)} \\
F_5(n) &:= -\frac{1}{2}\overline{z}\sqrt{(n_1+1)n_2(n_2-1)} \\
F_6(n) &:= -\frac{1}{2}\overline{w}\sqrt{n_2(n_2-1)} \\
F_7(n) &:= -\frac{1}{2}\overline{w}z\sqrt{n_1} \\
F_8(n) &:= \frac{1}{2}w\sqrt{(n_2+1)(n_2+2)} \\
F_9(n) &:= \frac{1}{2}\overline{z}w\sqrt{n_1+1} \\
F_{10}(n) &:= -\frac{1}{2}z\sqrt{n_1(n_2+1)(n_2+2)}\,.
\end{aligned}
$$

In the next step, we regroup the 17 terms into terms differing only by a shift: *Case 0:* Diagonal terms, involving F_1, F_2 and F_3,

$$C_0(n) := F_1(n)\,|n\rangle\langle n| + F_2(n)\,|n_1,n_2+2\rangle\langle n_1,n_2+2| + F_3(n)\,|n_1+1,n_2\rangle\langle n_1+1,n_2|\;.$$

Case 1: Terms of the form $|n_1+1,n_2-2\rangle\langle n_1,n_2|$, involving $\overline{F}_4$, F_5 and $\overline{F}_{10}$,

$$
\begin{aligned}
C_1(n) := \overline{F}_4(n)\,&|n_1+1,n_2\rangle\langle n_1,n_2+2| \\
&+ F_5(n)\,|n_1+1,n_2-2\rangle\langle n| + \overline{F}_{10}(n)\,|n\rangle\langle n_1-1,n_2+2|\;.
\end{aligned}
$$

Case 1': Terms of the form $|n_1,n_2\rangle\langle n_1+1,n_2-2|$, involving $F4$, $\overline{F}_5$ and F_{10},

$$
\begin{aligned}
C_{1'}(n) := F_4(n)\,&|n_1,n_2+2\rangle\langle n_1+1,n_2| \\
&+ \overline{F}_5(n)\,|n\rangle\langle n_1+1,n_2-2| + F_{10}(n)\,|n_1-1,n_2+2\rangle\langle n|\;.
\end{aligned}
$$

Case 2: Terms of the form $|n_1,n_2-2\rangle\langle n_1,n_2|$, involving F_6 and $\overline{F}_8$,

$$C_2(n) := F_6(n)\,|n_1,n_2-2\rangle\langle n| + \overline{F}_8(n)\,|n\rangle\langle n_1,n_2+2|\;.$$

Case 2': Terms of the form $|n_1,n_2\rangle\langle n_1,n_2-2|$, involving $\overline{F}_6$ and F_8,

$$C_{2'}(n) := \overline{F}_6(n)\,|n\rangle\langle n_1,n_2-2| + F_8(n)\,|n_1,n_2+2\rangle\langle n|\;.$$

Case 3: Terms of the form $|n_1-1,n_2\rangle\langle n_1,n_2|$, involving F_7 and $\overline{F}_9$,

$$C_3(n) := F_7(n)\,|n_1-1,n_2\rangle\langle n| + \overline{F}_9(n)\,|n\rangle\langle n_1+1,n_2|\;.$$

Case 3': Terms of the form $|n_1,n_2\rangle\langle n_1-1,n_2|$, involving $\overline{F}_7$ and F_9,

$$C_{3'}(n) := \overline{F}_7(n)\,|n\rangle\langle n_1-1,n_2| + F_9(n)\,|n_1+1,n_2\rangle\langle n|\;.$$

To summarize, we have decomposed $\mathcal{L}[b^2+za+w]^\dagger(|n\rangle\langle n|)$ into the sum

$$\mathcal{L}[b^2+za+w]^\dagger(|n\rangle\langle n|) = C_0(n)+C_1(n)+C_{1'}(n)+C_2(n)+C_{2'}(n)+C_3(n)+C_{3'}(n)\,. \tag{70}$$

Next, we introduce the functions $g_{j,l}:\mathbb{N}\to\mathbb{R}$, $j\in\{1,2\}$, $l\in\mathbb{N}$, as

$$g_{j,l}(x) = \begin{cases} f_j(x)-f_j(x-l) & x\geq l;\\ f_j(x) & l>x\geq 0;\\ 0 & 0>x\,. \end{cases}$$

Multiplying Equation (70) by $f(n)$ and summing over $n\in\mathbb{N}^2$, we find that

$$\mathcal{L}[b^2+za+w]^\dagger(f(N)) = \hat{C}_0+\hat{C}_1+\hat{C}_{1'}+\hat{C}_2+\hat{C}_{2'}+\hat{C}_3+\hat{C}_{3'}\,,$$

with

$$\begin{aligned}
\hat{C}_0 &\coloneqq \sum_n C_0(n) = \sum_n -\Big(f_1(n_1)g_{2,2}(n_2)n_2(n_2-1)+f_2(n_2)g_{1,1}(n_1)|z|^2 n_1\Big)\,|n\rangle\langle n|\\
\hat{C}_1 &\coloneqq \sum_n C_1(n)\\
&= \sum_n -\frac{\overline{z}}{2}\sqrt{(n_1+1)(n_2-1)n_2}\Big(f_1(n_1)g_{2,2}(n_2)+g_{1,1}(n_1+1)f_2(n_2-2)\Big)\,|n_1+1,n_2-2\rangle\langle n|\\
\hat{C}_{1'} &\coloneqq \hat{C}_1^\dagger\\
&= \sum_n -\frac{z}{2}\sqrt{(n_1+1)(n_2-1)n_2}\Big(f_1(n_1)g_{2,2}(n_2)+g_{1,1}(n_1+1)f_2(n_2-2)\Big)\,|n\rangle\langle n_1+1,n_2-2|\\
\hat{C}_2 &\coloneqq \sum_n C_2(n) = \sum_n -\frac{\overline{w}}{2}\sqrt{(n_2-1)n_2}f_1(n_1)g_{2,2}(n_2)\,|n_1,n_2-2\rangle\langle n|\\
\hat{C}_{2'} &\coloneqq \hat{C}_2^\dagger = \sum_n -\frac{w}{2}\sqrt{(n_2-1)n_2}f_1(n_1)g_{2,2}(n_2)\,|n\rangle\langle n_1,n_2-2|\\
\hat{C}_3 &\coloneqq \sum_n C_3(n) = \sum_n -\frac{\overline{w}z}{2}\sqrt{n_1}g_{1,1}(n_1)f_2(n_2)\,|n_1-1,n_2\rangle\langle n|\\
\hat{C}_{3'} &\coloneqq \hat{C}_3^\dagger = \sum_n -\frac{w\overline{z}}{2}\sqrt{n_1}g_{1,1}(n_1)f_2(n_2)\,|n\rangle\langle n_1-1,n_2|\ .
\end{aligned}$$

We will use an upper bound on $\hat{C}_0(n)$ in what follows:

$$\begin{aligned}
C_0(n) &= -\Big(f_1(n_1)g_{2,2}(n_2)(n_2-1)n_2+|z|^2 g_{1,1}(n_1)f_2(n_2)n_1\Big)\\
&\leq -\Big(f_1(n_1)g_{2,2}(n_2)1_{n_2\geq 2}((n_2+1)^2-3(n_2+1))+|z|^2 g_{1,1}(n_1)f_2(n_2)n_1\Big)\\
&\overset{(1)}{\leq} -\Big\{k_2 f_1(n_1)(n_2+1)^{k_2/2}1_{n_2\geq 2}\Big((n_2+1)-1_{k_2\geq 3}\frac{k_2}{2}-6\Big)\\
&\qquad +|z|^2 1_{n_1\geq 1}(n_1+1)^{k_1/2-1}f_2(n_2)n_1\Big\}\\
&\equiv C_{0'}(n)\,,
\end{aligned}$$

where (1) follows from Lemma C.2. We denote this upper bound by $\hat{C}_{0'}=\sum_n C_{0'}(n)\,|n\rangle\langle n|$. Recall that, by Lemma 4.3 and Remark 7,

$$\begin{aligned}
\mathcal{L}[a^2-\alpha^2]^\dagger(f(N)) &\leq -(N_1+\mathbb{1})^{k_1/2+1}(N_2+\mathbb{1})^{k_2/2}+\mu_{k_1}^{(2)}(N_2+\mathbb{1})^{k_2/2}\\
&= \sum_n\Big(-(n_1+1)f(n)+\mu_{k_1}^{(2)}\,f_2(n_2)\Big)\,|n\rangle\langle n| =: \hat{C}_4\,,
\end{aligned} \tag{71}$$

where $\mu^{(2)}_{k_1} = \Delta_2^{\nu}\left(\frac{(\nu-1)^{\nu-1}}{\nu^{\nu}}\right)$ with $\nu = \frac{k_1}{2}+1$ and $\Delta_2 = 6+2|\alpha|^2 k_1 2^{k_1/2-1}\sqrt{2}$. Therefore, the diagonal contribution of $\mathcal{L}^\dagger(f(N))$ can be controlled by

$$\begin{aligned}(C_{0'}+C_4)(n) &:= -\Big(k_2 f_1(n_1)(n_2+1)^{k_2/2+1}1_{n_2\geq 2} + |z|^2 1_{n_1\geq 1} f(n) + (n_1+1)^{k_1/2+1} f_2(n_2)\Big)\\ &\quad + k_2 f_1(n_1)(n_2+1)^{k_2/2} 1_{n_2\geq 2}\Big(1_{k_2\geq 3}\frac{k_2}{2}+6\Big)\\ &\quad + |z|^2 1_{n_1\geq 1}(n_1+1)^{k_2/2-1} f_2(n_2)\\ &\quad + \mu^{(2)}_{k_1}(n_2+1)^{k_2/2}\\ &\leq -\frac{1}{2}\Big(k_2 f_1(n_1)(n_2+1)^{k_2/2+1} + (n_1+1)^{k_1/2+1} f_2(n_2)\Big) + \Delta_0\\ &=: -x(n)\end{aligned}$$

where the constant Δ_0 is achieved by splitting off half of the negative leading order terms to control the lower order positive contributions (see for example the proof of Lemma 4.3). Next, we consider operators of the form

$$-x_1\,|e_1\rangle\langle e_1| - x_2\,|e_2\rangle\langle e_2| + y\,|e_2\rangle\langle e_1| + \overline{y}\,|e_1\rangle\langle e_2|\,, \tag{72}$$

where $\{e_1, e_2\}$ forms an orthonormal basis of a two-dimensional Hilbert space and $x_1, x_2 \in \mathbb{R}, y \in \mathbb{C}$. The operator in Equation (72) has the eigenvalues

$$\lambda_+ = \frac{-x_1-x_2+\sqrt{(x_1-x_2)^2+4|y|^2}}{2}, \quad \lambda_- = \frac{-x_1-x_2-\sqrt{(x_1-x_2)^2+4|y|^2}}{2} \tag{73}$$

Moreover,

$$\frac{-x_1-x_2+\sqrt{(x_1-x_2)^2+4|y|^2}}{2} \leq -\min\{x_1,x_2\} + |y|\,.$$

Using this bound, we control each of the off-diagonal operators $\hat{C}_i + \hat{C}_{i'}$, $i \in \{1,2,3\}$, in terms of $\frac{1}{4}X$, where $X = \sum_n x(n)\,|n\rangle\langle n|$.

$$\begin{aligned}-\frac{1}{4}X + \hat{C}_1 + \hat{C}_{1'} &:= \sum_n -\frac{1}{4}x(n)\,|n\rangle\langle n| + y_1\,|n\rangle\langle n_1+1, n_2-2| + \overline{y}_1\,|n_1+1,n_2-2\rangle\langle n|\\ &\leq \sum_{n|n_2\geq 2} -\frac{1}{8}x(n)\,|n\rangle\langle n| - \frac{1}{8}x(n_1+1,n_2-2)\,|n_1+1,n_2-2\rangle\langle n_1+1,n_2-2|\\ &\qquad + y_1\,|n\rangle\langle n_1+1,n_2-2| + \overline{y}_1\,|n_1+1,n_2-2\rangle\langle n|\end{aligned}$$

with

$$y_1 = y_1(n_1,n_2) = -\frac{z}{2}\sqrt{(n_1+1)(n_2-1)n_2}\Big(f_1(n_1)g_{2,2}(n_2) + g_{1,1}(n_1+1)f_2(n_2-2)\Big)$$

so that

$$-\frac{1}{4}X+\hat{C}_1+\hat{C}_{1'} \leq \sum_{n|n_2\geq 2}(-\min\{x_1,x_2\}+|y_1|)\,(|n\rangle\langle n|+|n_1+1,n_2-2\rangle\langle n_1+1,n_2-2|) \tag{74}$$

where $x_1 = \frac{1}{8}x(n_1,n_2)$ and $x_2 = \frac{1}{8}x(n_1+1,n_2-2)$. Moreover, for $n_2 \geq 2$

$$\begin{aligned}|y_1| &\overset{(1)}{\leq} \frac{|z|}{2}\sqrt{(n_1+1)(n_2-1)n_2}\Big(f_1(n_1)2k_2(n_2+1)^{k_2/2-1} + k_1(n_1+2)^{k_1/2-1}f_2(n_2-2)\Big)\\ &\leq |z|k_2(n_1+1)^{k_1/2+1/2}(n_2+1)^{k_2/2}\\ &\qquad + \frac{|z|k_1}{2}\sqrt{(n_2-1)^2+n_2-1}\,(n_1+2)^{k_1/2-1/2}f_2(n_2-2)\\ &\leq |z|k_2(n_1+1)^{k_1/2+1/2}(n_2+1)^{k_2/2} + |z|k_1\,(n_1+2)^{k_1/2-1/2}(n_2-1)^{k_2/2+1}\,,\end{aligned}$$

where (1) follows from Lemma C.2. At this stage, we consider two cases:

Case (i): $x_2 \geq x_1$. In that case, $-\min\{x_1, x_2\} + |y_1| = -x_1 + |y_1|$, and therefore

$$-\min\{x_1, x_2\} + |y_1| \leq -\frac{1}{16}\Big(k_2 f_1(n_1)(n_2+1)^{k_2/2+1} + (n_1+1)^{k_1/2+1} f_2(n_2)\Big) + \frac{1}{8}\Delta_0 + \underbrace{|z| k_2 (n_1+1)^{k_1/2+1/2}(n_2+1)^{k_2/2}}_{=:A_1} + \underbrace{|z| k_1 (n_1+2)^{k_1/2-1/2}(n_2-1)^{k_2/2+1}}_{=:A_2}.$$

Note that the first positive non-constant term A_1 can be controlled with half the negative contribution in the first term by a constant using the same type of polynomial optimization as in the proof of Lemma 4.3. For the last term, i.e. A_2, we use the assumption

$$|z| k_1 2^{k_1/2-1/2} \leq |\alpha| k_1 2^{k_1/2-1/2} \leq \frac{1}{32} k_2, \tag{75}$$

which allows us to control A_2 with the other half of the first term, as we already did with A_1. Recall the definition $z = -\frac{\alpha}{2}(1 - e^{2i\pi t/T})$. Summarising the above considerations we can conclude the existence of a constant $\tilde{\Delta}_1'$ such that

$$-\min\{x_1, x_2\} + |y_1| \leq \tilde{\Delta}_1'.$$

Case (ii): $x_2 \leq x_1$. In that case $-\min\{x_1, x_2\} + |y_1| = -x_2 + |y_1|$, and therefore

$$-\min\{x_1, x_2\} + |y_1| = -\frac{1}{16}\Big(k_2 f_1(n_1+1)(n_2-1)^{k_2/2+1} + (n_1+2)^{k_1/2+1} f_2(n_2-2)\Big) + \frac{1}{8}\Delta_0 + |z| k_2 (n_1+1)^{k_1/2+1/2}(n_2+1)^{k_2/2} + |z| k_1 (n_1+2)^{k_1/2-1/2}(n_2-1)^{k_2/2+1}.$$

To upper bound the above, we use again the assumption (75), which implies the existence of a constant $\tilde{\Delta}_1'$ such that

$$-\min\{x_1, x_2\} + |y_1| \leq \tilde{\Delta}_1'.$$

Combining cases (i) and (ii) above, denoting $\Delta_1 := \max\{\tilde{\Delta}_1, \tilde{\Delta}_1'\}$ and plugging the bounds into (74), we arrive at

$$-\frac{1}{4}X + \hat{C}_1 + \hat{C}_{1'} \leq \Delta_1 \sum_{n | n_2 \geq 2} \left(|n\rangle\langle n| + |n_1+1, n_2-2\rangle\langle n_1+1, n_2-2| \right) \tag{76}$$

Next, we control $-\frac{1}{4}X + \hat{C}_2 + \hat{C}_{2'}$. Here, we have

$$y_2 = y_2(n_1, n_2) = -\frac{\overline{w}}{2}\sqrt{(n_2-1)n_2}\, f_1(n_1) g_{2,2}(n_2),$$

$x_1 = \frac{1}{8}x(n)$, and $x_2 = \frac{1}{8}x(n_1, n_2-2)$. By Lemma C.2, we have that

$$|y_2| \leq |w| \sqrt{(n_2-1)n_2}\, f_1(n_1) k_2 (n_2+1)^{k_2/2-1}.$$

Therefore, the negative contribution from $\min\{x_1, x_2\}$ has leading order in both variables n_1 and n_2, which implies the existence of a constant Δ_2 such that

$$-\min\{x_1, x_2\} + |y_2| \leq \Delta_2.$$

Hence,

$$-\frac{1}{4}X+\hat{C}_2+\hat{C}_{2'}\leq\Delta_2\sum_{n|n_2\geq 2}(|n\rangle\langle n|+|n_1,n_2-2\rangle\langle n_1,n_2-2|) \tag{77}$$

Finally, we consider $-\frac{1}{4}X+\hat{C}_3+\hat{C}_{3'}$. In this case,

$$y_3=y_3(n_1,n_2)=-\frac{\overline{w}z}{2}\sqrt{n_1}g_{1,1}(n_1)f_2(n_2)\,,$$

$x_1=\frac{1}{8}x(n)$, and $x_2=\frac{1}{8}x(n_1-1,n_2)$. Similarly to the above, we can argue the existence of a constant Δ_3 such that

$$-\max\{x_1,x_2\}+|y_3|\leq\Delta_3\,.$$

Hence,

$$-\frac{1}{4}X+\hat{C}_3+\hat{C}_{3'}\leq\Delta_3\sum_{n|n_1\geq 1}(|n\rangle\langle n|+|n_1-1,n_2\rangle\langle n_1-1,n_2| \tag{78}$$

Combining (76), (77) and (78), we have shown that

$$\begin{aligned}\mathcal{L}^\dagger(f(N))&\leq-\frac{X}{4}+\Delta_1\sum_{n|n_2\geq 2}(|n\rangle\langle n|+|n_1+1,n_2-2\rangle\langle n_1+1,n_2-2|)\\&+\Delta_2\sum_{n|n_2\geq 2}(|n\rangle\langle n|+|n_1,n_2-2\rangle\langle n_1,n_2-2|)\\&+\Delta_3\sum_{n|n_1\geq 1}(|n\rangle\langle n|+|n_1-1,n_2\rangle\langle n_1-1,n_2|)\\&\leq-\frac{X}{4}+2(\Delta_1+\Delta_2+\Delta_3)\mathbb{1}\\&=-\frac{1}{8}\Big(k_2f_1(N_1)(N_2+1)^{k_2/2+1}+(N_1+1)^{k_1/2+1}f_2(N_2)\Big)+\mu_{\mathbf{k}}\,\mathbb{1}\\&\leq-\frac{1+k_2}{8}f(N)+\mu_{\mathbf{k}}\mathbb{1}\,,\end{aligned}$$

with

$$\mu_{\mathbf{k}}:=\frac{\Delta_0}{4}+2(\Delta_1+\Delta_2+\Delta_3)\,.$$

The claim (69) finally follows from Proposition 4.1. □

5 Perturbation bounds

In this section, we establish a perturbative analysis at any time scale for the semigroups considered in Section 4. In finite dimensions, [65, Theorem 6] gives a quantitative bound which controls the perturbation of a quantum dynamical semigroup under the condition that the latter converges exponentially fast to a unique invariant state τ: for two generators $\mathcal{L}$ and $\mathcal{L}+\mathcal{K}$, if $\mathcal{L}$ satisfies $\|e^{t\mathcal{L}}-\mathrm{tr}(.)\,\tau\|_{1\to1}\leq ce^{-\omega t}$ for all $t\geq 0$ and some $c,\omega>0$, then

$$\forall\rho,\sigma\text{ states},\quad\left\|e^{t\mathcal{L}}(\rho)-e^{t(\mathcal{L}+\mathcal{K})}(\sigma)\right\|_1\leq\begin{cases}\|\rho-\sigma\|_1+t\|\mathcal{K}\|_{1\to1}\,, & t<\hat{t}\\ ce^{-\omega t}\|\rho-\sigma\|_1+\frac{\log(c)+1-ce^{-\omega t}}{\omega}\|\mathcal{K}\|_{1\to1}\,, & t\geq\hat{t}\end{cases}$$

where $\hat{t}:=\frac{\log(c)}{\omega}$. The result can be easily extended to the case of bounded generators in infinite dimensions, although proving the exponential decay for the semigroup generated by $\mathcal{L}$ is not easy. The situation becomes even trickier in the case of unbounded generators since the use of a Duhamel integral as in the proof in finite dimensions requires a proper justification. It is precisely these issues that we are interested in and want to address here.

5.1 Gaussian perturbations of the quantum Ornstein Uhlenbeck semigroup

The quantum Ornstein Uhlenbeck semigroup is well-known to correspond to a so-called beam-splitter channel of exponentially decreasing transmissivity $e^{-(\lambda^2-\mu^2)t}$ with unique Gaussian invariant state (see [21]):

$$\sigma := \frac{\lambda^2-\mu^2}{\mu^2}\sum_{k=0}^{\infty}\left(\frac{\mu^2}{\lambda^2}\right)^k |k\rangle\langle k| \ .$$

While quantitative statements about the convergence of this semigroup towards σ are known [18, 13, 14, 21], they do not necessarily imply convergence in trace distance in contrast to their finite-dimensional analogues. In contrast, the semigroup is known to contract a certain kind of quantum Wasserstein distance, which we introduce now. First, given a bounded, self-adjoint operator $X \in \mathcal{B}(\mathcal{H})$, we call X a Lipschitz observable if aX, $a^\dagger X$ are bounded, and if Xa and $Xa^\dagger$ are closable operators with bounded closures $\overline{Xa}$ and $\overline{Xa^\dagger}$. In this case, we denote by $\partial_a(X) := aX - \overline{Xa}$ and $\partial_{a^\dagger}(X) = a^\dagger X - \overline{Xa^\dagger}$. The Lipschitz constant of X is then defined as

$$\|X\|_{\mathrm{Lip}} := \max\left\{\|\partial_a(X)\|_\infty,\, \|\partial_{a^\dagger}(X)\|_\infty\right\}.$$

We denote the set of Lipschitz observables by Lip. Next, any $T \in \mathcal{T}_{1,\mathrm{sa}}$, we denote

$$\|T\|_{W_1} := \sup\left\{\mathrm{tr}\left[X\,T\right] :\, X \in \mathrm{Lip},\, \|X\|_{\mathrm{Lip}} \leq 1\right\} \ .$$

In [28, Proposition 6.4], the authors showed that, for any $T \in \mathcal{T}_{1,\mathrm{sa}}$ and $t > 0$,

$$\|e^{t\mathcal{L}_{\mathrm{qOU}}}(T)\|_1 \leq \sqrt{\frac{e^{-(\lambda^2-\mu^2)t}}{1-e^{-(\lambda^2-\mu^2)t}}}\left(\|a\sigma-\sigma a\|_1 + \|a^\dagger\sigma - \sigma a^\dagger\|_1\right)\|T\|_{W_1}\,. \tag{79}$$

Moreover, using the canonical commutation relations, one can also prove the following identities (see e.g. [14], or [28, Proposition 6.2]): for any two states $\rho_1, \rho_2 \in \mathcal{T}_{1,\mathrm{sa}}$,

$$\|e^{t\mathcal{L}_{\mathrm{qOU}}}(\rho_1-\rho_2)\|_{W_1} \leq e^{-\frac{(\lambda^2-\mu^2)t}{2}}\,\|\rho_1-\rho_2\|_{W_1}\,. \tag{80}$$

In the next proposition, we use these conditions to find a perturbation bound for any Gaussian perturbation of the quantum Ornstein Uhlenbeck semigroup.

Proposition 5.1 *Let $(\mathcal{L}_{\mathrm{qOU}}, \mathcal{T}_f)$ be the generator of the quantum Ornstein Uhlenbeck semigroup with $\lambda > \mu \geq 0$ and $(\varepsilon\mathcal{L}_G, \mathcal{T}_f) := (\varepsilon\mathcal{L}[\gamma a + \eta a^\dagger], \mathcal{T}_f)$ a Gaussian perturbation with $\gamma, \eta \in \mathbb{R}$, $\varepsilon > 0$. Then, assuming $\lambda^2 - \mu^2 + |\gamma|^2 - |\eta|^2 > 0$, $\mathcal{L}_{\mathrm{qOU}} + \varepsilon\mathcal{L}_G$ generates a positivity and Sobolev preserving semigroup on $W^{k,1}$ for $k \geq 1$, and there exist uniformly bounded functions $C(\varepsilon), D(\varepsilon)$ depending on $\lambda, \mu, |\eta|, |\gamma|$ such that, for all $t \geq 0$ and states $\rho \in W^{2,1}$*

$$\left\|\left(e^{t\mathcal{L}_{\mathrm{qOU}}} - e^{t(\mathcal{L}_{\mathrm{qOU}}+\varepsilon\mathcal{L}_G)}\right)(\rho)\right\|_1 \leq \varepsilon\, C(\varepsilon)\,\max\left\{\|\rho\|_{W^{2,1}}, D(\varepsilon)\right\}. \tag{81}$$

Proof. The generation of a Sobolev preserving semigroup was already stated in Lemma 4.2 for $\mathcal{L}_{\mathrm{qOU}}$ and its proof can easily be extended to $\mathcal{L}_{\mathrm{qOU}} + \varepsilon\mathcal{L}_G$. For instance, given a state $\rho \in \mathcal{T}_f$, one can show that

$$\mathrm{tr}[\mathcal{L}_G(\rho)(N+\mathbb{1})] \leq -(|\gamma|^2 - |\eta|^2)\ \mathrm{tr}[\rho N] + |\eta|^2\,,$$

We have also seen in the proof of Lemma 4.2 that $\mathrm{tr}[\rho\mathcal{L}_{\mathrm{qOU}}] \leq -(\lambda^2-\mu^2)\,\mathrm{tr}[\rho N] + \mu^2$, so that

$$\mathrm{tr}[(\mathcal{L}_{\mathrm{qOU}} + \varepsilon\mathcal{L}_G)(\rho)(N+\mathbb{1})] \leq -(\lambda^2 - \mu^2 + \varepsilon|\gamma|^2 - \varepsilon|\eta|^2)\,\mathrm{tr}[\rho(N+\mathbb{1})] + \lambda^2 + \varepsilon|\gamma|^2\,.$$

Therefore, by Proposition 4.1 we have that, as long as $\lambda^2 - \mu^2 + \varepsilon|\gamma|^2 - \varepsilon|\eta|^2 > 0$, for all states $\rho \in W^{2,1}$, $\rho \geq 0$ and all $t \geq 0$,

$$\|e^{t(\mathcal{L}_{\mathrm{qOU}}+\varepsilon\mathcal{L}_G)}(\rho)\|_{W^{2,1}} \leq \max\left\{\|\rho\|_{W^{2,1}}, \frac{\lambda^2+\varepsilon|\gamma|^2}{\lambda^2-\mu^2+\varepsilon|\gamma|^2-\varepsilon|\eta|^2}\right\}.$$

Next, for $\rho \in \mathcal{T}_f$ and $0 < u < t$, and denoting $\mathcal{L} \equiv \mathcal{L}_{\mathrm{qOU}}$ and $\widetilde{\mathcal{L}} \equiv \mathcal{L}_{\mathrm{qOU}} + \varepsilon\mathcal{L}_G$,

$$\left\|\left(e^{t\mathcal{L}} - e^{t\widetilde{\mathcal{L}}}\right)(\rho)\right\|_1 \leq \left\|e^{u\mathcal{L}}\left(e^{(t-u)\mathcal{L}} - e^{(t-u)\widetilde{\mathcal{L}}}\right)(\rho)\right\|_1 + \left\|\left(e^{u\mathcal{L}} - e^{u\widetilde{\mathcal{L}}}\right)e^{(t-u)\widetilde{\mathcal{L}}}(\rho)\right\|_1 \equiv A + B\,.$$

We use Equation (79), so that

$$\begin{aligned} A &\leq c_u\left\|\left(e^{(t-u)\mathcal{L}} - e^{(t-u)\widetilde{\mathcal{L}}}\right)(\rho)\right\|_{W_1} \\ &\leq c_u\,\varepsilon\int_0^{t-u}\left\|e^{s\mathcal{L}}\mathcal{L}_G\,e^{(t-u-s)\widetilde{\mathcal{L}}}(\rho)\right\|_{W_1} ds \\ &= c_u\,\varepsilon\int_0^{t-u} e^{-\frac{(\lambda^2-\mu^2)s}{2}}\left\|\mathcal{L}_G\,e^{(t-u-s)\widetilde{\mathcal{L}}}(\rho)\right\|_{W_1} ds\,, \end{aligned}$$

where $c_u := \sqrt{\frac{e^{-(\lambda^2-\mu^2)u}}{1-e^{-(\lambda^2-\mu^2)u}}}\left(\|\partial_a(\sigma)\|_1 + \|\partial_{a^\dagger}(\sigma)\|_1\right)$. Moreover, denoting $\widetilde{\rho}_v := e^{v\widetilde{\mathcal{L}}}(\rho)$ and $b = \gamma a + \eta a^\dagger$, since $\widetilde{\rho}_v \in W^{2,1}$ for all $v \geq 0$,

$$\begin{aligned} \left\|\mathcal{L}_G\,\widetilde{\rho}_v\right\|_{W_1} &= \sup_{\|X\|_{\mathrm{Lip}}\leq 1} \operatorname{tr}[X\mathcal{L}_G\widetilde{\rho}_v] \\ &= \frac{1}{2}\sup_{\|X\|_{\mathrm{Lip}}\leq 1} \operatorname{tr}\left[\partial_{b^\dagger}(X)b\widetilde{\rho}_v - \partial_b(X)\,\widetilde{\rho}_v b^\dagger\right] \\ &\leq (|\eta|+|\gamma|)\left(\|b\widetilde{\rho}_v\|_1 + \|\widetilde{\rho}_v b^\dagger\|_1\right) \\ &\leq (|\eta|+|\gamma|)\left(\|b\,(N+\mathbb{1})^{-\frac{1}{2}}\| + \|(N+\mathbb{1})^{-\frac{1}{2}}\,b^\dagger\|\right)\|\widetilde{\rho}_v\|_{W^{2,1}} \\ &\leq (|\eta|+|\gamma|)\left(\|b\,(N+\mathbb{1})^{-\frac{1}{2}}\| + \|(N+\mathbb{1})^{-\frac{1}{2}}\,b^\dagger\|\right) \\ &\qquad\cdot\max\left\{\|\rho\|_{W^{2,1}}, \frac{\lambda^2+\varepsilon|\gamma|^2}{\lambda^2-\mu^2+\varepsilon|\gamma|^2-\varepsilon|\eta|^2}\right\}. \end{aligned}$$

On the other hand,

$$\begin{aligned} B &\leq \varepsilon\int_0^u \|e^{s\mathcal{L}}\mathcal{L}_G e^{(t-s)\widetilde{\mathcal{L}}}(\rho)\|_1\,ds \\ &\leq u\varepsilon\max_{s\in[0,u]}\|\mathcal{L}_G e^{(t-s)\widetilde{\mathcal{L}}}(\rho)\|_1 \\ &\leq u\varepsilon\|\mathcal{L}_G\circ\mathcal{W}^{-2}\|_{\mathcal{T}_1\to\mathcal{T}_1}\max_{s\in[0,u]}\|e^{(t-s)\widetilde{\mathcal{L}}}(\rho)\|_{W^{2,1}} \\ &\leq u\varepsilon\|\mathcal{L}_G\circ\mathcal{W}^{-2}\|_{\mathcal{T}_1\to\mathcal{T}_1}\max\left\{\|\rho\|_{W^{2,1}}, \frac{\lambda^2+\varepsilon|\gamma|^2}{\lambda^2-\mu^2+\varepsilon|\gamma|^2-\varepsilon|\eta|^2}\right\}. \end{aligned}$$

The result follows from a simple bound on $\|\mathcal{L}_G\circ\mathcal{W}^{-2}\|_{\mathcal{T}_1\to\mathcal{T}_1}$ with the help of Hölder's inequality and a standard density argument, and by choosing u appropriately. □

5.2 Photon-dissipation and CAT qubits

As mentioned before, one crucial property of the underlying evolution in continuous error correction is that it is exponentially converging to the code-space. In the spirit of [65], we prove a large time perturbation result for the l-photon dissipation perturbed by a Hamiltonian evolution. This can be generalized to dissipative perturbations. First, we recall the exponential convergence of the l-photon dissipation defined by $L = a^l - \alpha^l$ ([5, Theorem 2]):

$$\operatorname{tr}\Big[Le^{t\mathcal{L}_l}(\rho)L^\dagger\Big] \leq e^{-l!t}\operatorname{tr}\Big[L\rho L^\dagger\Big]\,. \tag{82}$$

Additionally, it is shown that there is a unique limit $\overline{\rho}$ of $e^{t\mathcal{L}_l}(\rho)$ for $t \to \infty$. We show large-time perturbation bounds by combining this bound with our established generation theory for Sobolev and positivity-preserving Markov semigroups. We start with the l-photon dissipation perturbed by the Hamiltonian introduced in Lemma 4.3, i.e. $H = p_H(a, a^\dagger)$ with $d_H \leq 2(l-1)$.

Theorem 5.2 *Let $\mathcal{L}_l$ be the generator of the l-photon dissipation and $p_H \in \mathbb{C}[X, Y]$ with $\deg(p_H) = d_H \leq 2(l-1)$ such that $H = p_H(a, a^\dagger)$ is a symmetric operator. Then, there exist explicit constants $c, \gamma > 0$ such that for $\varepsilon \geq 0$ and all states $\rho \in W^{2(l+d_H+2),1}$*

$$\Big|\operatorname{tr}\Big[L\left(e^{t\mathcal{L}_l}(\rho) - e^{t(\mathcal{L}_l+\varepsilon\mathcal{H}[H])}(\rho)\right)L^\dagger\Big]\Big| \leq \varepsilon c\left(1 - e^{-l!t}\right)\max\{\gamma, \|\rho\|_{W^{2(l+d_H+2),1}}\}\,.$$

Proof. The proof consists in applying Lemma 4.6 in combination with Equation (82). Let $\rho \in \mathcal{T}_f$ and $\mathcal{W}^k = (N+\mathbb{1})^{k/4}\cdot(N+\mathbb{1})^{k/4}$, then

$$\begin{aligned}
&\operatorname{tr}\Big[L\Big(e^{t\mathcal{L}_l}(\rho) - e^{t(\mathcal{L}_l+\varepsilon\mathcal{H}[H])}(\rho)\Big)L^\dagger\Big]\\
&\quad= \varepsilon\int_0^t \operatorname{tr}\Big[Le^{s\mathcal{L}_l}\mathcal{W}^{-2(l+2)}\mathcal{W}^{2(l+2)}\mathcal{H}[H]e^{(t-s)(\mathcal{L}_l+\varepsilon\mathcal{H}[H])}(\rho)L^\dagger\Big]ds\\
&\quad\overset{(1)}{\leq} \varepsilon\int_0^t \operatorname{tr}\Big[Le^{s\mathcal{L}_l}\mathcal{W}^{-2(l+2)}(\mathbb{1})L^\dagger\Big]ds\, 2\Lambda d_H^{l+2}\sqrt{d_H!}\max\left\{\gamma_\varepsilon,\ \|\rho\|_{W^{2(l+d_H+2),1}}\right\}\\
&\quad\overset{(2)}{\leq} \varepsilon\frac{\pi^2}{3}\Lambda d_H^2\sqrt{d_H!}(1+|\alpha|^l\,(l+1)\sqrt{l!}+|\alpha|^{2l})\frac{1}{l!}(1-e^{-l!t})\max\left\{\gamma_\varepsilon,\ \|\rho\|_{W^{2(l+d_H+2),1}}\right\}\\
&\quad=: \varepsilon c(1-e^{-l!t})\max\left\{\gamma_\varepsilon,\ \|\rho\|_{W^{2(l+d_H+2),1}}\right\}
\end{aligned}$$

where Λ denotes the larges coefficient of $\mathcal{H}[H]$ in absolute value. Note that the Bochner integral in the calculation is well-defined by the boundedness of the integrand w.r.t. s and the same argumentation as Theorem 2.11. Besides the boundedness above, the Sobolev preserving property of the involved semigroups implies by construction that the integral is Sobolev preserving. Therefore, the integral commutes with the map $x \mapsto \operatorname{tr}\Big[LxL^\dagger\Big]$ for $x \in W^{2(l+d_H+2),1}$. In (1) we used that $Le^{s\mathcal{L}_l}(\cdot)L^\dagger$ preserves positivity and

$$\begin{aligned}
\mathcal{W}^{2(l+1)}\mathcal{H}[H]e^{t-s(\mathcal{L}_l+\varepsilon\mathcal{H}[H]}(\rho) &\leq \Big\|\mathcal{W}^{2(l+2)}\mathcal{H}[H]e^{(t-s)(\mathcal{L}_l+\varepsilon\mathcal{H}[H]}(\rho)\Big\|_\infty\mathbb{1}\\
&\leq \Big\|\mathcal{W}^{2(l+2)}\mathcal{H}[H]e^{(t-s)(\mathcal{L}_l+\varepsilon\mathcal{H}[H]}(\rho)\Big\|_1\mathbb{1}\\
&\leq 2\Lambda d_H^{l+2}\sqrt{d_H!}\Big\|e^{(t-s)(\mathcal{L}_l+\varepsilon\mathcal{H}[H]}(\rho)\Big\|_{W^{2(l+d_H+2)}}\mathbb{1}\\
&\leq 2\Lambda d_H^{l+2}\sqrt{d_H!}\max\{\gamma_\varepsilon, \|\rho\|_{W^{2(l+d_H+2)}}\}\mathbb{1}\,.
\end{aligned}$$

In the above estimation, we used $(N+\mathbb{1})^{-d_H}$ to control $\mathcal{H}[H]$. We further used the Sobolev preserving property of the semigroup from Lemma 4.6. Finally, applying the decay (q.v. Equation (82)) to $\mathcal{W}^{-2(l+1)}(\mathbb{1}) = (N+\mathbb{1})^{-(l+2)}$ we estimated in (2):

$$\int_0^t \operatorname{tr}\Big[Le^{s\mathcal{L}_l}\mathcal{W}^{-2(l+2)}(\mathbb{1})L^\dagger\Big]ds \le \int_0^t e^{-l!t}\operatorname{tr}\Big[L(N+\mathbb{1})^{-(l+2)}L^\dagger\Big] \le (1-e^{-l!t})(1+|\alpha|^l\,(l+1)\sqrt{l!}+|\alpha|^{2l})\frac{\pi^2}{6}$$

with the bound

$$\|L(N+1)^{-l-2}L^\dagger\|_1 \le (1+|\alpha|^l\,(l+1)\sqrt{l!}+|\alpha|^{2l})\frac{\pi^2}{6}$$

that follows from

$$\begin{aligned} L^\dagger L &= (a^\dagger)^l a^l - \overline{\alpha}^l a^l - \alpha^l (a^\dagger)^l + |\alpha|^{2l} \\ &\le (N-(l-1)\mathbb{1})\cdots(N-\mathbb{1})N + |\alpha|^l\,(l+1)\sqrt{(N+l\mathbb{1})\cdots(N+\mathbb{1})} + |\alpha|^{2l} \\ &\le (N+\mathbb{1})^l + |\alpha|^l\,(l+1)\sqrt{l!}\,(N+\mathbb{1})^{l/2} + |\alpha|^{2l}\,. \end{aligned} \tag{83}$$

and the $\left\|(N+\mathbb{1})^{-2}\right\|_1 = \frac{\pi^2}{6}$. This concludes that claim. □

Special cases of the above result include the X or $Z(\theta)$ gate.

Corollary 5.3 ($Z(\theta)$-gate) *Let $\mathcal{L}_2$ be the 2-photon dissipation. Then, for all $\varepsilon \in [0,1]$ and all states $\rho \in W^{10,1}$*

$$\Big|\operatorname{tr}\Big[L\left(e^{t\mathcal{L}_2}(\rho) - e^{t(\mathcal{L}_2+\varepsilon\mathcal{H}[a+a^\dagger])}(\rho)\right)L^\dagger\Big]\Big| \le \varepsilon 2(1+6|\alpha|^2+|\alpha|^4)(1-e^{-2t})\max\{\gamma_\varepsilon, \|\rho\|_{W^{10,1}}\}$$

where $\gamma_\varepsilon = \frac{1}{25}\left(6+\sqrt{2}\,2^6\,5\,|\alpha|^2+\varepsilon 4^5\sqrt{24}\right)^6$.

5.3 Application: entropic and capacity continuity bounds

Here, we provide one basic application to the perturbation bounds found in this section. We recall the definition of the energy-constrained diamond norm:

Definition 5.4 (see [61, 68]) Given $E \ge 0$ and any two completely positive, trace-preserving maps $\mathcal{N}, \mathcal{M} : \mathcal{T}_1(\mathcal{H}_m) \to \mathcal{T}_1(\mathcal{H}_m)$, their energy constrained diamond norm distance is defined as

$$\|\mathcal{N}-\mathcal{M}\|_\diamond^E := \sup_{\mathcal{H}_r} \sup_{\substack{\rho\in\mathcal{D}(\mathcal{H}_m\otimes\mathcal{H}_r) \\ \operatorname{tr}\left[\rho_m N_m\right]\le E}} \left\|(\mathcal{N}-\mathcal{M})\otimes \mathrm{id}_R(\rho)\right\|_1, \tag{84}$$

where $N_m := \sum_{i=1}^m a_i^\dagger a_i$ denotes the total photon number operator, and where the supremum is over all bipartite states ρ_{mr} on $\mathcal{H}_\otimes\mathcal{H}_r$ with reduced state ρ_m on $\mathcal{H}_m$ of average total photon number at most E, for some arbitrary separable Hilbert space $\mathcal{H}_r$.

Not that we denote the set of quantum states over a separable Hilbert space as $\mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{T}_{1,\mathrm{sa}}(\mathcal{H}) \;:\; \rho \ge 0, \operatorname{tr}[\rho] = 1\}$ in this section. In other words, the energy-constrained diamond norm is a measure of distinguishability between quantum channels with entanglement assistance, and where the input states used for this task are restricted to a physically relevant set of energy-limited states. By the same reasoning as for the usual diamond norm, the supremum in (84) can be restricted to $\mathcal{H}_r \cong \mathcal{H}_m$, and the optimization can be restricted to pure states on $\mathcal{H}_m \otimes \mathcal{H}_r$. Moreover, it turns out that the above definition is equivalent to one where the input state is energy limited on both m and r, as introduced in [55]:

$$\|\!|\mathcal{N}-\mathcal{M}\|\!|_\diamond^E := \sup_{\substack{\rho\in\mathcal{D}(\mathcal{H}_m\otimes\mathcal{H}_{m'})\\ \operatorname{tr}\left[\rho(N_m+N_{m'})\right]\le E}} \left\|(\mathcal{N}-\mathcal{M})\otimes \mathrm{id}_{m'}(\rho)\right\|_1 ,$$

for $\mathcal{H}_{m'} \cong \mathcal{H}_m$, and where the difference with (84) lies in the input states in the above optimization are energy constrained in both their inputs. Clearly, from the definitions we have that

$$\|\!|\mathcal{N}-\mathcal{M}\|\!|_\diamond^E \le \|\mathcal{N}-\mathcal{M}\|_\diamond^E \le \|\!|\mathcal{N}-\mathcal{M}\|\!|_\diamond^{2E} . \tag{85}$$

The second bound above simply results from taking a state $\rho_{mm'}$ for which $\operatorname{tr}\left[\rho_m N_m\right] \le E$ and unitarily rotating the second register so that $\rho'_m = \rho_m$, and therefore $\operatorname{tr}\left[\rho_{m'} N_{m'}\right] \le E$, too.

Similarly, in the next lemma, we establish a straightforward connection between the energy-constrained diamond norm distance between two channels and certain norms between Sobolev spaces:

Lemma 5.5 *For any two completely positive, trace preserving maps $\mathcal{N},\mathcal{M} : \mathcal{T}_1(\mathcal{H}_1) \to \mathcal{T}_1(\mathcal{H}_1)$ and $E \ge 0$,*

$$\|\mathcal{N}-\mathcal{M}\|_\diamond^E = (1+E) \sup_{\rho\in\mathcal{D}} \frac{\|(\mathcal{N}-\mathcal{M})\otimes \mathrm{id}_{M'}(\rho)\|_1}{\|\rho\|_{W^{(2,0),1}}} ,$$

where $\mathcal{D} := \mathcal{D}(\mathcal{H}_m \otimes \mathcal{H}_{m'})$ above. Similar identities hold in multi-mode settings.

Proof. This is direct from

$$\begin{aligned}\|\mathcal{N}-\mathcal{M}\|_\diamond^E &= \sup_{\substack{\rho\in\mathcal{D}\\ \operatorname{tr}[\rho N]\le E}} \|(\mathcal{N}-\mathcal{M})\otimes \mathrm{id}(\rho)\|_1 \\ &= \sup_{\substack{\rho\in\mathcal{D}\\ \|\rho\|_{W^{(2,0),1}}\le 1+E}} \|(\mathcal{N}-\mathcal{M})\otimes \mathrm{id}(\rho)\|_1 .\end{aligned}$$

□

Lemma 5.5 can be used in combination with perturbation bounds and entropic continuity bounds like those derived e.g. in [69, 68, 62, 61] to control the deviation of energy-constrained channel capacities in presence of a Lindbladian perturbation. Since these considerations go beyond the scope of the present paper, we do not pursue them further. Instead, we provide a basic illustration of the method we propose in the case of Gaussian perturbations of Gaussian semigroups by combining Proposition 5.1 with Lemma 5.5.

Corollary 5.6 *Let $(\mathcal{L}_{\mathrm{qOU}}, \mathcal{T}_f)$ be the generator of the quantum Ornstein Uhlenbeck semigroup with $\lambda > \mu \ge 0$ and $(\varepsilon\mathcal{L}_G, \mathcal{T}_f) := (\varepsilon\mathcal{L}[\gamma a + \eta a^\dagger], \mathcal{T}_f)$ a Gaussian perturbation with $\gamma, \eta \in \mathbb{R}$, $\varepsilon > 0$. Then, assuming $\lambda^2 - \mu^2 + |\gamma|^2 - |\eta|^2 > 0$ as in Proposition 5.1, there exist uniformly bounded functions $C(\varepsilon)$, $D(\varepsilon)$ such that, for all $t \ge 0$:*

$$\left\| e^{t\mathcal{L}_{\mathrm{qOU}}} - e^{t(\mathcal{L}_{\mathrm{qOU}}+\varepsilon\mathcal{L}_G)} \right\|_\diamond^E \le (1+E)\varepsilon\, C(\varepsilon) \max\left\{1, D(\varepsilon)\right\} . \tag{86}$$

6 Discussion and open questions

In this paper, we proved a generation theorem for bosonic Sobolev preserving quantum dynamical semigroups, which we also extended to time-dependent evolutions on bosonic multi-mode systems. The property of Sobolev preservation allowed all-time perturbation bounds, which we later applied to bosonic Gaussian semigroups, the quantum Ornstein Uhlenbeck semigroup, and dissipative CAT-qubit dynamics. As a consequence, we also established convergence results for the adherent points in the limit of large times. For this, we only required two physical assumptions on the generator (i) that it is in GKSL form with a Hamiltonian and jump-operators defined by polynomials in the annihilation and creation operators; and (ii) a bound on the output moments of the generator in terms of its input-moments. An interesting question for further work would be to further generalize the result by allowing different Sobolev spaces on which the semigroups are defined, and by this enlarge the class of generators covered by the generation theorem. In this line of research, a more detailed consideration of the Toffoli gate would be interesting.

7 Declarations

Competing interests:

T.M. and C.R. received funding from the Munich Center for Quantum Sciences and Technology, C.R. from the Humboldt Foundation, and all authors from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 470903074 – TRR 352. The authors have no further relevant financial or non-financial interests to disclose.

Data availability statement:

No datasets were generated or analyzed during the current study.

8 References

[1] J. Aasi et al. 'Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light'. In: *Nature Photonics* 7.8 (July 2013), pp. 613–619. ISSN: 1749-4893. DOI: 10.1038/nphoton.2013.177.

[2] J. Agredo, F. Fagnola, and D. Poletti. 'Gaussian Quantum Markov Semigroups on a One-Mode Fock Space: Irreducibility and Normal Invariant States'. In: *Open Systems & Information Dynamics* 28.01 (Mar. 2021), p. 2150001. ISSN: 1793-7191. DOI: 10.1142/s1230161221500013.

[3] J. Alber. 'On implemented semigroups'. In: *Semigroup Forum* 63.3 (July 2001), pp. 371–386. ISSN: 1432-2137. DOI: 10.1007/s002330010082.

[4] R. Azouit, A. Sarlette, and P. Rouchon. 'Convergence and adiabatic elimination for a driven dissipative quantum harmonic oscillator'. In: *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, Dec. 2015, pp. 6447–6453. DOI: 10.1109/cdc.2015.7403235.

[5] R. Azouit, A. Sarlette, and P. Rouchon. 'Well-posedness and convergence of the Lindblad master equation for a quantum harmonic oscillator with multi-photon drive and damping'. In: *ESAIM: Control, Optimisation and Calculus of Variations* 22.4 (Aug. 2016), pp. 1353–1369. ISSN: 1262-3377. DOI: 10.1051/cocv/2016050.

[6] S. Becker, N. Datta, L. Lami, and C. Rouzé. 'Convergence Rates for the Quantum Central Limit Theorem'. In: *Communications in Mathematical Physics* 383.1 (Feb. 2021), pp. 223–279. ISSN: 1432-0916. DOI: 10.1007/s00220-021-03988-1.

[7] C. Berdou, A. Murani, U. Réglade, W. Smith, M. Villiers, J. Palomo, M. Rosticher, A. Denis, P. Morfin, M. Delbecq, T. Kontos, N. Pankratova, F. Rautschke, T. Peronnin, L.-A. Sellem, P. Rouchon, A. Sarlette, M. Mirrahimi, P. Campagne-Ibarcq, S. Jezouin, R. Lescanne, and Z. Leghtas. 'One Hundred Second Bit-Flip Time in a Two-Photon Dissipative Oscillator'. In: *PRX Quantum* 4.2 (June 2023), p. 020350. ISSN: 2691-3399. DOI: 10.1103/prxquantum.4.020350.

[8] O. Bratteli and D. W. Robinson. *Operator Algebras and Quantum Statistical Mechanics*. Springer Berlin Heidelberg, 1981. ISBN: 9783662090893. DOI: 10.1007/978-3-662-09089-3.

[9] O. Bratteli and D. W. Robinson. *Operator Algebras and Quantum Statistical Mechanics 1*. Springer Berlin Heidelberg, 1987. ISBN: 9783662025208. DOI: 10.1007/978-3-662-02520-8.

[10] S. L. Braunstein and P. van Loock. 'Quantum information with continuous variables'. In: *Reviews of Modern Physics* 77.2 (June 2005), pp. 513–577. ISSN: 1539-0756. DOI: 10.1103/revmodphys.77.513.

[11] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret. 'Quantum error correction of a qubit encoded in grid states of an oscillator'. In: *Nature* 584.7821 (Aug. 2020), pp. 368–372. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2603-3.

[12] R. Carbone and F. Fagnola. 'The Feller Property of a Class of Quantum Markov Semigroups II'. In: *Quantum Probability and Infinite-Dimensional Analysis*. WORLD SCIENTIFIC, Jan. 2003, pp. 57–76. DOI: 10.1142/9789812704290`0005.

[13] R. Carbone and E. Sasso. 'Hypercontractivity for a quantum Ornstein–Uhlenbeck semigroup'. In: *Probability Theory and Related Fields* 140.3–4 (May 2007), pp. 505–522. ISSN: 1432-2064. DOI: 10.1007/s00440-007-0073-2.

[14] E. A. Carlen and J. Maas. 'Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance'. In: *Journal of Functional Analysis* 273.5 (Sept. 2017), pp. 1810–1869. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2017.05.003.

[15] C. Chamberland, K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller, G. Refael, J. Preskill, L. Jiang, A. H. Safavi-Naeini, O. Painter, and F. G. Brandão. 'Building a Fault-Tolerant Quantum Computer Using Concatenated Cat Codes'. In: *PRX Quantum* 3.1 (Feb. 2022), p. 010329. ISSN: 2691-3399. DOI: 10.1103/prxquantum.3.010329.

[16] A. M. Chebotarev. 'A Priori Estimates for Quantum Dynamical Semigroups'. In: *Theoretical and Mathematical Physics* 134.2 (2003), pp. 160–165. ISSN: 0040-5779. DOI: 10.1023/a:1022219919058.

[17] A. Chebotarev and F. Fagnola. 'Sufficient Conditions for Conservativity of Minimal Quantum Dynamical Semigroups'. In: *Journal of Functional Analysis* 153.2 (Mar. 1998), pp. 382–404. ISSN: 0022-1236. DOI: 10.1006/jfan.1997.3189.

[18] F. Cipriani, F. Fagnola, and J. M. Lindsay. 'Spectral Analysis and Feller Property for Quantum Ornstein-Uhlenbeck Semigroups'. In: *Communications in Mathematical Physics* 210.1 (Mar. 2000), pp. 85–105. ISSN: 1432-0916. DOI: 10.1007/s002200050773.

[19] J. B. Conway. *A Course in Functional Analysis*. Springer New York, 2007. ISBN: 9781475743838. DOI: 10.1007/978-1-4757-4383-8.

[20] E. Davies. 'Quantum dynamical semigroups and the neutron diffusion equation'. In: *Reports on Mathematical Physics* 11.2 (Apr. 1977), pp. 169–188. ISSN: 0034-4877. DOI: 10.1016/0034-4877(77)90059-3.

[21] G. De Palma and S. Huber. 'The conditional entropy power inequality for quantum additive noise channels'. In: *Journal of Mathematical Physics* 59.12 (Dec. 2018). ISSN: 1089-7658. DOI: 10.1063/1.5027495.

[22] C. Eichler, D. Bozyigit, C. Lang, M. Baur, L. Steffen, J. M. Fink, S. Filipp, and A. Wallraff. 'Observation of Two-Mode Squeezing in the Microwave Frequency Domain'. In: *Physical Review Letters* 107.11 (Sept. 2011), p. 113601. ISSN: 1079-7114. DOI: 10.1103/physrevlett.107.113601.

[23] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer-Verlag, 2000. ISBN: 0387984631. DOI: 10.1007/b97696.

[24] L. C. Evans. *Partial differential equations*. Second edition. Graduate studies in mathematics 19. Literaturverzeichnis: Seite 689-701. - Index. Providence, Rhode Island: American Mathematical Society, 2010. 712 pp. ISBN: 9780821849743.

[25] F. Fagnola. 'Quantum Markov semigroups'. In: *Proyecciones (Antofagasta)* 18.3 (1999), pp. 29–74. ISSN: 0717-6279. DOI: 10.22199/s07160917.1999.0003.00004.

[26] F. Fagnola, R. Rebolledo, and C. Saavedra. 'Quantum flows associated to master equations in quantum optics'. In: *Journal of Mathematical Physics* 35.1 (Jan. 1994), pp. 1–12. ISSN: 1089-7658. DOI: 10.1063/1.530788.

[27] E. Flurin, V. V. Ramasesh, S. Hacohen-Gourgy, L. S. Martin, N. Y. Yao, and I. Siddiqi. 'Observing Topological Invariants Using Quantum Walks in Superconducting Circuits'. In: *Physical Review X* 7.3 (Aug. 2017), p. 031023. ISSN: 2160-3308. DOI: 10.1103/physrevx.7.031023.

[28] L. Gao and C. Rouzé. *Ricci curvature of quantum channels on non-commutative transportation metric spaces*. 2021. DOI: 10.48550/ARXIV.2108.10609.

[29] R. Gordon. 'Survey Article Riemann Integration in Banach Spaces'. In: *Rocky Mountain Journal of Mathematics* 21.3 (Sept. 1991). ISSN: 0035-7596. DOI: 10.1216/rmjm/1181072923.

[30] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. 'Completely positive dynamical semigroups of N-level systems'. In: *Journal of Mathematical Physics* 17.5 (May 1976), pp. 821–825. ISSN: 1089-7658. DOI: 10.1063/1.522979.

[31] D. Gottesman, A. Kitaev, and J. Preskill. 'Encoding a qubit in an oscillator'. In: *Physical Review A* 64.1 (June 2001), p. 012310. ISSN: 1094-1622. DOI: 10.1103/physreva.64.012310.

[32] J. Guillaud, J. Cohen, and M. Mirrahimi. 'Quantum computation with cat qubits'. In: *SciPost Physics Lecture Notes* (June 2023). ISSN: 2590-1990. DOI: 10.21468/scipostphyslectnotes.72.

[33] J. Guillaud and M. Mirrahimi. 'Repetition Cat Qubits for Fault-Tolerant Quantum Computation'. In: *Physical Review X* 9.4 (Dec. 2019), p. 041053. ISSN: 2160-3308. DOI: 10.1103/physrevx.9.041053.

[34] T. Heinosaari, A. S. Holevo, and M. M. Wolf. *The semigroup structure of Gaussian channels*. 2009. DOI: 10.48550/ARXIV.0909.0408.

[35] E. Hille and R. S. Phillips. *Functional analysis and semi-groups*. Rev. ed (Online-Ausg.) AMS ebook collection. Bibliography: p. 781-796. - Electronic reproduction; Provid-

ence, Rhode Island; American Mathematical Society; 2012. - Description based on print version record. Providence: American Mathematical Society, 2012. 1808 pp. ISBN: 9780821833957.

[36] A. Holevo and R. Werner. 'Evaluating capacities of bosonic Gaussian channels'. In: *Physical Review A* 63.3 (Feb. 2001), p. 032312. ISSN: 1094-1622. DOI: 10.1103/physreva.63.032312.

[37] K. Huang, H. Le Jeannic, J. Ruaudel, V. B. Verma, M. D. Shaw, F. Marsili, S. W. Nam, E. Wu, H. Zeng, Y.-C. Jeong, R. Filip, O. Morin, and J. Laurat. 'Optical Synthesis of Large-Amplitude Squeezed Coherent-State Superpositions with Minimal Resources'. In: *Physical Review Letters* 115.2 (July 2015), p. 023602. ISSN: 1079-7114. DOI: 10.1103/physrevlett.115.023602.

[38] R. L. Hudson and K. R. Parthasarathy. 'Quantum Ito's formula and stochastic evolutions'. In: *Communications in Mathematical Physics* 93.3 (Sept. 1984), pp. 301–323. ISSN: 1432-0916. DOI: 10.1007/bf01258530.

[39] A. Joshi, K. Noh, and Y. Y. Gao. 'Quantum information processing with bosonic qubits in circuit QED'. In: *Quantum Science and Technology* 6.3 (Apr. 2021), p. 033001. ISSN: 2058-9565. DOI: 10.1088/2058-9565/abe989.

[40] T. Kato. *Perturbation Theory for Linear Operators*. Springer Berlin Heidelberg, 1995. ISBN: 9783642662829. DOI: 10.1007/978-3-642-66282-9.

[41] E. Kreyszig. *Introductory functional analysis with applications*. Wiley classics library ed. Wiley classics library. New York [u.a.]: Wiley, 1989. 688 pp. ISBN: 0471507318.

[42] Y. Kurochkin, A. S. Prasad, and A. I. Lvovsky. 'Distillation of The Two-Mode Squeezed State'. In: *Physical Review Letters* 112.7 (Feb. 2014), p. 070402. ISSN: 1079-7114. DOI: 10.1103/physrevlett.112.070402.

[43] L. Lafleche. 'On quantum Sobolev inequalities'. In: *Journal of Functional Analysis* 286.10 (May 2024), p. 110400. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2024.110400.

[44] L. Lami and M. M. Wilde. 'Exact solution for the quantum and private capacities of bosonic dephasing channels'. In: *Nature Photonics* 17.6 (Apr. 2023), pp. 525–530. ISSN: 1749-4893. DOI: 10.1038/s41566-023-01190-4.

[45] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret. 'Confining the state of light to a quantum manifold by engineered two-photon loss'. In: *Science* 347.6224 (Feb. 2015), pp. 853–857. ISSN: 1095-9203. DOI: 10.1126/science.aaa2085.

[46] G. Lindblad. 'On the generators of quantum dynamical semigroups'. In: *Communications in Mathematical Physics* 48.2 (June 1976), pp. 119–130. ISSN: 1432-0916. DOI: 10.1007/bf01608499.

[47] K. C. McCormick, J. Keller, S. C. Burd, D. J. Wineland, A. C. Wilson, and D. Leibfried. 'Quantum-enhanced sensing of a single-ion mechanical oscillator'. In: *Nature* 572.7767 (July 2019), pp. 86–90. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1421-y.

[48] V. Meyer, M. A. Rowe, D. Kielpinski, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. 'Experimental Demonstration of Entanglement-Enhanced Rotation Angle Estimation Using Trapped Ions'. In: *Physical Review Letters* 86.26 (June 2001), pp. 5870–5873. ISSN: 1079-7114. DOI: 10.1103/physrevlett.86.5870.

[49] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin. 'New Class of Quantum Error-Correcting Codes for a Bosonic Mode'.

In: *Physical Review X* 6.3 (July 2016), p. 031006. ISSN: 2160-3308. DOI: 10.1103/physrevx.6.031006.

[50] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret. 'Dynamically protected cat-qubits: a new paradigm for universal quantum computation'. In: *New Journal of Physics* 16.4 (Apr. 2014), p. 045014. ISSN: 1367-2630. DOI: 10.1088/1367-2630/16/4/045014.

[51] T. Möbus and C. Rouzé. 'Optimal Convergence Rate in the Quantum Zeno Effect for Open Quantum Systems in Infinite Dimensions'. In: *Annales Henri Poincaré* 24.5 (Oct. 2022), pp. 1617–1659. ISSN: 1424-0661. DOI: 10.1007/s00023-022-01241-6.

[52] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. 'Extending the lifetime of a quantum bit with error correction in superconducting circuits'. In: *Nature* 536.7617 (July 2016), pp. 441–445. ISSN: 1476-4687. DOI: 10.1038/nature18949.

[53] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier. 'Generating Optical Schrödinger Kittens for Quantum Information Processing'. In: *Science* 312.5770 (Apr. 2006), pp. 83–86. ISSN: 1095-9203. DOI: 10.1126/science.1122858.

[54] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer New York, 1983. ISBN: 9781461255611. DOI: 10.1007/978-1-4612-5561-1.

[55] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi. 'Fundamental limits of repeaterless quantum communications'. In: *Nature Communications* 8.1 (Apr. 2017). ISSN: 2041-1723. DOI: 10.1038/ncomms15043.

[56] G. Prato and G. Geymonat. *Hyperbolicity: Lectures given at the Centro Internazionale Matematico Estivo (C.I.M.E.), held in Cortona (Arezzo), Italy, June 24 – July 2, 1976*. Springer Berlin Heidelberg, 2011. ISBN: 9783642111051. DOI: 10.1007/978-3-642-11105-1.

[57] J. Preskill. *Quantum computation*. University Lecture. 2021. URL: http://theory.caltech.edu/~preskill/ph219/ph219_2021-22.html.

[58] C. Reimer, M. Kues, P. Roztocki, B. Wetzel, F. Grazioso, B. E. Little, S. T. Chu, T. Johnston, Y. Bromberg, L. Caspani, D. J. Moss, and R. Morandotti. 'Generation of multiphoton entangled quantum states by means of integrated frequency combs'. In: *Science* 351.6278 (Mar. 2016), pp. 1176–1180. ISSN: 1095-9203. DOI: 10.1126/science.aad8532.

[59] M. Rosati, A. Mari, and V. Giovannetti. 'Narrow bounds for the quantum capacity of thermal attenuators'. In: *Nature Communications* 9.1 (Oct. 2018). ISSN: 2041-1723. DOI: 10.1038/s41467-018-06848-0.

[60] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf. 'Fault-tolerant detection of a quantum error'. In: *Science* 361.6399 (July 2018), pp. 266–270. ISSN: 1095-9203. DOI: 10.1126/science.aat3996.

[61] M. E. Shirokov. 'On the Energy-Constrained Diamond Norm and Its Application in Quantum Information Theory'. In: *Problems of Information Transmission* 54.1 (Jan. 2018), pp. 20–33. ISSN: 1608-3253. DOI: 10.1134/s0032946018010027.

[62] M. E. Shirokov. 'Tight uniform continuity bounds for the quantum conditional mutual information, for the Holevo quantity, and for capacities of quantum channels'. In: *Journal of Mathematical Physics* 58.10 (Oct. 2017). ISSN: 1089-7658. DOI: 10.1063/1.4987135.

[63] I. Siemon, A. S. Holevo, and R. F. Werner. 'Unbounded Generators of Dynamical Semigroups'. In: *Open Systems & Information Dynamics* 24.04 (Dec. 2017), p. 1740015. ISSN: 1793-7191. DOI: 10.1142/s1230161217400157.

[64] B. Simon. *Operator Theory: A Comprehensive Course in Analysis. Barry.* Online-Ausg. Vol. part 4. AMS ebook collection. Providence, Rhode Island: AMS, American Mathematical Society, 2015. 1749 pp. ISBN: 9781470427634.

[65] O. Szehr and M. M. Wolf. 'Perturbation bounds for quantum Markov processes and their fixed points'. In: *Journal of Mathematical Physics* 54.3 (Mar. 2013). ISSN: 1089-7658. DOI: 10.1063/1.4795112.

[66] M. Takeoka, S. Guha, and M. M. Wilde. 'Fundamental rate-loss tradeoff for optical quantum key distribution'. In: *Nature Communications* 5.1 (Oct. 2014). ISSN: 2041-1723. DOI: 10.1038/ncomms6235.

[67] M. M. Wilde, M. Tomamichel, and M. Berta. 'Converse Bounds for Private Communication Over Quantum Channels'. In: *IEEE Transactions on Information Theory* 63.3 (Mar. 2017), pp. 1792–1817. ISSN: 1557-9654. DOI: 10.1109/tit.2017.2648825.

[68] A. Winter. *Energy-constrained diamond norm with applications to the uniform continuity of continuous variable channel capacities.* 2017. DOI: 10.48550/ARXIV.1712.10267.

[69] A. Winter. 'Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints'. In: *Communications in Mathematical Physics* 347.1 (Mar. 2016), pp. 291–313. ISSN: 1432-0916. DOI: 10.1007/s00220-016-2609-8.

[70] M. M. Wolf, D. Pérez-García, and G. Giedke. 'Quantum Capacities of Bosonic Channels'. In: *Physical Review Letters* 98.13 (Mar. 2007), p. 130501. ISSN: 1079-7114. DOI: 10.1103/physrevlett.98.130501.

[71] J. Zhang, M. Um, D. Lv, J.-N. Zhang, L.-M. Duan, and K. Kim. 'NOON States of Nine Quantized Vibrations in Two Radial Modes of a Trapped Ion'. In: *Physical Review Letters* 121.16 (Oct. 2018), p. 160502. ISSN: 1079-7114. DOI: 10.1103/physrevlett.121.160502.

[72] L. Zhong, E. P. Menzel, R. Di Candia, P. Eder, M. Ihmig, A. Baust, M. Haeberlein, E. Hoffmann, K. Inomata, T. Yamamoto, Y. Nakamura, E. Solano, F. Deppe, A. Marx, and R. Gross. 'Squeezing with a flux-driven Josephson parametric amplifier'. In: *New Journal of Physics* 15.12 (Dec. 2013), p. 125013. ISSN: 1367-2630. DOI: 10.1088/1367-2630/15/12/125013.

A Semigroup perturbation theory

Here, we prove Theorem 2.11 stated in the preliminary section.

Theorem A.1 *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ and $(\mathcal{L}+\mathcal{K}, \mathcal{D}(\mathcal{L}+\mathcal{K}))$ be two generators of C_0-semigroups on $\mathcal{X}$, for an operator $(\mathcal{K}, \mathcal{D}(\mathcal{K}))$. Moreover, let $(\mathcal{W}, \mathcal{D}(\mathcal{W}))$ be an invertible operator on $\mathcal{X}$ with bounded inverse, such that $\mathcal{D}(\mathcal{W})$ is an $\mathcal{L}+\mathcal{K}$-admissible subspace (see Definition 2.9) and such that $\mathcal{K}\mathcal{W}^{-1}$ is bounded. Then, for all $x \in \mathcal{D}(\mathcal{W})$,*

$$\|(e^{t\mathcal{L}} - e^{t(\mathcal{L}+\mathcal{K})})x\|_{\mathcal{X}} \leq t\|\mathcal{K}\mathcal{W}^{-1}\|_{\mathcal{X}\to\mathcal{X}} \int_0^1 \|e^{(1-s)t\mathcal{L}}\|_{\mathcal{X}\to\mathcal{X}} \, \|e^{st(\mathcal{L}-\mathcal{K})}\|_{\mathcal{W}\to\mathcal{W}} \, ds \|x\|_{\mathcal{W}}.$$

and especially

$$(e^{t\mathcal{L}} - e^{t(\mathcal{L}+\mathcal{K})})x = t\int_0^1 e^{(1-s)t\mathcal{L}}\mathcal{K}e^{st(\mathcal{L}-\mathcal{K})}x \, ds.$$

Proof. For simplicity, we assume that $t = 1$. To be able to use the integral form, we start by considering the following vector-valued map:

$$[0,1] \ni s \mapsto e^{(1-s)\mathcal{L}}\mathcal{K}e^{s(\mathcal{L}+\varepsilon\mathcal{K})}x.$$

It is clear that $s \mapsto e^{(1-s)\mathcal{L}}$ is a strongly continuous map of bounded operators. Moreover,

$$\|\mathcal{K}e^{s(\mathcal{L}+\mathcal{K})}x - \mathcal{K}e^{(s+s')(\mathcal{L}+\mathcal{K})}x\|_{\mathcal{X}} \leq \|\mathcal{K}\mathcal{W}^{-1}\|_{\mathcal{X}} \, \|(e^{s(\mathcal{L}+\mathcal{K})}x - e^{(s+s')(\mathcal{L}+\mathcal{K})}x)\|_{\mathcal{W}}$$

shows that $s \mapsto \mathcal{K}e^{s(\mathcal{L}+\mathcal{K})}$ is strongly continuous because $s \mapsto e^{s(\mathcal{L}+\mathcal{K})}$ defines a C_0-semigroup on $(\mathcal{D}(\mathcal{W}), \|\cdot\|_{\mathcal{W}})$. Then, for every converging sequence $(s_n)_{n\in\mathbb{N}} \to s$ for $n \to \infty$ the set

$$\left\{\mathcal{K}e^{s_n(\mathcal{L}+\mathcal{K})}x \,|\, n \in \mathbb{N}\right\}$$

is relatively compact in $\mathcal{X}$. Therefore, the strong continuity of $s \mapsto e^{(1-s)\mathcal{L}}$ is equivalent to uniform continuity by [23, Prop. A3], the considered vector-valued map is continuous and we can use the fundamental theorem of calculus for the generalized Riemann integral so that

$$\|(e^{\mathcal{L}} - e^{\mathcal{L}+\mathcal{K}})x\|_{\mathcal{X}} \leq \varepsilon \int_0^1 \|e^{(1-s)\mathcal{L}}\mathcal{K}e^{s(\mathcal{L}+\varepsilon\mathcal{K})}x\|_{\mathcal{X}} ds$$

proves the theorem. □

B Bosonic single mode system

In this section, we prove some basic properties of polynomials of annihilation and creation operators. We shortly repeat the normal form of our polynomials in a and $a^\dagger$ given in Equation (14):

$$p(a\,, a^\dagger) = \sum_{i+2j\leq \deg(p)} \lambda_{ij}(a^\dagger)^i N^j + \sum_{k+2l\leq\deg(p)} \mu_{kl} N^l a^k$$

with coefficients $\lambda_{ij}, \mu_{kl} \in \mathbb{C}$. The modification considered in our proof (see Lemma 3.2) is given in Equation (16) by

$$\tilde{p}(a, a^\dagger) \coloneqq N^{2d} + p(a, a^\dagger)\,.$$

Then, we start proving a simple representation of a domain of $p(a, a^\dagger)$ and $\tilde{p}(a, a^\dagger)$ which extends the domain $\mathcal{H}_f$: For $n \in \mathbb{N}$

$$p(a\,, a^\dagger)\,|n\rangle = \sum_{i+2j\leq d} \lambda_{ij} n^j \sqrt{i!\binom{n+i}{i}}\,|n+i\rangle + \sum_{k+2l\leq d} \mu_{kl}(n-k)^l \sqrt{k!\binom{n}{k}}\,|n-k\rangle \qquad (87)$$

where $d := \deg(p)$. This directly implies for $|\phi\rangle = \sum_n \phi_n |n\rangle \in \mathcal{D}(N^{d/2})$

$$\begin{aligned} &p(a\,,a^\dagger)\,|\phi\rangle \\ &= \sum_{n=0}^{\infty} \sum_{i+2j\le d} \phi_n \lambda_{ij} n^j \sqrt{i!\binom{n+i}{i}}\,|n+i\rangle + \sum_{n=0}^{\infty}\sum_{k+2l\le d} \phi_n \mu_{kl} (n-k)^l \sqrt{k!\binom{n}{k}}\,|n-k\rangle \\ &= \sum_{n=0}^{\infty} \sum_{i+2j\le d} \left(\phi_{n-i}\lambda_{ij}(n-i)^j \sqrt{i!\binom{n}{i}} + \phi_{n+i}\mu_{ij} n^j \sqrt{i!\binom{n+i}{i}}\right) |n\rangle \ . \end{aligned}$$

Then, the leading order of the summands in n is maximal of order $d/2$ so that $\mathcal{D}(N^{d/2})$ is a domain of $p(a, a^\dagger)$. For the modified polynomial $\tilde{p}$, there is sequence of functions $R_n : \mathcal{H} \to \mathbb{R}$ with asymptotics strictly smaller than n^{4d} such that

$$\|\tilde{p}(a, a^\dagger)\,|\phi\rangle\,\|^2 = \sum_{n=0}^{\infty} |\phi_n|^2 n^{4d} + R(\phi)\,. \tag{88}$$

Having the domain above in mind, we can prove that p is closable and $\tilde{p}$ is a closed operator with core $\mathcal{H}_f$:

Lemma B.1 (Adjoint and core of polynomials of $a, a^\dagger$) *Let $p \in \mathbb{C}[X, Y]$ be a polynomial on $\mathbb{C}$ and $(p(a, a^\dagger), \mathcal{D}(N^{d/2}))$ the unbounded operator in normal form (15). Then, $p(a, a^\dagger)$ is closable and there is a $c \ge 0$ such that for all $|\phi\rangle \in \mathcal{D}(N^{d/2})$*

$$\|p(a, a^\dagger)\,|\phi\rangle\,\| \le c\|(\mathbb{1} + N)^{d/2}\,|\phi\rangle\,\|\,.$$

The modification $\tilde{p}(a, a^\dagger) = N^{2d} + p(a, a^\dagger)$ is closed with $\mathcal{D}(\tilde{p}(a, a^\dagger)) = \mathcal{D}(\tilde{p}(a, a^\dagger)^\dagger) = \mathcal{D}(N^{2d})$ and core $\mathcal{H}_f$.

Proof. First, note that the relative boundedness w.r.t. the number operator is a direct consequence of Equation (87). To prove that $p(a, a^\dagger)$ is closable, we show that $\mathcal{D}(N^{d/2}) \subset \mathcal{D}(p(a, a^\dagger)^\dagger)$: we recall that the adjoint is defined via boundedness of the functional

$$\mathcal{D}(p(a, a^\dagger)) \ni |\phi\rangle \mapsto \langle p(a, a^\dagger)\phi\,, \varphi\rangle$$

for $\phi \in \mathcal{D}(p(a, a^\dagger)^\dagger)$. Since for all $n, m \in \mathbb{N}$

$$\langle a\,n\,, m\rangle = \delta_{n,0}\delta_{n-1,m}\sqrt{n} = \delta_{n,m+1}\sqrt{m+1} = \langle n\,, a^\dagger\, m\rangle$$

and by the definition of the domains, we know that for all $|\phi\rangle\,, |\psi\rangle \in \mathcal{D}(N^{k+\frac{l}{2}})$

$$\langle N^k (a^\dagger)^l \phi\,, \psi\rangle = \langle \phi\,, a^l N^k \psi\rangle \quad \text{and} \quad \langle a^l N^k \phi\,, \psi\rangle = \langle \phi\,, N^k (a^\dagger)^l \psi\rangle\ .$$

By sesquilinearity of the scalar product and Equation (15), the above equations hold for all polynomials $p \in \mathbb{C}[X, Y]$ which proves $\mathcal{D}(N^{d/2}) \subset \mathcal{D}(p(a, a^\dagger)^\dagger)$. Since $\mathcal{D}(N^{d/2})$ is a dense subspace of $\mathcal{H}$, Theorem 7.1.1 in [64] shows that $p(a, a^\dagger)$ is closable. Next, we show that the modified polynomial

$$(N + \mathbb{1})^{2d} + p(a, a^\dagger)$$

is already closed. Actually, we show $\mathcal{D}(N^{2d}) = \mathcal{D}(\tilde{p}(a, a^\dagger)^\dagger)$, i.e. the domain is maximal, by contradiction. Assume that there exists a $\varphi \in \mathcal{D}(\tilde{p}(a, a^\dagger)^\dagger) \backslash \mathcal{D}(\tilde{p}(a, a^\dagger))$ and define P_M to be the projection on the first M Fock basis elements. Then, we use the representation in

Equation (14) so that, denoting by $\tilde{p}^\dagger$ the polynomial where we took the complex conjugate of the coefficients and swapped the coordinates,

$$P_M\tilde{p}^\dagger(a,a^\dagger) = P_M(N+\mathbb{1})^{2d}P_M + P_M\left(\sum_{i+2j\le d}\overline{\lambda}_{ij}N^ja^iP_{M-i} + \overline{\mu}_{ij}(a^\dagger)^iN^jP_{M+i}\right)P_{M+d}\,.$$

We then can define the state

$$|\phi_M\rangle := \frac{P_M\tilde{p}^\dagger(a,a^\dagger)\,|\varphi\rangle}{\|P_M\tilde{p}^\dagger(a,a^\dagger)\,|\varphi\rangle\,\|} \in \mathcal{H}_f$$

Next, we use ϕ_M to get a lower bound on the operator norm of

$$\mathcal{D}(p(a,a^\dagger)) \ni |\phi\rangle \mapsto \langle \tilde{p}(a,a^\dagger)\phi\,,\varphi\rangle$$

by

$$\sup_{\|\phi\|=1} |\,\langle \tilde{p}(a,a^\dagger)\phi\,,\varphi\rangle\,| \ge |\,\langle \tilde{p}(a,a^\dagger)\phi_M\,,\varphi\rangle\,| = |\,\langle \phi_M\,, P_M\tilde{p}^\dagger(a,a^\dagger)\varphi\rangle\,| = \|P_M\tilde{p}^\dagger(a,a^\dagger)\varphi\|.$$

Now, by definition of $\tilde{p}$ and denoting by φ_n the coefficients of $|\varphi\rangle$ in the Fock basis,

$$\|P_M\tilde{p}^\dagger(a,a^\dagger)\varphi\|^2 = \left\|\sum_{n=0}^{M+d}\varphi_nP_M\left((N+\mathbb{1})^{2d} + p^\dagger(a,a^\dagger)\right)|n\rangle\right\|^2,$$

where $p^\dagger$ is defined similarly to $\tilde{p}^\dagger$. By assumption $\phi \notin \mathcal{D}(N^{2d})$, so that the above sequence is diverging for $M\to\infty$ to infinity (see Equation (88)). This contradicts the assumption and shows $\mathcal{D}(\tilde{p}(a,a^\dagger)^\dagger) = \mathcal{D}(N^{2d})$ as well as $\tilde{p}(a,a^\dagger)^\dagger = \tilde{p}^\dagger(a,a^\dagger)$. Moreover, $p(a,a^\dagger)$ is by Theorem 7.1.1 in [64] a closed operator. Since $\{|n\rangle\}_{n\in\mathbb{N}}$ is an orthonormal basis and N a multiplication operator on that basis, we can immediately conclude that $\mathcal{H}_f$ is a core for N and further for all $(\mathbb{1}+N)^k$, $k\ge 0$. Since $\tilde{p}(a,a^\dagger)$ is closed w.r.t. $\mathcal{D}(N^{p/2+1})$, $\mathcal{H}_f$ is also a core of $\tilde{p}(a,a^\dagger)$. □

Having in mind that polynomials of the annihilation and creation are closed operators on certain domains, we use the canonical commutation relation $[a,a^\dagger] = \mathbb{1}$ in the following lemma:

Lemma B.2 *Let $l\in\mathbb{N}$, then the following hold on $\mathcal{H}_f$ and can be extended to maximal domains by taking limits*

$$\begin{aligned}(a^\dagger)^la^l &= (N-(l-1)\mathbb{1})(N-(l-2)\mathbb{1})\cdots(N-\mathbb{1})N\,,\\ a^l(a^\dagger)^l &= (N+\mathbb{1})(N+2\mathbb{1})\cdots(N+(l-1)\mathbb{1})(N+l\mathbb{1})\,.\end{aligned} \tag{89}$$

Proof. The above equalities can be proven by induction over $l\in\mathbb{N}$. The cases $l\in\{0,1\}$ are trivial by definition. Next assume that the equation holds for $l\in\mathbb{N}$, then by Equation (49)

$$(a^\dagger)^{l+1}a^{l+1} = (a^\dagger)^lNa^l = (N-l)(a^\dagger)^la^l = (N-l)\cdots N$$

which finishes the proof by induction. The second expression can be proven by induction as well, and the induction start is again clear by the CCR. Next, we assume the equation for $l\in\mathbb{N}$. Then, Equation (49) shows

$$a^{l+1}(a^\dagger)^{l+1} = a^l(N+\mathbb{1})(a^\dagger)^l = a^l(a^\dagger)^l(N+(l+1)\mathbb{1}) = (N+\mathbb{1})\cdots(N+(l+1)\mathbb{1})$$

which completes the induction. □

Lemma B.3 *Let $\ell_1, \ell_2, k_1, k_2 \in \mathbb{N}$ with $\min\{\ell_1, k_1\} = \min\{\ell_2, k_2\} = 0$, $z \in \mathbb{C}$, and $h : \mathbb{N}^2 \to \mathbb{R}$ a positive function that is increasing in each of its variables. Then,*

$$\begin{aligned} z a_1^{\ell_1} a_2^{\ell_2} h(N_1, N_2)(a_1^\dagger)^{k_1}(a_2^\dagger)^{k_2} + \overline{z} a_1^{k_1} a_2^{k_2} h(N_1, N_2)(a_1^\dagger)^{\ell_1}(a_2^\dagger)^{\ell_2} \\ \leq 2|z|\widetilde{h}_{m_1,m_2}(N_1 + m_1 I, N_2 + m_2 I)\,, \end{aligned}$$

where $m_1 := \max\{\ell_1, k_1\}$, $m_2 := \max\{\ell_2, k_2\}$ and

$$\widetilde{h}_{m_1,m_2}(n_1, n_2) = \prod_{j=n_1-m_1+1}^{n_1} \sqrt{j} \prod_{i=n_2-m_2+1}^{n_2} \sqrt{i}\;\; h(n_1, n_2)\, 1_{n_1 \geq m_1} 1_{n_2 \geq m_2}\,,$$

where we introduced the notation $1_{x \geq m}$ for the indicator function on the set $\{x : x \geq m\}$, and where by convention we take $\prod_{i=a}^{b} = 1$ when $a > b$.

Proof. We define $K := z a_1^{\ell_1} a_2^{\ell_2} h(N_1, N_2)(a_1^\dagger)^{k_1}(a_2^\dagger)^{k_2} + \overline{z} a_1^{k_1} a_2^{k_2} h(N_1, N_2)(a_1^\dagger)^{\ell_1}(a_2^\dagger)^{\ell_2}$ and represent it in the 2-mode Fock basis:

$$\begin{aligned} K &= \sum_{n_1,n_2} h(n_1, n_2)\left(z a_1^{\ell_1} a_2^{\ell_2}\, |n_1, n_2\rangle\langle n_1, n_2|\,(a_1^\dagger)^{k_1}(a_2^\dagger)^{k_2} + \overline{z} a_1^{k_1} a_2^{k_2}\, |n_1, n_2\rangle\langle n_1, n_2|\,(a_1^\dagger)^{\ell_1}(a_2^\dagger)^{\ell_2}\right) \\ &= \sum_{\substack{n_1 \geq m_1 \\ n_2 \geq m_2}} g_{\substack{\ell_1,\ell_2 \\ k_1,k_2}}(n_1, n_2) \\ &\qquad \left(z\,|n_1 - \ell_1, n_2 - \ell_2\rangle\langle n_1 - k_1, n_2 - k_2| + \overline{z}\,|n_1 - k_1, n_2 - k_2\rangle\langle n_1 - \ell_1, n_2 - \ell_2|\right)\,, \end{aligned}$$

where

$$\begin{aligned} &g_{\substack{\ell_1,\ell_2 \\ k_1,k_2}}(n_1, n_2) \\ &:= h(n_1, n_2)\sqrt{n_1 \ldots (n_1 - \ell_1 + 1) n_1 \ldots (n_1 - k_1 + 1) n_2 \ldots (n_2 - \ell_2 + 1) n_2 \ldots (n_2 - k_2 + 1)}\,. \end{aligned}$$

By assumption, since $\min\{\ell_1, k_1\} = \min\{\ell_2, k_2\} = 0$, we have that $g_{\substack{\ell_1,\ell_2 \\ k_1,k_2}}(n_1, n_2) = \widetilde{h}_{m_1,m_2}(n_1, n_2)$ for $n_1 \geq m_1$, $n_2 \geq m_2$, and

$$\begin{aligned} K = \sum_{n_1,n_2 \in \mathbb{N}} &\widetilde{h}_{m_1,m_2}(n_1 + m_1, n_2 + m_2) \\ &\left(z\,|n_1 + k_1, n_2 + k_2\rangle\langle n_1 + \ell_1, n_2 + \ell_2| + \overline{z}\,|n_1 + \ell_1, n_2 + \ell_2\rangle\langle n_1 + k_1, n_2 + k_2|\right)\,. \end{aligned}$$

Next, we consider the constituents of the above sum individually. Note that the operator

$$z\,|n_1 + k_1, n_2 + k_2\rangle\langle n_1 + \ell_1, n_2 + \ell_2| + \overline{z}\,|n_1 + \ell_1, n_2 + \ell_2\rangle\langle n_1 + k_1, n_2 + k_2| \tag{90}$$

$$\begin{pmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & * & 0 \\ 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \end{pmatrix}.$$

can be embedded into an operator on a two-dimensional space of the form

$$z\,|e_1\rangle\langle e_2| + \overline{z}\,|e_2\rangle\langle e_1|\,,$$

where $|e_1\rangle$ and $|e_2\rangle$ are orthonormal vectors. For $|e_1\rangle = |e_2\rangle$, $z + \overline{z} \leq 2|z|$ shows

$$z\,|e_1\rangle\langle e_2| + \overline{z}\,|e_2\rangle\langle e_1| \leq |z|(|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|)\,.$$

In the case $|e_1\rangle \neq |e_2\rangle$, we have

Eigenvalue	Eigenvectors
$\lvert z\rvert$	$\lvert\psi\rangle = \frac{1}{\sqrt{2}\lvert z\rvert}(\lvert z\rvert\,\lvert e_1\rangle + z\,\lvert e_2\rangle)$
$-\lvert z\rvert$	$\lvert\varphi\rangle = \frac{1}{\sqrt{2}\lvert z\rvert}(\lvert z\rvert\,\lvert e_1\rangle - z\,\lvert e_2\rangle)$

so that

$$z\,|e_2\rangle\langle e_1| + \overline{z}\,|e_1\rangle\langle e_2| = |z|\,|\psi\rangle\langle\psi| - |z|\,|\varphi\rangle\langle\varphi| \leq |z|\,|\psi\rangle\langle\psi| \leq |z|(|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|)\,.$$

This allows us to estimate

$$\begin{aligned} K &\leq \sum_{n_1,n_2\in\mathbb{N}} \widetilde{h}_{m_1,m_2}(n_1+m_1, n_2+m_2)|z|(|n_1+k_1, n_2+k_2\rangle\langle n_1+k_1, n_2+k_2| \\ &\qquad\qquad + |n_1+l_1, n_2+l_2\rangle\langle n_1+l_1, n_2+l_2|) \\ &\leq 2|z|\widetilde{h}_{m_1,m_2}(N_1+m_1 I, N_2+m2I) \end{aligned}$$

employing the monotonicity of $\widetilde{h}_{m_1,m_2}$ in both arguments in the last step. □

C Inequalities for power functions

Many bounds in Sections 4 and 5 can be deduced from bounds on real-valued functions acting on the spectrum of the number operator N. Especially, the following functions, first introduced in Equation (51), will require special attention: Let $l, k \in \mathbb{N}$, $f(x) = (x+1)^{k/2} 1_{x\geq -1}$, and

$$g_l(x) = \begin{cases} f(x) - f(x-l) & x \geq l-1; \\ f(x) & l-1 > x \geq 0; \\ 0 & 0 > x\,. \end{cases} \tag{91}$$

Lemma C.1 *Let g_l be defined in Equation (91) for $l, k \in \mathbb{N}$. Then, for all $k \geq 2$ and $x \in \mathbb{R}$*

$$g_l(x) \leq g_{l+1}(x)\,, \tag{92}$$

$$g_l(x-l) \leq g_l(x)\,. \tag{93}$$

Proof. By the monotonicity and non-negativity of $f(x) = (x+1)^{k/2} 1_{x\geq -1}$,

$$\left.\begin{matrix} x \geq l-1 & f(x) - f(x-l) \\ l-1 > x \geq 0 & f(x) \\ 0 > x & 0 \end{matrix}\right\} = g_l(x) \leq g_{l+1}(x) = \begin{cases} f(x) - f(x-(l+1)) & x \geq l \\ f(x) & l > x \geq 0 \\ 0 & 0 > x \end{cases},$$

which proves Inequality 92. For Inequality 93, we consider the following cases:

$$\left.\begin{matrix} f(x-l) - f(x-2l) \\ f(x-l) \\ 0 \\ 0 \end{matrix}\right\} = g_l(x-l) \leq g_l(x) = \begin{cases} f(x) - f(x-l) & x \geq 2l-1 \\ f(x) - f(x-l) & 2l-1 > x \geq l-1 \\ f(x) & l-1 > x \geq 0 \\ 0 & 0 > x \end{cases}.$$

For $x < l-1$, the inequalities are clear by the non-negativity of f. The case $l-1 \leq x < 2l-1$ follows by

$$2\frac{f(x-l)}{f(x)} = 2\left(1 - \frac{l}{x+1}\right)^{k/2} \leq 2\left(1-\frac{1}{2}\right)^{k/2} = 1$$

and the last case $x \geq 2l-1$ follows by monotonicity of g_l:

$$\frac{2}{k}g_l'(x-l) = (x-l)^{k/2-1} - (x-2k)^{k/2-1} \geq 0\,.$$

□

Next, we prove upper and lower bounds for g_l:

Lemma C.2 *Let $g_l : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be defined in Equation (91) for $l \in \mathbb{N}$. Then, for all $x \in \mathbb{R}$ and $k \in \mathbb{N}$,*

$$\left.\begin{array}{ll} x \geq l-1 & (x+1)^{k/2-1}\frac{kl}{2} - 1_{k\geq 3}(x+1)^{k/2-2}\,\frac{(kl)^2}{8} \\ x \geq l-1 & (x+1)^{k/2-1}l \\ l-1 > x \geq 0 & (x+1)^{k/2} \\ 0 > x & 0 \end{array}\right\} \leq g_l(x)$$

and

$$g_l(x) \leq \begin{cases} \frac{kl}{2}\left(1+1_{k=1}\right)(x+1)^{k/2-1} & x \geq 0 \\ (x+1)^{k/2} & x \geq 0 \\ 0 & 0 > x \end{cases}.$$

Proof. The case $k=0$ is trivial. We start with the upper bounds. By monotonicity of g_l, it is enough to prove the first upper bound just for $x \geq l-1$. For $k=1$,

$$g_l(x) = (x+1)^{-1/2}\frac{l}{2}\int_0^1 \left(1 - s\frac{l}{x+1}\right)^{-1/2} ds \leq (x+1)^{-1/2}\frac{l}{2}\int_0^1 (1-s)^{-1/2}\, ds = (x+1)^{-1/2}l.$$

For $k \geq 2$,

$$g_l(x) = \frac{k}{2}\int_0^l (x+1-s)^{k/2-1}ds \leq \frac{kl}{2}(x+1)^{k/2-1}\,,$$

which finishes the proof of the first upper bound. The other two bounds are obvious by definition. Next, we consider the lower bounds. The case $x < l-1$ is trivial so we are left with proving

$$(x+1)^{k/2-1}\frac{kl}{2} - \delta_{k\geq 3}(x+1)^{k/2-2}\frac{(kl)^2}{8} \leq g_l(x)$$

for $x \geq l-1$. For $k=1$, the integral representation can be lower bounded as

$$g_l(x) = (x+1)^{-1/2}\frac{l}{2}\int_0^1\left(1 - s\frac{l}{x+1}\right)^{-1/2} ds \geq (x+1)^{-1/2}\frac{l}{2}\,.$$

For $k=2$, it is again easy to calculate the quantity $g_l(x) = l$, and for $k=3$

$$\begin{aligned} g_l(x) &= (x+1)^{1/2}\frac{3l}{2}\int_0^1\left(1 - s_1\frac{l}{x+1}\right)^{1/2} ds_1 \\ &= (x+1)^{1/2}\frac{3l}{2} - (x+1)^{-1/2}l^2\frac{3}{4}\iint_0^1 s_1\left(1 - s_1 s_2\frac{l}{x+1}\right)^{-1/2} ds_2 ds_1 \\ &\geq (x+1)^{1/2}\frac{3l}{2} - (x+1)^{-1/2}l^2\frac{3}{4}\iint_0^1 s_1\left(1 - s_1 s_2\right)^{-1/2} ds_2 ds_1 \\ &= (x+1)^{1/2}\frac{3l}{2} - (x+1)^{-1/2}\frac{l^2}{2}. \end{aligned}$$

Finally, the case $k \geq 4$ is given by

$$\begin{aligned}
g_l(x) &= (x+1)^{k/2-1}\frac{kl}{2}\int_0^1 \left(1 - s_1\frac{l}{x+1}\right)^{k/2-1} ds_1 \\
&= (x+1)^{k/2-1}\frac{kl}{2} - (x+1)^{k/2-2}l^2\frac{k(k-2)}{4}\iint_0^1 s_1\left(1 - s_1 s_2\frac{l}{x+1}\right)^{k/2-2} ds_2 ds_1 \\
&\geq (x+1)^{k/2-1}\frac{kl}{2} - (x+1)^{k/2-2}l^2\frac{k(k-2)}{4}\int_0^1 s_1 ds_1 \\
&\geq (x+1)^{k/2-1}\frac{kl}{2} - (x+1)^{k/2-2}\frac{(kl)^2}{8}
\end{aligned}$$

which proves the first non-trivial lower bound for $x \geq l-1$. Next, we consider

$$g_l(x) \geq (x+1)^{k/2-1}l\,.$$

The inequality is obvious for $k < 2$ by the same idea as before and for $k \geq 2$

$$\begin{aligned}
g_l(x) &= (x+1)^{k/2-1}\frac{kl}{2}\int_0^1 \left(1 - s_1\frac{l}{x+1}\right)^{k/2-1} ds_1 \\
&\geq (x+1)^{k/2-1}\frac{kl}{2}\int_0^1 (1-s_1)^{k/2-1} ds_1 \\
&\geq (x+1)^{k/2-1}l
\end{aligned}$$

which ends the proof. □

Lemma C.3 *Let $l \in \mathbb{N}$ and $x \geq l$, then*

$$\begin{aligned}
(x+1)^l - \frac{(l+1)l}{2}(x+1)^{l-1} \leq \quad ((x+1)-l)\cdots((x+1)-1) \leq \quad & (x+1)^l \\
(x+1)^l \leq \quad (x+1)\cdots(x+1+(l-1) \leq \quad & l!(x+1)^l
\end{aligned}$$

Proof. To prove Lemma C.3, we redefine $y = x+1$ and rewrite the first product as

$$p_l(y) \coloneqq (y-l)\cdots(y-1) \eqqcolon y^l - \frac{(l+1)l}{2}y^{l-1} + r_{l-2}(y)$$

where r_{l-2} is a polynomial of degree $l-2$. The proof idea is to show that $r_{l-2}(y)$ is non-negative for all $y \geq l+1$, which proves the inequality. The non-negativity of the polynomial r_{l-2} can be proven by induction over l: The statement is directly clear for $l = 1$ and $l = 2$. Next, we assume that r_{l-2} is non-negative for all $x \geq l+1$ and show that r_{l-1} is for all $x \geq l+2$.

$$\begin{aligned}
p_{l+1}(y) &= (y-(l+1))p_l(y) \\
&= (y-(l+1))\left(y^l - \frac{(l+1)l}{2}y^{l-1} + r_{l-2}(y)\right) \\
&= y^{l+1} - \frac{(l+1)(l+2)}{2}y^l + \frac{(l+1)^2 l}{2}y^{l-1} + (y-(l+1))r_{l-2}(y) \\
&= y^{l+1} - \frac{(l+1)(l+2)}{2}y^l + r_{l-1}(y).
\end{aligned}$$

For the second product $(x+1)\cdots((x+1)+l-1)$ the lower bound is clear and the upper bound follows by

$$(x+1)(x+2)\cdots((x+1)+l-1) = (l-1)!\left(\frac{x}{1}+1\right)\cdots\left(\frac{x}{l}+1\right) \leq l!(x+1)^l.$$

□

D Technical lemmas for the quantum Sobolev spaces

Lemma D.1 (Continuity of $G(z)$) *Let $k_0 < k_1 \in \mathbb{R}_+$ and $T : W^{k_j,1} \to W^{k_j,1}$, be a linear map with $\|T\|_{W^{k_j,1}\to W^{k_j,1}} \leq M_j$, bounded by $M_j \geq 0$ for $j = 1,2$ respectively. Further let $\theta \in [0,1]$ and $k_\theta = (1-\theta)k_0 + \theta k_1$ and $x \in \mathcal{T}_f$, then the map*

$$\begin{aligned} G : S \coloneqq \{z \in \mathbb{C} \,:\, 0 \leq \mathrm{Re}(z) \leq 1\} &\to \mathcal{T}_{1,\mathrm{sa}} \\ z \mapsto G(z) &= (N+\mathbb{1})^{\frac{k(z)}{4}} T\Big((N+\mathbb{1})^{\frac{k_\theta - k(z)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k(z)}{4}}\Big)(N+\mathbb{1})^{\frac{k(z)}{4}} \end{aligned}$$

with $k(z) = (1-z)k_0 + zk_1$, is well-defined, uniformly bounded and continuous.

Proof. To prove the claim, we decompose G using the following auxiliary functions:

$$\begin{aligned} G_1 : S \times W^{k_1,1} &\to \mathcal{T}_{1,\mathrm{sa}} \\ (z,y) &\mapsto (N+\mathbb{1})^{\frac{k(z)}{4}} y (N+\mathbb{1})^{\frac{k(z)}{4}} \end{aligned} \tag{94}$$

and

$$\begin{aligned} G_2 : S &\to \mathcal{T}_f \subset W^{k_1,1} \\ z &\mapsto (N+\mathbb{1})^{\frac{k_\theta - k(z)}{4}} x (N+\mathbb{1})^{\frac{k_\theta - k(z)}{4}} . \end{aligned} \tag{95}$$

We clearly have that $G_1(z,\cdot) : W^{k_1,1} \to \mathcal{T}_{1,\mathrm{sa}}$ is a bounded linear map for all $z \in S$, since

$$\|G_1(z,y)\|_1 = \left\|(N+\mathbb{1})^{\frac{\mathrm{Re}(k(z))}{4}} y (N+\mathbb{1})^{\frac{\mathrm{Re}(k(z))}{4}}\right\|_1 \leq \left\|(N+\mathbb{1})^{\frac{k_1}{4}} y (N+\mathbb{1})^{\frac{k_1}{4}}\right\|_1 = \|y\|_{W^{k_1,1}} \tag{96}$$

where we used that $k_0 \leq \mathrm{Re}(k(z)) \leq k_1$ and $(N+\mathbb{1})^{i\frac{\mathrm{Im}(k(z))}{4}}$ is a unitary that can be absorbed into the norm. Next, we will show that $G_1(\cdot,y) : S \to \mathcal{T}_{1,\mathrm{sa}}$ is continuous for all $y \in W^{k_1,1}$. For that first note that, for $y \in \mathcal{T}_f$, the claim follows directly from the continuity of $z \mapsto (n+1)^{\frac{k(z)}{4}}$ with $n \in \mathbb{N}$ as a map from S to $\mathbb{C}$. This is because all the involved operators can be considered finite dimensional using a cut-off of the Fock-basis. For a general $y \in W^{k_1,1}$, we find $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{T}_f$, s.t. $y_n \to y$ in $W^{k_1,1}$, hence for all $n \in \mathbb{N}$

$$\begin{aligned} &\lim_{z\to z_0} \|G_1(z,y) - G_1(z_0,y)\|_1 \\ &\quad\leq \lim_{z\to z_0} \|G_1(z,y-y_n)\|_1 + \|G_1(z,y_n) - G_1(z_0,y_n)\|_1 + \|G_1(z_0,y_n-y)\|_1 \\ &\quad\leq \lim_{z\to z_0} \|G_1(z,y_n) - G_1(z_0,y_n)\|_1 + 2\|y-y_n\|_{W^{k_1,1}} \\ &\quad\leq 2\|y-y_n\|_{W^{k_1,1}} , \end{aligned}$$

where we used Equation (96). Taking the limit $n \to \infty$ concludes the claim that $G_1(\cdot,y) : S \to \mathcal{T}_{1,\mathrm{sa}}$ is continuous for all $y \in W^{k_1,1}$. We further have that G_2 as a map from S to $W^{k_1,1}$ is continuous, since $x \in \mathcal{T}_f$ and the maps $z \mapsto (n+1)^{\frac{k(z)}{4}}$ for $n \in \mathbb{N}$ are continuous as maps $S \to \mathbb{C}$. This suffices since $x \in \mathcal{T}_f$, hence all involved operators can be made finite dimensional via a cut-off in the Fock-basis again.

We can now write

$$G(z) = G_1(z, T(G_2(z)))$$

where $T(G_2(z)) \in W^{k_1,1}$ as $T : W^{k_1,1} \to W^{k_1,1}$ and $G_2(z) \in \mathcal{T}_f \subset W^{k_1,1}$ for all $z \in S$. This not only gives us that G is well-defined but also allows us to get

$$\begin{aligned}\|G(z)\|_1 &= \|G_1(z, T(G_2(z)))\|_1 \\ &\leq \|T(G_2(z))\|_{W^{k_1,1}} \\ &\leq \|T\|_{W^{k_1,1}\to W^{k_1,1}} \|G_2(z)\|_{W^{k_1,1}} \\ &\leq \|T\|_{W^{k_1,1}\to W^{k_1,1}} \left\|(N+\mathbb{1})^{\frac{k_\theta-k_0}{4}} x (N+\mathbb{1})^{\frac{k_\theta-k_0}{4}}\right\|_{W^{k_1,1}}\end{aligned}$$

where we again used Equation (96), giving us a bound independent of z. Further, using again Equation (96) we can conclude continuity, since

$$\begin{aligned}\lim_{z\to z_0} \|G(z) - G(z_0)\|_1 &\leq \lim_{z\to z_0} \|G_1(z, T\{G_2(z) - G_2(z_0)\})\|_1 \\ &\qquad + \lim_{z\to z_0} \|G_1(z, T(G_2(z_0))) - G_1(z_0, T(G_2(z_0)))\|_1 \\ &\leq \lim_{z\to z_0} \|T\|_{W^{k_1,1}\to W^{k_1,1}} \|G_2(z) - G_2(z_0)\|_{W^{k_1,1}} \\ &\qquad + \lim_{z\to z_0} \|G_1(z, T(G_2(z_0))) - G_1(z_0, T(G_2(z_0)))\|_1 \\ &= 0\end{aligned}$$

where in addition we used the continuity of $G_1(\cdot, y) : S \to \mathcal{T}_{1,\text{sa}}$ for $y \in W^{k_1,1}$ and $G_2 : S \mapsto W^{k_1,1}$. □

Lemma D.2 (Differentiability of G(z)) *Let $k_0 < k_1 \in \mathbb{R}_+$, $T : W^{k_j,1} \to W^{k_j,1}$, be a linear map with $\|T\|_{W^{k_j,1}\to W^{k_j,1}} \leq M_j$, bounded by $M_j \geq 0$ for $j = 1, 2$ respectively. Further let $\theta \in [0, 1]$ and $k_\theta = (1-\theta)k_0 + \theta k_1$ and $x \in \mathcal{T}_f$, then the map*

$$\begin{aligned}G : S &:= \{z \in \mathbb{C} \; : \; 0 \leq \mathrm{Re}(z) \leq 1\} \to \mathcal{T}_{1,\text{sa}} \\ z &\mapsto G(z) = (N+\mathbb{1})^{\frac{k(z)}{4}} T\left((N+\mathbb{1})^{\frac{k_\theta-k(z)}{4}} x (N+\mathbb{1})^{\frac{k_\theta-k(z)}{4}}\right)(N+\mathbb{1})^{\frac{k(z)}{4}}\end{aligned}$$

with $k(z) = (1-z)k_0 + zk_1$, is holomorphic on $\mathring{S} := \{z \in \mathbb{C} \; : \; 0 < \mathrm{Re}(z) < 1\}$.

Proof. To prove the claim, we follow a similar strategy as with Lemma D.1. We will again use the auxiliary functions Equation (94) and Equation (95). We begin by showing that for a fixed $y \in W^{k,1}$, $G_1(\cdot, y) : S \to \mathcal{T}_{1,\text{sa}}$ is holomorphic on $\mathring{S}$ and initially even simplify to the case $y \in \mathcal{T}_f$. In this setting, all operators involved can be assumed to be linear maps on a finite subspace, by just taking a cut-off in the Fock-basis as we did before. This allows us to Taylor expand around $z_0 \in \mathring{S}$

$$(N+\mathbb{1})^{\frac{k(z)}{4}} y (N+\mathbb{1})^{\frac{k(z)}{4}} = G_1(z_0, y) + G_1'(z_0, y)(z - z_0) + \int_{[z_0,z]} G_1''(\omega, y)(\omega - z_0)\, d\omega$$

where the integral is a path integral along the line segment $[z_0, z]$ and

$$\begin{aligned}G_1'(z_0, y) = \frac{k_0 - k_1}{4} \Big(&\log(N+\mathbb{1})(N+\mathbb{1})^{\frac{k(z_0)}{4}} y (N+\mathbb{1})^{\frac{k(z_0)}{4}} \\ &+ (N+\mathbb{1})^{\frac{k(z_0)}{4}} y (N+\mathbb{1})^{\frac{k(z_0)}{4}} \log(N+\mathbb{1})\Big)\end{aligned}$$

and

$$G_1''(\omega, y) = \left(\frac{k_0 - k_1}{4}\right)^2 \Big(\log^2(N+\mathbb{1})(N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}} + 2\log(N+\mathbb{1})(N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}} \log(N+\mathbb{1}) + (N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}} \log^2(N+\mathbb{1})\Big)$$

are linear in y. From this representation, we can immediately deduce holomorphy of $G_1(\cdot, y) : S \to \mathcal{T}_{1,\mathrm{sa}}$ at $z_0 \in \mathring{S}$ and hence on all of $\mathring{S}$. To lift holomorphy from $y \in \mathcal{T}_f$ to $y \in W^{k_1,1}$, we note that for $z_0 \in \mathring{S}$ there exists $C_{z_0} \geq 0$ such that for $y \in \mathcal{T}_f$

$$\left\|G'(z_0, y)\right\|_1 \leq C_{z_0} \|y\|_{W^{k_1,1}} \tag{97}$$

and further for $\omega \in B_\varepsilon(z_0) \coloneqq \{z \in \mathbb{C} \ : \ |z - z_0| < \varepsilon\} \subset \mathring{S}$ there exists $C_{\varepsilon,z_0} \geq 0$ such that

$$\left\|G''(\omega, y)\right\|_1 \leq C_{\varepsilon,z_0} \|y\|_{W^{k_1,1}} \,. \tag{98}$$

We will only show that given ω as above,

$$\left\|\log^2(N+\mathbb{1})(N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}}\right\|_1 \leq \tilde{C}_{\varepsilon,z_0} \|y\|_{W^{k_1,1}} \,. \tag{99}$$

Using the same reasoning for the other terms of Equation (97) and Equation (98) in combination with triangle inequality immediately gives the claims. Note first that we can reduce $k(\omega)$ to its real part since the imaginary part only produces a unitary $(N+\mathbb{1})^{i\,\mathrm{Im}(k(\omega))}$ that can be absorbed into the norm. We call the real part $r(\omega)$ for now. Since $\omega \in B_\varepsilon(z_0) \subset \mathring{S}$ we find a $\delta_\varepsilon > 0$ independent of ω, such that $|r(\omega) - k_1| < \delta_\varepsilon$ or more precisely $r(\omega) - k_1 \leq -\delta_\varepsilon$. Hence using Hölder's inequality, we can deduce

$$\begin{aligned}
&\left\|\log^2(N+\mathbb{1})(N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}} y (N+\mathbb{1})^{\frac{k(\omega)}{4}}\right\|_1 \\
&\leq \left\|\log^2(N+\mathbb{1})(N+\mathbb{1})^{\frac{r(\omega)-k_1}{4}}\right\|_\infty \left\|(N+\mathbb{1})^{\frac{r(\omega)-k_1}{4}}\right\|_\infty \|y\|_{W^{k_1,1}} \\
&\leq \left\|\log^2(N+\mathbb{1})(N+\mathbb{1})^{-\frac{\delta_\varepsilon}{4}}\right\|_\infty \left\|(N+\mathbb{1})^{-\frac{\delta_\varepsilon}{4}}\right\|_\infty \|y\|_{W^{k_1,1}} \\
&\leq \left\|\log^2(N+\mathbb{1})(N+\mathbb{1})^{-\frac{\delta_\varepsilon}{4}}\right\|_\infty \|y\|_{W^{k_1,1}}
\end{aligned}$$

where we used that $x \mapsto e^{kx}$ for $k \geq 0$ is monotone and further that $(N+\mathbb{1})^{-\frac{\delta_\varepsilon}{4}}$ is a contraction. Lastly, we have that $x \mapsto \frac{\log^2(x+1)}{(x+1)^{\frac{\delta_\varepsilon}{4}}}$ is a bounded function for $x \geq 0$ with a bound we call $\tilde{C}_{\delta_\varepsilon}$. This allows us to estimate $\left\|\log^2(N+\mathbb{1})(N+\mathbb{1})^{-\frac{\delta_\varepsilon}{4}}\right\|_\infty \leq \tilde{C}_{\delta_\varepsilon}$, which concludes Equation (99) and therefore also Equation (97) and Equation (98).

For a general $y \in W^{k_1,1}$ and $z_0 \in \mathring{S}$ Equation (97) allows us to conclude that $G_1'(z_0, y) \in \mathcal{T}_{1,\mathrm{sa}}$ is well defined. Further, for $z \in B_\varepsilon(z_0)$ and $(y_n)_{n\in\mathbb{N}} \subset \mathcal{T}_f$ with $y_n \to y$ in $W^{k_1,1}$, we have for all $n \in \mathbb{N}$

$$\begin{aligned}
\left\|\frac{G_1(z, y_n) - G_1(z_0, y_n)}{z - z_0} - G_1'(z_0, y_n)\right\|_1 &\leq \frac{1}{|z - z_0|} \int\limits_{[z_0,z]} \|G_1''(\omega, y_n)\|_1 |(\omega - z_0) d\omega| \\
&\leq C_{\varepsilon,z_0} |z - z_0| \|y_n\|_{W^{k_1,1}}
\end{aligned} \tag{100}$$

where we used the expansion and Equation (98). Now we can take the limit $n \to \infty$ on both sides, as all objects involved are stable w.r.t. that limit (using Lemma D.1 and Equation (97)). We get

$$\left\| \frac{G_1(z,y) - G_1(z_0,y)}{z - z_0} - G_1'(z_0,y) \right\|_1 \leq C_{\varepsilon,z_0} |z - z_0| \|y\|_{W^{k_1,1}} \tag{101}$$

which immediately lets us deduce holomorphy of $G_1(\cdot,y) : S \to \mathcal{T}_{1,\mathrm{sa}}$ on $\mathring{S}$ for $y \in W^{k_1,1}$.

For $G_2 : S \to W^{k_1,1}$ the holomorphy immediately follows from the fact that $x \in \mathcal{T}_f$, which again allows reducing the analysis to a finite-dimensional subspace by taking a cut-off in the Fock basis again. Lastly, we have that $T(G_2(z)) \in W^{k_1,1}$ for all $z \in S$, which finally gives us that for $z_0 \in \mathring{S}$ and for $z \in B_\varepsilon(z_0) \subset \mathring{S}$

$$\begin{aligned}
&\left\| \frac{G(z) - G(z_0)}{z - z_0} - (G_1'(z_0, T(G_2(z_0))) + G_1(z_0, T\{G_2'(z_0)\}) \right\| \\
&\leq \left\| \frac{G_1(z, T(G_2(z_0))) - G_1(z_0, T(G_2(z_0)))}{z - z_0} - G_1'(z_0, T(G_2(z_0))) \right\|_1 \\
&\quad + \|T\|_{W^{k_1,1} \to W^{k_1,1}} \left\| \frac{G_2(z) - G_2(z_0)}{z - z_0} - G_2'(z_0) \right\|_{W^{k_1,1}}
\end{aligned}$$

where we used linearity of $G_1(z,\cdot)$, $G_1'(z,\cdot)$ and T. In addition, we used the bound on $G_1(z,\cdot)$ from Equation (96) and G_2' to denote the derivative of G_2. Now the differentiability of $G_1(\cdot,y)$ and G_2 at z_0 immediately gives the differentiability of G at z_0, which concludes the proof as $z_0 \in \mathring{S}$ was arbitrary. □

E Technical lemmas for the generation theorem

Lemma E.1 *For $d \geq 0$ and $\varepsilon > 0$, define the operator*

$$\mathcal{I}_{d,\varepsilon} : \mathcal{T}_f \to \mathcal{T}_f, \qquad x \mapsto \mathcal{I}_{d,\varepsilon}(x) := -\varepsilon\{(N + \mathbb{1})^{4d}, x\}\,.$$

For $\lambda \geq 0$, we have that $\lambda - \mathcal{I}_{d,\varepsilon} : \mathcal{T}_f \to \mathcal{T}_f$ is bijective. While for all $k \in \mathbb{R}_+$ and $x \in \mathcal{T}_f$ one further has

$$\left\| (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x) \right\|_{W^{k,1}} \leq \frac{1}{\lambda + 2\varepsilon} \|x\|_{W^{k,1}}\,; \tag{1}$$

$$\left\| \mathcal{I}_{d,\varepsilon} \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x) \right\|_{W^{k,1}} \leq 2\|x\|_{W^{k,1}}\,. \tag{2}$$

Proof. For $\lambda \geq 0$ define the following linear operator

$$\begin{aligned}
&(\lambda - \mathcal{I}_{d,\varepsilon})^{-1} : \mathcal{T}_f \to \mathcal{T}_f, \\
x = \sum_{\text{finite}} x_{nm} \, |n\rangle\!\langle m| &\mapsto (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x) := \sum_{\text{finite}} x_{nm} \frac{1}{\varepsilon(n+1)^{4d} + \varepsilon(m+1)^{4d} + \lambda} |n\rangle\!\langle m|\,.
\end{aligned}$$

or alternatively

$$\begin{aligned}
(\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x) &= \int_0^\infty e^{-(\varepsilon(N+\mathbb{1})^{4d} + \lambda/2)s} x e^{-(\varepsilon(N+\mathbb{1})^{4d} + \lambda/2)s} ds \\
&= \int_0^\infty \sum_{\text{finite}} e^{-(\varepsilon(n+\mathbb{1})^{4d} + \lambda/2)s} x_{nm} e^{-(\varepsilon(m+\mathbb{1})^{4d} + \lambda/2)s} \, |n\rangle\!\langle m| \, ds\,.
\end{aligned}$$

The integral representation allows us to deduce that $(\lambda - \mathcal{I}_{d,\varepsilon})^{-1}$ preserves positivity. Using the first expression it is straightforward to show that this map is indeed the inverse to $\lambda - \mathcal{I}_{d,\varepsilon} : \mathcal{T}_f \to \mathcal{T}_f$. The bound $\|(\lambda - \mathcal{I}_{d,\varepsilon})^{-1}x\|_{W^{k,1}} \leq \frac{1}{\lambda+2\varepsilon}\|x\|_{W^{k,1}}$ can be shown, using the integral representation and Hölder inequality:

$$\begin{aligned}\left\|(\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x)\right\|_{W^{k,1}} &\leq \int_0^\infty \left\|e^{-\varepsilon(N+\mathbb{1})^{4d}+\lambda/2)s}(N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4}e^{-\varepsilon(N+\mathbb{1})^{4d}+\lambda/2)s}\right\|_1 \\ &\leq \int_0^\infty \left\|e^{-\varepsilon(N+\mathbb{1})^{4d}+\lambda/2)s}\right\|_\infty^2 ds\|x\|_{W^{k,1}} \\ &= \int_0^\infty e^{-(2\varepsilon+\lambda)s}\|x\|_{W^{k,1}} = \frac{1}{\lambda+2\varepsilon}\|x\|_{W^{k,1}}\,.\end{aligned}$$

Issues arising from the unbounded nature of N can be ignored in the above estimations, as we can take a finite cut-off in the Fock basis due to $x \in \mathcal{T}_f$. For (2), we have that on $\mathcal{T}_f$, $-\mathcal{I}_{d,\varepsilon} \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1} = \mathbb{1} - \lambda(\lambda - \mathcal{I}_{d,\varepsilon})^{-1}$, i.e. for $x \in \mathcal{T}_f$

$$\left\|(\mathbb{1} - \lambda(\lambda - \mathcal{I}_{d,\varepsilon})^{-1})x\right\|_{W^{k,1}} = \left\|-\mathcal{I}_{d,\varepsilon} \circ (\lambda - \mathcal{I}_{d,\varepsilon})^{-1}(x)\right\|_{W^{k,1}} \tag{102}$$

where the LHS can be upper bounded by $(1 + \frac{\lambda}{\lambda+2\varepsilon})\|x\|_{W^{k,1}} \leq 2\|x\|_{W^{k,1}}$ using (1). This proves the last claim. □

Lemma E.2 *For $p \in \mathbb{C}[X,Y]$ a polynomial of degree d and*

$$A : \mathcal{H}_f \to \mathcal{H}, \qquad |\psi\rangle = \sum_{\textit{finite}} \psi_n |n\rangle \mapsto p(a, a^\dagger)|\psi\rangle = \sum_{\textit{finite}} \psi_n p(a, a^\dagger)|n\rangle\,,$$

we get that for all $k \in \mathbb{R}_+$ and $d' \geq d$

$$\begin{aligned}B_1 : \mathcal{H}_f \to \mathcal{H}, \qquad |\psi\rangle \mapsto (N+\mathbb{1})^k A (N+\mathbb{1})^{-k-d'}|\psi\rangle \\ B_2 : \mathcal{H}_f \to \mathcal{H}, \qquad |\psi\rangle \mapsto (N+\mathbb{1})^{-k-d'} A (N+\mathbb{1})^{k}|\psi\rangle\end{aligned}$$

are bounded and therefore can be uniquely extended to a bounded map on $\mathcal{H}$, with the same bound.

Proof. Since the proof for B_1 and B_2 are almost completely analogous, we will only show it here for B_1. The canonical commutation relation allows us to rewrite A as a finite linear combination of monomials of the form $(a^\dagger)^i N^j$ and $a^i N^j$ with $i + j/2 \leq d$. Now by triangle inequality for the norm on $\mathcal{H}$ and since the sum of these monomials comprising A are finite, for the claim to be true it suffices to show that

$$(N+\mathbb{1})^k (a^\dagger)^i N^j (N+\mathbb{1})^{-k-d} : \mathcal{H}_f \to \mathcal{H}, \qquad (N+\mathbb{1})^k a^i N^j (N+\mathbb{1})^{-k-d} : \mathcal{H}_f \to \mathcal{H}$$

are bounded, and hence can be uniquely extended to a bounded map on $\mathcal{H}$. We only give the argument for the first map, since it is almost completely analogous to the second one. Let $|\psi\rangle = \sum_{n=0}^{M} \psi_n |n\rangle \in \mathcal{H}_f$, then

$$\begin{aligned}|\varphi\rangle \coloneqq (N+\mathbb{1})^k (a^\dagger)^i N^j (N+\mathbb{1})^{-k-d'}|\psi\rangle &= \sum_{n=0}^{M} \psi_n (N+\mathbb{1})^k (a^\dagger)^i N^j (N+\mathbb{1})^{-k-d'}|n\rangle \\ &= \sum_{n=0}^{M} \psi_n \frac{(n+1+j)^k}{(n+1)^k} \frac{n^j \prod_{l=1}^{i}\sqrt{n+l}}{(n+1)^{d'}}|n+i\rangle\,.\end{aligned}$$

Hence

$$\begin{aligned}\|\varphi\|^2 &= \sum_{n=0}^{M} \frac{(n+1+j)^{2k}}{(n+1)^{2k}} \frac{n^{2j}\prod_{l=1}^{i}(n+l)}{(n+1)^{2d'}}|\psi_n|^2 \\ &\le \sum_{n=0}^{M} \frac{(n+1+d)^{2k}}{(n+1)^{2k}} \frac{(n+d)^{2j}(n+d)^{i}}{(n+1)^{2d}}|\psi_n|^2 \\ &\le \sum_{n=0}^{M} d^{2k}d^{2d}|\psi_n|^2 \\ &= d^{2(k+d)}\|\psi\|^2\,,\end{aligned}$$

where we used that $i + j/2 \le d$ and $d \le d'$. Hence $(N+\mathbb{1})^k(a^\dagger)^i N^j (N+\mathbb{1})^{-k-d'} : \mathcal{H}_f \to \mathcal{H}$ is bounded by d^{k+d} and can be uniquely extended to a bounded linear map on $\mathcal{H}$. This concludes the claim. □

Lemma E.3 *Let $K \in \mathbb{N}$. For $i = 1, \dots, K$ let $p_{i,1}, p_{i,2} \in \mathbb{C}[X,Y]$ polynomials of degree $d_{i,1}$, $d_{i,2}$ such that*

$$\mathcal{A} : \mathcal{T}_f \to \mathcal{T}_f, \qquad x \mapsto \mathcal{A}(x) = \sum_{i=1}^{K} A_{i,1} x A_{i,2} = \sum_{i=1}^{K} p_{i,1}(a,a^\dagger)x\, p_{i,2}(a,a^\dagger)$$

where the action of $p_{1/2}(a,a^\dagger)$ on x is defined via the action of a and $a^\dagger$ on $|n\rangle\langle m|$. We then have that for all $k \ge 0$, $d \ge \max_{i=1,\dots,K} \max\{d_{i,1}, d_{i,2}\}$ there exists $C_k \ge 0$, s.t. for all $\varepsilon \ge 0$ and $\forall x \in \mathcal{T}_f$

$$\|\mathcal{A}(x)\|_{W^{k,1}} \le \varepsilon \left\| \{(N+\mathbb{1})^{4d}, x\} \right\|_{W^{k,1}} + \frac{C_k}{\varepsilon}\|x\|_{W^{k,1}}\,.$$

Proof. Let $k \in \mathbb{R}_+$. The first step is to show that there exists $c_k \ge 0$, s.t. for all $x \in \mathcal{T}_f$

$$\|\mathcal{A}(x)\|_{W^{k,1}} \le c_k \left\|(N+\mathbb{1})^d x (N+\mathbb{1})^d\right\|_{W^{k,1}}\,. \tag{103}$$

The argument reduces to showing that for $i = 1, \dots, K$ and $x \in \mathcal{T}_f$

$$\left\|(N+\mathbb{1})^{k/4} A_{i,1} x A_{i,2} (N+\mathbb{1})^{k/4}\right\|_1 \le c_{i,k} \left\|(N+\mathbb{1})^d x (N+\mathbb{1})^d\right\|_{W^{k,1}}$$

since the sum comprising $\mathcal{A}$ is finite. Note that the trace norm on the LHS is the one on the trace-class operators since the argument might not necessarily be self-adjoint. For $x \in \mathcal{T}_f$, we have

$$\begin{aligned}&\left\|(N+\mathbb{1})^{k/4} A_{i,1} x A_{i,2} (N+\mathbb{1})^{k/4}\right\|_1 \\ &\quad = \|(N+\mathbb{1})^{k/4} A_{i,1} (N+\mathbb{1})^{-k/4-d} (N+\mathbb{1})^{k/4+d} x \\ &\qquad\qquad (N+\mathbb{1})^{k/4+d} (N+\mathbb{1})^{-k/4-d} A_{i,2} (N+\mathbb{1})^{k/4}\|_1 \\ &\quad \le \left\|(N+\mathbb{1})^{k/4} A_{i,1} (N+\mathbb{1})^{-k/4-d}\right\|_\infty \|x\|_{W^{k+4d,1}} \left\|(N+\mathbb{1})^{-k/4-d} A_{i,2} (N+\mathbb{1})^{k/4}\right\|_\infty \\ &\quad \le c_{i,k} \left\|(N+\mathbb{1})^d x (N+\mathbb{1})^d\right\|_{W^{k,1}}\end{aligned}$$

where we used Lemma E.2 to argue that the operators involved are bounded and we can employ Hölder's inequality to split them off. Subsequently, we replaced the operator norms with the constant $c_{i,k}$. Now in the second step we show that for all $\varepsilon > 0$ and $x \in \mathcal{T}_f$

$$\left\|(N+\mathbb{1})^d x (N+\mathbb{1})^d\right\|_{W^{k,1}} \le \varepsilon \left\|\{(N+\mathbb{1})^{4d}, x\}\right\|_{W^{k,1}} + \frac{1}{4\varepsilon}\|x\|_{W^{k,1}}\,.$$

Combining this with Equation (103) then immediately provides the claim. Therefore, let $\lambda > 0$ and $x \in \mathcal{T}_f$ with $x \geq 0$. We find

$$\begin{aligned}\Big\|(N+1)^d(\lambda-\mathcal{I}_{d,1})^{-1}(x)(N+\mathbb{1})^d\Big\|_{W^{k,1}} &= \operatorname{tr}\Big[(\lambda-\mathcal{I}_{d,1})^{-1}\{(N+\mathbb{1})^{k/4+d}x(N+\mathbb{1})^{k/4+d}\}\Big]\\ &= \int_0^\infty \operatorname{tr}\Big[e^{-(2(N+\mathbb{1})^{4d}+\lambda)s}(N+\mathbb{1})^{2d}(N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4}\Big]ds\\ &= \operatorname{tr}\Big[\int_0^\infty e^{-(2(N+\mathbb{1})^{4d}+\lambda)s}(N+\mathbb{1})^{2d}ds\,(N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4}\Big]\\ &= \operatorname{tr}\left[\frac{(N+\mathbb{1})^{2d}}{(N+\mathbb{1})^{4d}+\lambda}(N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4}\right],\end{aligned}$$

where we used the map $\mathcal{I}_{d,1}$ from Lemma E.1, the integral representation of its resolvent $(\lambda-\mathcal{I}_{d,1})^{-1}$ and that this resolvent preserves positivity. Further, we applied the cyclicity of the trace and conveniently suppressed issues that might arise from the unbounded nature of N by taking a cut-off in the Fock basis. This is justified by $x \in \mathcal{T}_f$. Now we can use that $(N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4} \geq 0$ to bound the RHS of the above chain of inequalities to get

$$\begin{aligned}\Big\|(N+1)^d(\lambda-\mathcal{I}_{d,1})^{-1}(x)(N+\mathbb{1})^d\Big\|_{W^{k,1}} &\leq \sup_{s\geq 1}\frac{s}{s^2+\lambda}\operatorname{tr}\Big[(N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4}\Big]\\ &= \sup_{s\geq 1}\frac{s}{s^2+\lambda}\|x\|_{W^{k,1}}\,.\end{aligned}$$

For a general $x \in \mathcal{T}_f$, we can set $y = (N+\mathbb{1})^{k/4}x(N+\mathbb{1})^{k/4}$ decompose into $y = y_+ - y_-$ the positive and negative part of y respectively and then set $x_\pm = (N+\mathbb{1})^{-k/4}y_\pm(N+\mathbb{1})^{-k/4}$. We clearly have that $x_\pm \in \mathcal{T}_f$, $x = x_+ - x_-$ and $x_\pm \geq 0$ as $(N+\mathbb{1})^{-k/4}\cdot(N+\mathbb{1})^{-k/4}$ preserves positivity. This allows us to apply what we have shown above to obtain

$$\begin{aligned}\|(N+1)^d(\lambda-\mathcal{I}_{d,1})^{-1}(x)(N+\mathbb{1})^d\|_{W^{k,1}} &\leq \Big\|(N+1)^d(\lambda-\mathcal{I}_{d,1})^{-1}(x_+)(N+\mathbb{1})^d\Big\|_{W^{k,1}}\\ &\quad + \Big\|(N+1)^d(\lambda-\mathcal{I}_{d,1})^{-1}(x_-)(N+\mathbb{1})^d\Big\|_{W^{k,1}}\\ &\leq \sup_{s\geq 1}\frac{s}{s^2+\lambda}(\|x_+\|_{W^{k,1}}+\|x_-\|_{W^{k,1}})\\ &= \sup_{s\geq 1}\frac{s}{s^2+\lambda}(\|y_+\|_1+\|y_-\|_1)\\ &= \sup_{s\geq 1}\frac{s}{s^2+\lambda}\|y\|_1\\ &= \sup_{s\geq 1}\frac{s}{s^2+\lambda}\|x\|_{W^{k,1}}\end{aligned}$$

Lastly we can use the bijectivity of $(\lambda-\mathcal{I}_{d,1})$ on $\mathcal{T}_f$ (q.v. Lemma E.1) and triangle inequality to conclude

$$\Big\|(N+\mathbb{1})^d x(N+\mathbb{1})^d\Big\|_{W^{k,1}} \leq \sup_{s\geq 0}\frac{s}{s^2+\lambda}\|\mathcal{I}_{d,1}(x)\|_{W^{k,1}} + \lambda\sup_{s\geq 0}\frac{s}{s^2+\lambda}\|x\|_{W^{k,1}}\,. \tag{104}$$

Choosing $\lambda = \frac{1}{4\varepsilon^2}$, we find that $\sup_{s\geq 1}\frac{s}{s^2+\lambda} < \varepsilon$, hence

$$\Big\|(N+\mathbb{1})^d x(N+\mathbb{1})^d\Big\|_{W^{k,1}} \leq \varepsilon\Big\|\{(N+\mathbb{1})^{4d},x\}\Big\|_{W^{k,1}} + \frac{1}{4\varepsilon}\|x\|_{W^{k,1}} \tag{105}$$

□

Lemma E.4 (Interpolation Lemma) *Let $k_0 < k_1 \in \mathbb{R}_+$ and $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ an operator on $W^{k_j,1}$, $j = 0, 1$ respectively. Further, assume that the closure of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ defines a strongly continuous semigroup $(\mathcal{P}_t^j)_{t\geq 0}$ with*

$$\left\|\mathcal{P}_t^j\right\|_{W^{k_j,1}\to W^{k_j,1}} \leq M_j e^{\omega_j t} \quad \forall t \geq 0 \tag{106}$$

in both spaces, respectively. Then for $\theta \in [0, 1]$ and $k_\theta = \theta k_1 + (1-\theta)k_0$ the following are true

1. *The closure of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ defines a strongly continuous semigroup on $W^{k_\theta,1}$ with*

$$\left\|\mathcal{P}_t^\theta\right\|_{W^{k_\theta,1}\to W^{k_\theta,1}} \leq M_0^{1-\theta} M_1^\theta e^{(\omega_{k_0}(1-\theta)+\omega_{k_1}\theta)t} \quad \forall t \geq 0\,. \tag{107}$$

2. *$(\mathcal{P}_t^\theta)_{t\geq 0}$ agrees with $(\mathcal{P}_t^j)_{t\geq 0}$ on $W^{k_j,1} \cap W^{k_\theta,1}$ for $j = 1, 2$.*

Proof. We begin with the second claim and only cover $k_j = k_0$ as the other case only requires minor changes that are left to the reader. Let $\theta \in (0, 1)$ and the closure of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ the generator of $(\mathcal{P}_t^0)_{t\geq 0}$ and $(\mathcal{P}_t^\theta)_{t\geq 0}$ on the respective spaces. Since $k_0 < k_\theta$ and hence $W^{k_\theta,1} \Subset W^{k_0,1}$, we have that the closure of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ in $W^{k_\theta,1}$ agrees with the restriction of the closure in $W^{k_0,1}$. As the semigroup is completely determined by its generator, we find that the semigroups agree on $W^{k_\theta,1}$.
For the first claim, note that the semigroups $(\mathcal{P}_t^j)_{t\geq 0}$, $j = 1, 2$ agree on $W^{k_1,1}$ by Item 2, which allows us to employ the Stein-Weiss theorem for Bosonic Sobolev spaces (Theorem 2.14) to conclude Equation (107). It remains to check that the families of bounded maps $(\mathcal{P}_t^\theta)_{t\geq 0}$ are strongly continuous semigroups generated by the closure of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ on $W^{k_\theta,1}$. We have that $\mathcal{P}_0^\theta = \mathbb{1}$ and $\mathcal{P}_t^\theta \mathcal{P}_s^\theta = \mathcal{P}_{t+1}^\theta$ $\forall t, s \geq 0$ as a consequence of $\mathcal{P}_t^0|_{W^{k_\theta,1}} = \mathcal{P}_t^\theta$ $\forall t$ (this equality holds by Theorem 2.14). The strong continuity follows, due to $\mathcal{P}_t^1 = \mathcal{P}_t^\theta|_{W^{k_1,1}}$, as for $x \in W^{k_1,1}$

$$\lim_{t\to 0}\left\|\mathcal{P}_t^\theta(x) - x\right\|_{W^{k_\theta,1}} = \lim_{t\to 0}\left\|\mathcal{P}_t^1(x) - x\right\|_{W^{k_\theta,1}} \leq \lim_{t\to 0}\left\|\mathcal{P}_t^1(x) - x\right\|_{W^{k_1,1}} = 0$$

where we used $W^{k_1,1} \Subset W^{k_\theta,1}$ and that $\mathcal{P}_t^1$ is a strongly continous semigroup on $W^{k_1,1}$. For general $x \in W^{k_\theta,1}$, we find $(x_n)_{n\in\mathbb{N}} \subset W^{k_\theta,1}$ converging to x in $W^{k_\theta,1}$ and for all $n \in \mathbb{N}$

$$\begin{aligned}\lim_{t\to 0}\left\|\mathcal{P}_t^\theta(x) - x\right\|_{W^{k_\theta,1}} &\leq \lim_{t\to 0}\left[(1 + M_0^{1-\theta} M_1^\theta e^{(\omega_{k_0}(1-\theta)+\omega_{k_1}\theta)t})\|x - x_n\|_{W^{k_\theta,1}} + \left\|\mathcal{P}_t^\theta(x_n) - x_n\right\|\right]\\ &\leq (1 + M_0^{1-\theta} M_1^\theta)\|x - x_n\|_{W^{k_\theta,1}}\end{aligned}$$

which concludes the strong continuity. It remains to argue that the closure of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ on $W^{k_\theta,1}$ is indeed the generator of $(\mathcal{P}_t^\theta)_{t\geq 0}$. By [23, Sec. II.2.3], we find that the restriction of the generator $(\hat{\mathcal{L}}, \mathcal{D}(\hat{\mathcal{L}}))$ of $(\mathcal{P}_t^\theta)_{t\geq 0}$ to $W^{k_1,1}$ is the generator of $(\mathcal{P}_t^1)_{t\geq 0}$ with $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ being a core for this restricted generator on $W^{k_1,1}$ by assumption. This in particular means that for $\lambda > \omega_k$, $\lambda - \mathcal{L} : \mathcal{D}(\mathcal{L}) \to W^{k_1,1}$ has a dense range in $W^{k_1,1}$, which further allows us to conclude that $\lambda - \mathcal{L} : \mathcal{D}(\mathcal{L}) \to W^{k_\theta,1}$ has a dense range in $W^{k_\theta,1}$. Now as it is ω_{k_θ}-quasi dissipative (being the restriction of the generator of the semigroup $(\mathcal{P}_t^\theta)_{t\geq 0}$) we can conclude that indeed $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is closable in $W^{k_\theta,1}$ with the closure $(\hat{\mathcal{L}}, \mathcal{D}(\hat{\mathcal{L}}))$ (c.f. [23, Proposition 3.14]). □

Lemma E.5 (Approximation Lemma) *Let $K \in \mathbb{N}$. For $i = 1, \ldots, K$ let $p_{i,1}, p_{i,2} \in \mathbb{C}[X,Y]$ polynomials of degree $d_{i,1}$, $d_{i,2}$ and $\{a_{i,n}\}_{n\in\mathbb{N}} \subset \mathbb{C}$ convergent sequences with limits $a_i \in \mathbb{C}$ such that $\{(\mathcal{A}_n, \mathcal{D}(\mathcal{A}_n) = \mathcal{T}_f)\}_{n\in\mathbb{N}}$ is an operator sequence with*

$$\mathcal{A}_n : \mathcal{T}_f \to \mathcal{T}_f, \quad x \mapsto \mathcal{A}_n(x) \coloneqq \sum_{i=1}^{K} a_{i,n} A_{i,1}\, x\, A_{i,2} \coloneqq \sum_{i=1}^{K} a_{i,n} p_{i,1}(a, a^\dagger)\, x\, p_{i,2}(a, a^\dagger) \tag{108}$$

with $A_{i,1} = p_{i,1}(a, a^\dagger)$ and $A_{i,2} = p_{i,2}(a, a^\dagger)$ (see Lemma E.3). If for all $k \in \mathbb{R}_+$ there exists M_k, ω_k such that for all $n \in \mathbb{N}$ the closure of $(\mathcal{A}_n, \mathcal{D}(\mathcal{A}_n))$ generates a strongly continuous semigroup $(\mathcal{P}_t^n)_{t\geq 0}$ on $W^{k,1}$ with

$$\|\mathcal{P}_t^n\|_{W^{k,1}\to W^{k,1}} \leq M_k e^{\omega_k t} \quad \forall t \in \mathbb{R}\,, \tag{109}$$

then the closure of $(\mathcal{A}, \mathcal{D}(\mathcal{A}) = \mathcal{T}_f)$, the pointwise limit of $(\mathcal{A}_n, \mathcal{D}(\mathcal{A}_n))$, defines a strongly continuous semigroup on $W^{k,1}$ for $k \geq 0$ as well. We further get that the semigroups generated by the closure of $(\mathcal{A}_n, \mathcal{D}(\mathcal{A}_n))$ converge uniformly (in time) on compact intervals to the semigroup generated by the closure of $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ and that Equation (109) also holds for the limiting semigroup.

Proof. Let $k \in \mathbb{R}_+$. To prove the lemma, we first note that $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ is densely defined and the pointwise limit of $\{(\mathcal{A}_n, \mathcal{D}(\mathcal{A}_n))\}_{n\in\mathbb{N}}$. To employ the second Trotter-Kato approximation theorem, which implies the claim (see the version in [23, Thm. III.4.9]), we need to show that there exists $\lambda > 0$ such that $(\lambda - \mathcal{A}, \mathcal{D}(\mathcal{A}))$ has dense range in $W^{k,1}$. We will do so by showing that the closure of the range contains $\mathcal{T}_f$ which is a dense subset of $W^{k,1}$. Therefore let $\lambda > \max\{\omega_k, \omega_{k+4d}\}$ (with $\omega_.$ from Equation (109) and d the maximal degree of the polynomials but at least one, i.e. $d = \max_{i=1,\ldots,K} \max\{d_{i,1}, d_{i,2}, 1\}$). By assumption, we have that for all $n \in \mathbb{N}$ the operator $(\lambda - \mathcal{A}_n, \mathcal{D}(\mathcal{A}_n))$ has dense range in $W^{k+4d,1}$, meaning in particular that for any $\xi \in \mathcal{T}_f$ we find a sequence $\{x_{n,m}\}_{m\in\mathbb{N}}$ which is convergent in $W^{k+4d,1}$ and further

$$\lim_{m\to\infty} \|(\lambda - \mathcal{A}_n)(x_{n,m}) - \xi\|_{W^{k+4d,1}} = 0\,. \tag{110}$$

In addition, we can choose the sequence such that for all $m \in \mathbb{N}$, $\|(\lambda - \mathcal{A}_n)(x_{n,m})\|_{W^{k+4d,1}} \leq \|\xi\|_{W^{k+4d,1}} + 1$. Due to the ω_{k+4d}-quasi dissipativity of $\mathcal{A}_n$ (it is a generator of a strongly continuous semigroup with a bound given in Equation (109)) this immediately implies $\|x_{n,m}\|_{W^{k+4d,1}} \leq M_{k+4d} \frac{\|\xi\|_{W^{k+4d,1}}+1}{\lambda - \omega_{k+4d}} =: c_\xi$, i.e. the set $\{x_{n,m}\}_{n,m\in\mathbb{N}}$ is bounded in $W^{k+4d,1}$. We now have that for $n, m \in \mathbb{N}$

$$\begin{aligned}
\|(\lambda - \mathcal{A})(x_{n,m}) - \xi\|_{W^{k,1}} &\leq \|(\lambda - \mathcal{A}_n)(x_{n,m}) - \xi\|_{W^{k,1}} + \|(\mathcal{A} - \mathcal{A}_n)(x_{n,m})\|_{W^{k,1}} \\
&\leq \|(\lambda - \mathcal{A}_n)(x_{n,m}) - \xi\|_{W^{k,1}} + \sum_{i=1}^{K} c_{i,k} |a_i - a_{i,n}| \|x_{n,m}\|_{W^{k+4d,1}} \\
&\leq \|(\lambda - \mathcal{A}_n)(x_{n,m}) - \xi\|_{W^{k,1}} + \sum_{i=1}^{K} |a_i - a_{i,n}| c_{i,k} c_\xi \\
&\leq \|(\lambda - \mathcal{A}_n)(x_{n,m}) - \xi\|_{W^{k,1}} + C \sum_{i=1}^{K} |a_i - a_{i,n}|\,.
\end{aligned} \tag{111}$$

In the first line we used triangle inequality, in the second one the explicit form of $\mathcal{A}$ and $\mathcal{A}_n$ and then that there exists $c_{k,i} \geq 0$ such that

$$\left\|(N+\mathbb{1})^{k/4} A_{i,1} x_{n,m} A_{i,2} (N+\mathbb{1})^{k/4}\right\|_1 \leq c_{k,i} \left\|(N+\mathbb{1})^d x_{n,m} (N+\mathbb{1})^d\right\|_{W^{k,1}}$$

as in the proof of Lemma E.3. Lastly we used the uniform bound c_ξ and set $C = \max_{i=1,\dots,K} c_{i,k} c_\xi$. By a proper choice of a subsequence of $\{x_{n,m}\}_{n,m\in\mathbb{N}}$, we get that the RHS of Equation (111) vanishes. Since $\{x_{n,m}\}_{n,m\in\mathbb{N}}$ is bounded in $W^{k+4d,1}$ it is in particular precompact in $W^{k,1}$ (as of the compact embedding of the Sobolev spaces), meaning we can further choose the aforementioned sequence to be convergent in $W^{k,1}$. Let us call it $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{T}_f$. To summarise, for the chosen λ and $\xi \in \mathcal{T}_f$ arbitrary we have constructed a sequence $\{y_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(\mathcal{A})$ which is convergent in $W^{k,1}$ and further $\{(\lambda - \mathcal{A})(y_n)\}_{n\in\mathbb{N}}$ converges to ξ in $W^{k,1}$. Hence the closure of the range of $(\lambda - \mathcal{A}, \mathcal{D}(\mathcal{A}))$ contains $\mathcal{T}_f$ a dense subset of $W^{k,1}$, which concludes the proof. □

Remark. In the above lemma, it suffices to assume that the semigroups are Sobolev preserving, as one can interpolate between the sequence elements to obtain semigroups for $k \in \mathbb{R}_+$.

Appendix B

Submitted articles

Conditional Independence of 1D Gibbs States with Applications to Efficient Learning

Álvaro M. Alhambra[†1], Ángela Capel[‡2,3], Paul Gondolf∗[§2], Alberto Ruiz-de-Alarcón[¶2,4], and Samuel O. Scalet[•3]

[1]Instituto de Física Teórica UAM/CSIC, C/ Nicolás Cabrera 13-15, Cantoblanco, 28049 Madrid, Spain
[2]Fachbereich Mathematik, Universität Tübingen, 72076 Tübingen, Germany
[3]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom
[4]Departamento de Análisis Matemático y Matemática Aplicada, Universidad Complutense de Madrid, Plaza de las Ciencias 3, 28040 Madrid, Spain.

July 9, 2025

Abstract

We show that spin chains in thermal equilibrium have a correlation structure in which individual regions are strongly correlated at most with their near vicinity. We quantify this with alternative notions of the conditional mutual information, defined through the so-called Belavkin-Staszewski relative entropy. We prove that these measures decay superexponentially at every positive temperature, under the assumption that the spin chain Hamiltonian is translation-invariant. Using a recovery map associated with these measures, we sequentially construct tensor network approximations in terms of marginals of small (sublogarithmic) size. As a main application, we show that classical representations of the states can be learned efficiently from local measurements with a polynomial sample complexity. We also prove an approximate factorization condition for the purity of the entire Gibbs state, which implies that it can be efficiently estimated to a small multiplicative error from a small number of local measurements. The results extend from strictly local to exponentially-decaying interactions above a threshold temperature, albeit only with exponential decay rates. As a technical step of independent interest, we show an upper bound to the decay of the Belavkin-Staszewski relative entropy upon the application of a conditional expectation.

[†]alvaro.alhambra@csic.es
[‡]angela.capel@uni-tuebingen.de
[§]paul.gondolf@uni-tuebingen.de
[¶]albertoruizdealarcon@ucm.es
[•]sos25@cam.ac.uk
∗ corresponding author

Contents

1 Introduction 2
1.1 Main results . . . 4
1.2 Previous work and future directions . . . 7

2 Preliminaries 7
2.1 Basic notation . . . 7
2.2 Entropy measures . . . 8
2.3 Spin chains in 1D, local Hamiltonians and Gibbs states . . . 10
2.4 Approximate factorisation of Gibbs states of local Hamiltonians in 1D . . . 11

3 Main results 14
3.1 An upper bound on the DPI for the Belavkin Staszewski entropy . . . 14
3.2 Superexponential decay of the BS-CMI . . . 16
3.3 Approximate factorisation of the purity . . . 19

4 Applications 23
4.1 Positive MPO approximations from recovery maps . . . 23
4.2 Reconstructing the positive MPO from local tomography . . . 27
4.3 Efficient estimation of the global purity . . . 30

A Proofs of approximate factorization of Gibbs states in 1D 38
A.1 Proof of Lemma 2.4 . . . 38
A.2 Proof of Lemma 2.5 . . . 40

B Extension of results to exponentially-decaying interactions 41
B.1 Exponential decay of the BS-CMI for exponentially-decaying interactions . . . 43
B.2 Efficient estimation of the global purity for exponentially-decaying interactions . . . 44
B.3 Learning of Gibbs states via MPO approximations for exponentially-decaying interactions . . . 44

C Lifting the upper bound on the DPI of the BS-entropy to channels 45

1 Introduction

It is now widely established that tools and ideas from quantum information theory can give fresh perspectives to studying complex quantum many-body systems. A notable way this happens is the systematic characterization of the correlations among their constituents, which often allows us to narrow down the complexity of the many-body states in specific ways.

One of the main ways of constraining those correlations is the area law, which states that the amount of information that two adjacent regions share is upper bounded by the size of their mutual boundary. This area law has been shown for gapped ground states [67, 41, 2] using entanglement entropy, and also for Gibbs states of finite temperatures and other classes of mixed states [18, 35, 5] using the mutual information [84, 54, 71]. Importantly, area laws can be linked with the efficiency of classical algorithms in the form of tensor network methods [81, 57, 54, 38]. Another related way in which many-body states are typically constrained is by a fast decay of correlations of distant regions. This is often stated in terms of connected correlation functions [42, 19, 6, 49, 36], but also with other quantifiers such as the mutual information [16] (although they are often equivalent

[22, 50, 17]), as well as measures of quantum entanglement [56]. This decay means that distant regions behave roughly independently of each other, so complex collective effects (such as long-range order or entanglement) are absent.

There is a third, arguably more refined, way in which correlations can be constrained in many-body systems: conditional independence [43, 58]. This notion aims to quantify how much the (possibly small) correlations between two separate regions A and C are mediated by a separating region B that shields one from the other, such as in Figure 1.

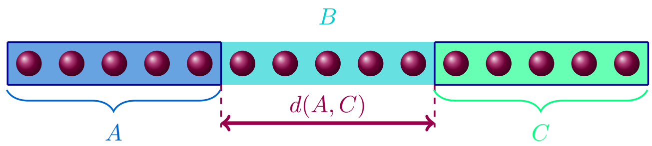


Figure 1: Regions A and C shielded by a region B.

In quantum systems, the notion of conditional independence of a state ρ is typically expressed through the quantum conditional mutual information (CMI) $I_\rho(A : C|B)$ [60], and its behaviour with the geometry of the regions A, B, C. This quantity has the following equivalent definitions

$$I_\rho(A : C|B) = S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) \tag{1}$$
$$= I_\rho(A : BC) - I_\rho(A : B) \tag{2}$$
$$= H_\rho(A|B) - H_\rho(A|BC). \tag{3}$$

That is, as a linear combination of subsystem entropies, or as a difference of mutual informations or conditional entropies. When $I_\rho(A : C|B) = 0$, we say that ρ_{ABC} is a quantum Markov state. Additionally, the fact that it is small enables us to conclude that the state ρ_{ABC} can be closely approximated from ρ_{AB} by applying a CPTP map on B alone [43, 34, 75].

The CMI has often appeared in the context of quantum many-body systems. As an analogue of the classical Hammersley-Clifford theorem [26], it is known that states in which this quantity vanishes whenever B shields A from C correspond to Gibbs states of commuting Hamiltonians [58, 21]. Additionally, an exponential decay of the CMI as B increasingly separates A and C has been shown for various instances of Gibbs states [47, 53], and matrix product states [77]. A fast decay with the distance between A and C guarantees the accuracy of local reconstruction maps [34, 46], which is linked to the efficiency of quantum algorithms for preparing Gibbs states [20], and of learning phases of matter [70]. The CMI is also prominently featured in the study of topological order at low energies [48, 74].

An arbitrarily small CMI is the main way of quantifying conditional independence, but quantum information theory provides us with a framework to systematically construct similar measures, for instance, through generalizations of Eq. (2) and Eq. (3). While their operational meaning might be entirely different [29, 10, 4], each possible definition comes with the potential for applications. In this paper, we study several such alternative notions of conditional independence. In contrast to the CMI, which is based on the Umegaki relative entropy, ours are largely based on the Belavkin-Staszewksi (BS) relative entropy [9]. We also study a measure in terms of local Rényi-2 entropies inspired by recent results on efficient learning of entropies and entanglement measures [80].

One of our main findings is that the decay of the BS-CMI in Gibbs states is strictly faster than what is believed to hold for any measure of bipartite correlations [6, 36, 16]: superexponential, as opposed to just exponential. This difference can be seen in analogy with the classical setting, in which Gibbs distributions have exactly zero CMI, while still having an exponential decay of correlations. In contrast, existing conjectures and results on the quantum CMI only display an exponential decay [20, 47, 53]. Our work thus suggests that in quantum Gibbs states there is a separation of the magnitude of conditional independence and decay of correlations, much like in the classical case.

We then study the application of our notions of conditional independence, and their fast decay, to devise efficient learning schemes for properties of quantum Gibbs states in one dimension. The problem of efficiently learning quantum Gibbs states from local measurements has seen tremendous progress recently, including methods whose efficiency is based on the decay of the CMI [70]. In particular, how to efficiently learn the Hamiltonian with a small number of samples and classical post-processing has been actively studied [7, 3, 40, 69, 51], including important experimental efforts [45], and efficient protocols at all temperatures [8]. Results along these lines have also appeared for matrix product states in 1D [27, 32]. With our techniques, based on alternative notions of conditional independence, we shift the focus away from Hamiltonian learning and instead study the problem of directly reconstructing tensor network approximations to the state, as well as the global purity.

1.1 Main results

In this paper, we focus on Gibbs states of local Hamiltonians in 1D at any inverse positive temperature $\beta > 0$. Given a finite chain $\Lambda = ABC$, with B shielding A from C as in Figure 1, and ρ a Gibbs state of a local, translation-invariant Hamiltonian on Λ, we investigate several notions of conditional independence between A and C conditioned on B, and show that they decay with the size of B.

1. Superexponential decay of the BS Conditional Mutual Information

We first consider generalisations of the CMI involving the Belavkin-Staszewski (BS) relative entropy [9], an upper bound for the Umegaki relative entropy, which is given by Eq. (12). For three notions of BS Conditional Mutual Information (BS-CMI), termed one-sided (Eq. (16)), two-sided (Eq. (17)) and reversed (Eq. (18)), we establish superexponential decay and thereby positively answer a conjecture from [16]. Note that this is faster than what is expected to hold for the CMI in this context [47, 55, 53]. We prove that all of these quantities exhibit the following scaling:

$$\widehat{I}^x_\rho(A\,;C|B) \leq c e^{\alpha|A|}\epsilon(|B|), \quad x \in \{\mathrm{os}, \mathrm{ts}, \mathrm{rev}\} \tag{4}$$

where A and C are separated by B. The function $\ell \mapsto \epsilon(\ell)$ decays superexponentially with ℓ, and c and α are universal for all intervals, only dependent on the inverse temperature and the range and strength of the interaction that gives rise to the local Hamiltonians, as well as the local dimension. The proof of this result is split into two parts, which constitute results of independent interest.

1.1. Upper bound on the DPI for the BS-entropy

The first step of the proof is an upper bound on the data-processing inequality (DPI) of the BS-entropy for conditional expectations. This bound is expressed in terms of a difference measure that establishes the relationship between ρ and its asymmetric BS-recovery condition [12], denoted by $\mathcal{B}^\sigma_{\mathcal{E}}(\cdot) := \sigma\,\mathcal{E}(\sigma)^{-1}(\cdot)$, and analogously with the roles of ρ and σ interchanged. More precisely, in Section 3.1 we show that

$$\widehat{D}(\rho\|\sigma) - \widehat{D}(\mathcal{E}(\rho)\|\,\mathcal{E}(\sigma)) \leq \begin{cases} C(\rho,\sigma,\mathcal{E})\,\|\rho\,(\mathcal{B}^\sigma_{\mathcal{E}}(\mathcal{E}(\rho)))^{-1} - \mathbb{1}\,\| \\ C'(\rho,\sigma,\mathcal{E})\,\|\sigma\,\left(\mathcal{B}^\rho_{\mathcal{E}}(\mathcal{E}(\sigma))\right)^{-1} - \mathbb{1}\,\| \end{cases}, \tag{5}$$

with $C(\rho, \sigma, \mathcal{E})$ and $C'(\rho, \sigma, \mathcal{E})$ multiplicative factors depending on ρ, σ and $\mathcal{E}$. This provides a complementary perspective on the strengthened data-processing inequality established in [12]. The reason is that it offers an upper limit on the disparity of the BS-DPI, explicitly featuring a difference between the states and their BS-recovery.

1.2. Decay of BS-CMI for Gibbs states of 1D local Hamiltonians

To prove Eq. (4), we combine the bounds of Eq. (5) with existing techniques of decay of correlations on Gibbs states of 1D local Hamiltonians, such as those in [16]. More specifically, in Section 2.4, we include a collection of technical results for these objects based on tight estimates of Araki's expansionals [6], which we use, jointly with [16, Theorem 5.1], to prove:

$$\left\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\right\| \leq \varepsilon(|B|)\,, \tag{6}$$

for ρ a Gibbs state on a possibly larger chain Λ' than $\Lambda = ABC$, with $\Lambda \subset \Lambda'$, and ρ_X its marginal on X for $X \in \{AB, B, BC, ABC\}$. Noticing that the BS-recovery of ρ_{ABC} for the partial trace in A is $\rho_{AB}\rho_B^{-1}\rho_{BC}$, we combine Eq. (6) with Eq. (5) to conclude Eq. (4).

2. Efficient estimation of the global purity of 1D local Hamiltonians

In the next part of the paper, we consider another notion of conditional independence, namely the purity. This quantity can be viewed as a version of the CMI defined in terms of Rényi-2 entropies. For this purity, using similar techniques and again results from [16] we resolve a conjecture from [80] and show that in the same setting as above, it decays exponentially with the size of $|B|$. We subsequently employ this to directly establish the sample and computational efficiency of the scheme to learn the purity from [80].

2.1. Approximate factorisation of the purity

In the first part of this result, contained in Section 3.3, we explicitly prove that

$$\left|\frac{\mathrm{Tr}_{AB}[\rho_{AB}^2]\,\mathrm{Tr}_{BC}[\rho_{BC}^2]}{\mathrm{Tr}_{ABC}[\rho_{ABC}^2]\,\mathrm{Tr}_B[\rho_B^2]} - 1\right| \leq c_p e^{-\alpha_p|B|}\,, \tag{7}$$

where ρ_{ABC} is a Gibbs state and the other states are marginals thereof. The constants c_p and $\alpha_p > 0$ are again universal and only depend on the range, inverse temperature, and strength of the interaction, as well as the local dimension.

2.2. Estimation of the global purity

Subsequently, in Section 4.3, we consider an N-partite system $\Lambda = A_1 \dots A_N$ and define an N-partite version of the purity (removing the Rényi entropy of the total chain), given by

$$P_2(\rho_{1:N}) = \frac{\prod_{j=1}^{N-1} \mathrm{Tr}_{j:j+1}[\rho_{j:j+1}^2]}{\prod_{j=2}^{N-1} \mathrm{Tr}_j[\rho_j^2]}\,. \tag{8}$$

Then, we show that, due to an iterated application of Eq. (7), we can obtain an approximation of $\mathrm{Tr}\left[\rho_{1:N}^2\right]$ by $P_2(\rho_{1:N})$ up to an error ε as long as each region A_i has size of order $\mathcal{O}(\log N/\varepsilon)$. Considering next some local measurements of each of the marginals of ρ up to an error δ, we show that there is an algorithm that outputs an estimate of $\mathrm{Tr}\left[\rho_{1:N}^2\right]$ up to a multiplicative error ε with a number of samples and classical post-processing cost of order $\mathrm{poly}(|\Lambda|/\varepsilon)$.

3. Learning of Gibbs states of 1D local Hamiltonians via Matrix Product Operator approximations

Finally, we apply our results for matrix product operator (MPO) reconstructions of Gibbs states and their efficient learning. Again, we consider a multipartite quantum state on an N-partite system. Specifically, the results on the decay of BS-CMI allow us to give an efficient MPO description of 1D Gibbs states. Furthermore, we show that this description can also be learned efficiently from local tomography. We find a polynomial sample and computational complexity in both the system size and inverse error.

3.1. Positive MPO approximations from recovery maps

Apart from the previously mentioned asymmetric recovery map, we introduce an alternative, symmetric, and thereby positive recovery map

$$\mathcal{R}_i(X) = \rho_i^{1/2}(\rho_i^{-1/2}\rho_{i:i+1}\rho_i^{-1/2})^{1/2}\rho_i^{-1/2}X\rho_i^{-1/2}(\rho_i^{-1/2}\rho_{i:i+1}\rho_i^{-1/2})^{1/2}\rho_i^{1/2}\,. \tag{9}$$

The recovery error is bounded by the BS-CMI with additional terms corresponding to the lowest eigenvalue of marginals of the thermal state and the maximal mutual information [71], all of which can be appropriately bounded for 1D Gibbs states. However, this map is not a quantum channel because it is not trace-preserving. This presents a challenge, as the non-contractive nature of the map could lead to exponential amplification of recovery errors from earlier steps in the reconstruction process. We overcome this issue by proving a Lipschitz constant on a concatenation of maps that is *independent* of the level of concatenation. We prove a representability result for the Gibbs state by the MPO obtained from concatenating these maps

$$\left\| \left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right) (\rho_1) - \rho_{1:N} \right\| \leq \varepsilon$$

with a subpolynomial bond dimension in $|\Lambda|/\varepsilon$.

3.2. Reconstruction of positive MPO from local tomography

Considering the explicit form of the recovery map Eq. (9) in terms of the marginals, we show how to learn this MPO representation and prove that the approximation error to the state is robust to small tomographic errors. We obtain the representation as an explicit formula of the estimated marginals. As sublogarithmically-sized marginals are sufficient for the reconstruction, using standard tomography results, the sample and computational cost incurred to reconstruct the state to ε error in 1-norm is again poly$(|\Lambda|/\varepsilon)$. A subpolynomial dependence in system size is also possible using an inherently translation-invariant formulation of the MPO, see Remark 4.10.

Bonus. Exponentially-decaying interactions

Many of the techniques used in our main results can or have been recently extended to Gibbs states of Hamiltonians with exponentially-decaying interactions above a threshold temperature, as seen in [64, 22, 17]. Therefore, a natural question is whether all results presented so far extend to that framework. In Appendix B we positively answer that question, with small but necessary modifications. The decay of the BS-CMI is now only exponential in $|B|$ and holds only above a critical temperature in line with previous results. In contrast, the estimation of the global purity follows the same arguments as for finite-range interactions. Moreover, we can provide a protocol to learn marginals of Gibbs states in 1D with translation-invariant, exponentially-decaying interactions, by MPO approximations, albeit with a slightly worse bond dimension than for finite-range. Nevertheless, as far as we know, this is the first MPO approximation for Gibbs states in this framework in the literature.

1.2 Previous work and future directions

It is interesting to compare our results to existing ones on the representability of thermal states by tensor networks. In general, it is known that representations with polynomial bond dimension exist for arbitrary lattices [62, 82]. Furthermore, in one dimension, an MPO description of finite Gibbs states with subpolynomial bond dimension has been shown in [54]. A feature common to all of these constructions is that they give explicit formulas in terms of the Hamiltonian. While this easily allows for a computation of the MPO given *knowledge of the systems' interactions*, it is unclear how to *learn* such a representation directly. To do that, a precise analysis of the approximation quality given the errors in the estimated Hamiltonian would be needed. Furthermore, even if such a result is achieved, the best existing Hamiltonian learning results [8, 3, 40] take polynomial time with degrees that are often impractically large. Defining our reconstruction in terms of measurable marginals, we circumvent these complications and provide a practical formula for the reconstruction. Another feature of our construction is that it directly works in the thermodynamic limit. Constructions explicitly involving the Hamiltonian would need to be truncated; however, an approach in this direction has been put forward in [1].

While a recent result in [32, 68], proposes a way of learning finitely correlated states (which can often be seen as equivalent to matrix product operators) and the application of this framework to the learning of Gibbs states, it is currently unclear whether this approach can yield an equivalent result to ours. The result is conditioned on bounds of singular values of a map involved in the MPO representation of thermal states, which are not known at present.

Finally, on a more conceptual level, the proof technique could be of independent interest in terms of extending the representability result to larger classes of states. We provide sufficient information theoretic criteria that guarantee efficient MPO approximations. It is a natural question to find out whether other types of states also satisfy these conditions. The overall idea displays a strong analogy to the technique used in the proof of representability of pure gapped ground states by matrix product states. There, an information-theoretic area law for Rényi entropies [41] implies efficient representability of the state by a matrix product state [81, 72] and is also involved in the later rigorous proof of an efficient algorithm for the ground-state energy [57].

Let us finally also comment on an alternative approach that might come to mind when looking at our construction: Using a decay of the standard quantum conditional mutual information and the corresponding recovery channels. The existence of such recovery channels has been shown [34] and since they are indeed quantum channels (CPTP-maps) their concatenation should be possible in an analogous way. There are two obstructions to this route. Firstly, the channels are not given by an equally simple explicit formula but involve an optimization procedure, which could be cumbersome to deal with on the learning side of our result [46, 76]. Secondly, no superexponential decay result is known or believed to hold for the conditional mutual information.

2 Preliminaries

This section is dedicated to introducing and reviewing the basic terminology, notions, and results of quantum systems, entropy measures, Gibbs states in quantum spin chains, and their *approximate factorization.*

2.1 Basic notation

First, all vector spaces, tensor products and operator spaces considered in this paper are defined over the field of complex numbers $\mathbb{C}$. For a quantum system labelled A, we let $\mathcal{H}_A$ denote the

Hilbert space.

A bipartite system AB arising as the composition of two systems A and B is described by the tensor product Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. In the specific context of quantum spin chains, consecutive letters representing intervals of qubits will generally imply that these intervals are adjacent and non-overlapping.

For $\mathcal{H}_A$ we use $\mathcal{B}(\mathcal{H}_A)$ to denote the space of bounded linear operators with $\|\cdot\|$ the operator norm. The trace map is $\mathrm{Tr} : \mathcal{B}(\mathcal{H}_A) \to \mathbb{C}$, sometimes with subindices if we want to emphasise the system it acts on. A similar notational convention is adapted to denote the identity map $\mathbb{1}$. In turn, for multipartite systems, we take $\mathrm{tr}_A : \mathcal{B}(\mathcal{H}_{AB}) \to \mathcal{B}(\mathcal{H}_B)$ to be the partial trace over A. The spaces of bounded linear operators can be equipped with the Schatten p-norms, defined by $\|X\|_p := (\mathrm{Tr}(|X|^p))^{1/p}$ for all $p \in [1, \infty)$, where $|X| := (X^*X)^{1/2}$ stands for the operator square root. We recall that the Schatten 1-norm coincides with the usual trace norm, defined as the sum of the singular values of the operator, and the Schatten 2-norm is the Hilbert-Schmidt norm arising from the Hilbert-Schmidt inner product on $\mathcal{B}(\mathcal{H}_A)$. Furthermore, in the limiting case $p \to +\infty$ the Schatten ∞-norm recovers the operator norm $\|\cdot\|_\infty = \|\cdot\|$, also characterized as the largest singular value of the operator. We recall that all Schatten p-norms are submultiplicative and unitarily invariant. Moreover, they are ordered in the sense that $\|X\|_p \geq \|X\|_q$ for any $1 \leq p \leq q \leq \infty$, and satisfy Hölder's inequality $\|XY\|_r \leq \|X\|_p \|Y\|_q$ for all $p, q, r \in [1, \infty]$ with $1/r = 1/p + 1/q$.

Over any system A we consider the non-negative and normalised set $\mathcal{S}(\mathcal{H}_A) \subset \mathcal{B}(\mathcal{H}_A)$ of states or density operators, i.e. non-negative operators ρ such that $\mathrm{Tr}[\rho] = 1$. In particular, we consider the maximally mixed state $\pi_A := d_A^{-1}\, \mathbb{1}_A$ on A, where $d_A = \dim \mathcal{H}_A$. We reserve the Greek letters ρ and σ for states. For any multipartite state ρ_{AB}, we denote the marginals that arise after application of the partial trace with a subindex that indicates the system they act on, i.e. $\rho_A := \mathrm{tr}_B[\rho_{AB}]$ and $\rho_B := \mathrm{tr}_A[\rho_{AB}]$.

As we will be relying on the notion of conditional expectation, we shortly want to introduce it here as well. For a von Neumann subalgebra $\mathcal{N}$ of $\mathcal{B}(\mathcal{H})$ there exists a unique map, called conditional expectation, which we will generally denote with $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$, and which satisfies the properties that it projects onto $\mathcal{N}$ orthogonally w.r.t the Hilbert-Schmidt inner product. An important example and the use case of such maps in this paper are partial traces followed by an embedding, i.e maps of the form $\pi_A \otimes \mathrm{tr}_A : \mathcal{B}(\mathcal{H}_{AB}) \to \mathbb{1}_A \otimes \mathcal{B}(\mathcal{H}_B) \subseteq \mathcal{B}(\mathcal{H}_{AB})$, where $X \in \mathcal{B}(\mathcal{H}_{AB})$ is mapped to $\pi_A \otimes \mathrm{tr}_A[X_{AB}]$.

Conditional expectations form a subset within the broader set of quantum channels, encompassing completely positive trace-preserving linear mappings. Note at last that through Stinespring's dilation theorem, every quantum channel can be represented by a composition of an isometry with a partial trace.

2.2 Entropy measures

We introduce several entropy measures that will appear throughout the text, and present some connections between them. Arguably, the most recognized measure among them is the *von Neumann entropy* of a state ρ, defined by the expression

$$S(\rho) := -\mathrm{Tr}[\rho \log(\rho)] \,. \tag{10}$$

Here and in the following we adopt the convention $0 \log 0 \equiv 0$. The von Neumann entropy captures the entropy of a quantum state and is a quantum analogous to the classical Shannon entropy [73]. Also inspired by the classical setting, one can generalise the Kullback-Leibler divergence [52] to the quantum setting. The most prominent of those generalizations is the *Umegaki relative entropy* [79].

For two quantum states ρ and σ, it is defined by the expression

$$D(\rho\|\sigma) := \begin{cases} \mathrm{Tr}[\rho\log\rho - \rho\log\sigma] & \text{if } \ker\sigma\subseteq\ker\rho\,,\\ +\infty & \text{otherwise}\,. \end{cases} \tag{11}$$

A less prominent one, which has however gained attention in the last years as tool to estimate channel capacities [31] and the decay of correlation measures for Gibbs state in 1D [16] is the *Belavkin-Staszewski relative entropy* [9] or BS-entropy for short. For any two states ρ and σ, it is defined as

$$\widehat{D}(\rho\,\|\,\sigma) := \begin{cases} \mathrm{Tr}\big[\rho\log\big(\rho^{1/2}\sigma^{-1}\rho^{1/2}\big)\big] & \text{if } \ker\sigma\subseteq\ker\rho\,,\\ +\infty & \text{otherwise}\,. \end{cases} \tag{12}$$

The reason for terming those entropic measures generalisations is that if the involved states commute the quantum measure reduces to their classical analogue. For example, in the case of the Umegaki and BS-entropy, one would recover the Kullback-Leibler divergence. Hence, both Umegaki and BS-entropy agree in the commuting case, while the BS-entropy is strictly bigger than the Umegaki relative entropy in the case of non-commuting states [44]. Note further that both divergences can be defined for general positive semidefinite matrices, and we will use those definitions under slight abuse of notation.

Whilst not featured in our decay results, we will make use of the maximal Rényi divergence in our results on MPO constructions [28]

$$D_\infty(\rho\|\sigma) = \log\inf\left\{\lambda : \rho\le\lambda\sigma\right\}\,,$$

which arises as a limit of a wider family of sandwiched Rényi entropies [61].

As we are mostly concerned with the analysis of quantities involving the BS-entropy, we skip the definitions of the analogues for the relative, respectively von Neumann entropy, and only note that conditional entropy, mutual information, and conditional mutual information were initially defined in terms of the Shannon entropy and then rewritten using the Umegaki relative entropy. Inspired by this latter representation, the BS-entropy analogues are obtained by replacing the Umegaki with the BS-entropy. We refer to [14, 15, 71, 85] for partial or complete definitions of these quantities.

For any bipartite state ρ_{AB}, we define the *BS-conditional entropy* by

$$\widehat{H}_\rho(A\,|\,B) := -\widehat{D}(\rho_{AB}\,\|\,\mathbb{1}_A\otimes\rho_B), \tag{13}$$

and the *BS-mutual information* by

$$\widehat{I}_\rho(A:B) := \widehat{D}(\rho_{AB}\,\|\,\rho_A\otimes\rho_B). \tag{14}$$

Equivalently, we also introduce the *maximal mutual information* [71]

$$I_\infty(A:B) := D_\infty(\rho_{AB}\|\rho_A\otimes\rho_B). \tag{15}$$

The definition of the BS-conditional mutual information (BS-CMI) is relatively more ambiguous and hence we recall the following definitions from [14]. Let us consider a tripartite quantum system ABC. Then, for a quantum state ρ_{ABC} we define the *one-sided BS-conditional mutual information* between the systems A and C conditioned on the system B by the expression

$$\widehat{I}^{\mathrm{os}}_\rho(A\,;C\,|\,B) := \widehat{D}(\rho_{ABC}\,\|\,\pi_A\otimes\rho_{BC}) - \widehat{D}(\rho_{AB}\,\|\,\pi_A\otimes\rho_B), \tag{16}$$

and the *two-sided BS-conditional mutual information* as

$$\widehat{I}^{\mathrm{ts}}_\rho(A\,;C\,|\,B) := \widehat{D}(\rho_{ABC}\,\|\,\rho_A\otimes\rho_{BC}) - \widehat{D}(\rho_{AB}\,\|\,\rho_A\otimes\rho_B). \tag{17}$$

Lastly, we define the *reversed BS-conditional mutual information* by

$$\widehat{I}^{\mathrm{rev}}_{\rho}(A\,;C\mid B):=\widehat{D}(\pi_A\otimes\rho_{BC}\,\|\,\rho_{ABC})-\widehat{D}(\pi_A\otimes\rho_B\,\|\,\rho_{AB}),\tag{18}$$

where we recall that π_A stands for the maximally mixed state on A.

2.3 Spin chains in 1D, local Hamiltonians and Gibbs states

We write $\Lambda\Subset\mathbb{Z}$ for Λ being a finite subset of $\mathbb{Z}$, use $|\Lambda|$ to denote its cardinality and $\mathrm{diam}(\Lambda):=\max\{x-y:x,y\in\Lambda\}$ its diameter. With every site $x\in\mathbb{Z}$ we associate a finite-dimensional Hilbert space $\mathcal{H}_x:=\mathbb{C}^d$ with corresponding linear operators $\mathcal{A}_x:=\mathcal{B}(\mathcal{H}_x)$. We generalise this concept to finite sets, where for $\Lambda\Subset\mathbb{Z}$, we define the Hilbert space $\mathcal{H}_\Lambda:=\bigotimes_{x\in\Lambda}\mathcal{H}_x$ and the algebra of linear operators as $\mathcal{A}_\Lambda:=\mathcal{B}(\mathcal{H}_\Lambda)=\bigotimes_{x\in\Lambda}\mathcal{B}(\mathcal{H}_x)$. The last equality holds due to the finite-dimensional nature of $\mathcal{H}_\Lambda$. For $\Lambda'\subseteq\Lambda\Subset\mathbb{Z}$, we can therefore consider $X\in\mathcal{A}_{\Lambda'}$ as an element of $\mathcal{A}_\Lambda$ by identifying X with $X\otimes\mathbb{1}_{\Lambda\setminus\Lambda'}$. This identification is kept implicit and allows us to define the *algebra of local observables* for general $\Sigma\subseteq\mathbb{Z}$ simply by

$$\mathcal{A}_\Sigma:=\overline{\bigcup_{\Lambda\Subset\Sigma}\mathcal{A}_\Lambda}^{\|\cdot\|},\tag{19}$$

where the closure is taken with respect to the operator norm. Now an *interaction* on $\Sigma\subseteq\mathbb{Z}$ is defined as

$$\Phi:\{\Lambda\Subset\Sigma\}\to\mathcal{A}_{\mathbb{Z}},\quad \Lambda\mapsto\Phi(\Lambda)\in\mathcal{A}_\Lambda\quad\text{with}\quad\Phi(\Lambda)=\Phi(\Lambda)^*\,,\tag{20}$$

and its corresponding local Hamiltonian on any $\Lambda\Subset\Sigma$ as

$$H_\Lambda:=\sum_{\Lambda'\subseteq\Lambda}\Phi(\Lambda')\,.\tag{21}$$

In this paper, we focus on *finite-range interactions*, characterised by the existence of parameters $R>0$ and $J>0$ such that $\Phi(\Lambda)=0$ whenever the diameter of Λ exceeds R, and

$$\sum_{\Lambda\Subset\mathbb{Z}\,:\,x\in\Lambda}\|\Phi(\Lambda)\|\le J\,,$$

for all $x\in\mathbb{Z}$. In the literature, it is common to alternatively impose the condition that for all $\Lambda\Subset\mathbb{Z}$, $\|\Phi(\Lambda)\|\le\widetilde{J}$, which can be related to J by observing that, due to the interaction vanishing if the diameter of the input exceeds R,

$$\sum_{\Lambda\Subset\mathbb{Z}\,:\,x\in\Lambda}\|\Phi(\Lambda)\|=\sum_{k=0}^{R}\;\sum_{\Lambda\in\mathcal{P}(\{x-k,\dots,x-k+R\}\setminus\{x\})}\|\Phi(\{x\}\cup\Lambda)\|\le(R+1)2^R\widetilde{J}\,.$$

In addition to the finite range, we further require interactions to exhibit *translation invariance*. In Appendix B, we extend all our results to the setting of exponentially-decaying interactions, whose formalism we introduce therein.

Lastly, we define the *Gibbs state* for a Hamiltonian $H\in\mathcal{B}(\mathcal{H})$ with $H=H^*$ over an arbitrary (finite-dimensional) Hilbert space $\mathcal{H}$ at inverse temperature $\beta\in(0,\infty)$ as

$$\rho^\beta_{\mathcal{H}}[H]:=\frac{e^{-\beta H}}{\mathrm{Tr}[e^{-\beta H}]}\,.\tag{22}$$

In Section 3.2, i.e., when we prove superexponential decay of Gibbs states of local translation-invariant Hamiltonians on a quantum spin system, we drop the index of the Hilbert space, absorb the inverse temperature into the Hamiltonian and write for $\Lambda \Subset \Sigma \subseteq \mathbb{Z}$ just

$$\rho^{\Lambda} := \rho^{\beta}_{\mathcal{H}_\Lambda}[H_\Lambda] \,. \tag{23}$$

Notice that Gibbs states of finite β are always full rank.

2.4 Approximate factorisation of Gibbs states of local Hamiltonians in 1D

We now present the technical lemmas we need for the decay of correlations and uniformity of 1D Gibbs states. While they hold at all temperatures for any finite-range interaction, the constants involved depend on range, interaction strength and temperature. This dependence is mostly the same for all involved quantities, so we introduce the following convenient notation.

Let $\Lambda \Subset \mathbb{Z}$ be a finite interval. Let us split Λ into two subintervals X and Y so that $\Lambda = X \cup Y$ or $\Lambda = XY$ for short and write

$$E_{X,Y} \,:=\, e^{-\,H_{XY}}\, e^{\,H_X \,+\, H_Y} \,. \tag{24}$$

Note that $E^{*}_{X,Y} \,=\, e^{\,H_X \,+\, H_Y}\, e^{-\,H_{XY}}$ and $E^{-1}_{X,Y} \,=\, e^{-H_X \,-\, H_Y}\, e^{\,H_{XY}}$ and that we absorbed β into the Hamiltonian for better readability. The following proposition is extracted from [16, 64] and contains an alternative formulation of Araki's results for estimates on expansionals [6].

Proposition 2.1 ([16, Corollary 3.4])**.** *Let Φ be an interaction of finite-range R and strength J over $\mathbb{Z}$, at any inverse temperature $\beta < \infty$, which is further translation invariant. Then the following hold:*

(i) There is an absolute constant $\mathcal{G} > 1$ depending only on J, R and β such that, for any finite interval $\Lambda = XY \Subset \mathbb{Z}$ split into two subintervals X and Y, we have:

$$\|E_{X,Y}\| \,,\, \left\|E^{-1}_{X,Y}\right\| \;\leq\; \mathcal{G} \,.$$

(ii) There is a positive and decreasing function $\ell \mapsto \delta(\ell)$ with superexponential decay and depending on J, R and β such that if we add two intervals $\widetilde{X}$ and $\widetilde{Y}$ adjacent to X and Y, respectively, so that we get a larger interval $\widetilde{X}XY\widetilde{Y}$, then

$$\left\|E^{-1}_{X,Y} - E^{-1}_{\widetilde{X}X,Y\widetilde{Y}}\right\| \,,\, \left\|E_{X,Y} - E_{\widetilde{X}X,Y\widetilde{Y}}\right\| \;\leq\; \delta(\ell) \,.$$

for any $\ell \in \mathbb{N}$ such that $\ell \,\leq\, |X| \,,\, |Y|$.

Note in addition that for $\Lambda' \subset \Lambda \Subset \mathbb{Z}$ with local Hamiltonian $H_{\Lambda'}$ supported in $\mathcal{H}_{\Lambda'}$ but lifted to $\mathcal{H}_\Lambda$, the map

$$\mathcal{A}_\Lambda \to \mathcal{A}_\Lambda \,, \quad Q \mapsto \operatorname{tr}_{\Lambda'}[e^{-H_{\Lambda'}}Q] = \operatorname{tr}_{\Lambda'}[e^{-\frac{1}{2}H_{\Lambda'}}Qe^{-\frac{1}{2}H_{\Lambda'}}] \tag{25}$$

where $\operatorname{tr}_{\Lambda'}$ denotes the partial trace in Λ', has the following property:

$$\left\|\operatorname{tr}_{\Lambda'}[e^{-H_{\Lambda'}}Q]\right\| \leq \left\|\operatorname{tr}_{\Lambda'}[e^{-H_{\Lambda'}}]\right\| \|Q\| = \operatorname{Tr}_{\Lambda'}[e^{-H_{\Lambda'}}]\|Q\| \,. \tag{26}$$

As a consequence,

$$\mathcal{A}_\Lambda \to \mathcal{A}_\Lambda \,, \quad Q \mapsto \operatorname{tr}_{\Lambda'}[\rho^{\Lambda'}Q] \tag{27}$$

is positive, unital and hence contractive. This observation is essential for the following result, in which we are considering an interval Λ split into three subintervals $\Lambda = ABC$, where B shields A from C, as illustrated in Figure 2.

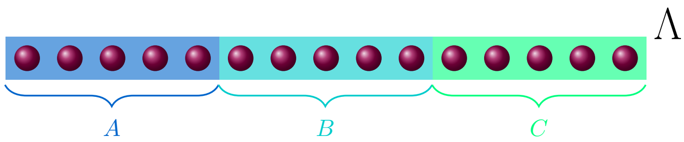


Figure 2: Representation of an interval Λ split into three subintervals $\Lambda = ABC$, where B shields A from C.

Proposition 2.2 ([16, Section 3.4])**.** *In the context of Proposition 2.1, let $\Lambda = ABC \Subset \mathbb{Z}$ (we admit the possibility of some subintervals being empty). Then, there is an absolute constant $\mathcal{C}$ depending only on the strength J, range R of the interaction and inverse temperature β, such that*

$$\left\| \operatorname{tr}_B[\rho^B Q] \right\|, \left\| \operatorname{tr}_B[\rho^B Q]^{-1} \right\| \leq \mathcal{C} \,, \qquad Q \in \{E^*_{B,C}\,, E_{B,C}\,, E^*_{A,B}\,, E_{A,B}\}\,, \tag{28}$$

$$\left\| \operatorname{tr}_{AB}[\rho^{AB} Q] \right\|, \left\| \operatorname{tr}_{AB}[\rho^{AB} Q]^{-1} \right\| \leq \mathcal{C} \,, \qquad Q \in \{E^{*\,-1}_{A,B}\,, E^{-1}_{A,B}\}\,, \tag{29}$$

$$\left\| \operatorname{tr}_B\left[\rho^B E^*_{A,B} E^*_{AB,C}\right] \right\|, \left\| \operatorname{tr}_B\left[\rho^B E^*_{A,B} E^*_{AB,C}\right]^{-1} \right\| \leq \mathcal{C}\,. \tag{30}$$

These two propositions are instrumental in the proof of the following result, which provides an approximate factorization of the Gibbs state of a finite-range, translation-invariant Hamiltonian.

Theorem 2.3 ([16, Eq. (17) in Theorem 5.1])**.** *For Φ a finite-range, translation-invariant interaction over $\mathbb{Z}$ there exists a positive function $\ell \mapsto \varepsilon(\ell)$ with superexponential decay in ℓ, depending only on J, R, and β such that for every $\Lambda \Subset \mathbb{Z}$ split into three subintervals $\Lambda = ABC$, where B shields A from C and $|B| \geq \ell$ for its local Gibbs state $\rho^\Lambda =: \rho_{ABC}$ it holds that*

$$\left\| \rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1} \right\| \leq \varepsilon(\ell)\,. \tag{31}$$

The form of the superexponentially-decaying function is given by

$$\varepsilon(\ell) = \mathcal{C}_1 \frac{\mathcal{C}_2^{1+\lfloor \ell/2 \rfloor}}{(\lfloor \lfloor \ell/2 \rfloor / R \rfloor + 1)!}$$

for constants $\mathcal{C}_1, \mathcal{C}_2$ that only depend on J, R, and β.

Note that, in Eq. (31), only ρ_{ABC} denotes a Gibbs state of a local Hamiltonian while ρ_{AB}, ρ_{BC} and ρ_B are marginals after partially tracing out systems. We further suppressed tensoring with identity, i.e., ρ_{AB} for example has to be understood as $\rho_{AB} \otimes \mathbb{1}_C$.

In the rest of the paper, we will need a refined version of this theorem for Gibbs states on $\Lambda = A'ABCC'$, for which we only compare the marginals in AB, B, BC and ABC as above. To generalize this, however, we need to introduce the following lemmata. The first constitutes a generalization of Eq. (28) in Proposition 2.2.

Lemma 2.4. *Let $\Lambda \Subset \mathbb{Z}$ be a finite interval split into five subintervals $\Lambda = A'ABCC'$ as in Figure 3. Then, for $\rho = \rho^\Lambda$ the Gibbs state on Λ, and for every $\beta > 0$, the following holds:*

$$\left\| \operatorname{tr}_{AC}[\rho^A \otimes \rho^C Q]^{-1} \right\| \leq \mathcal{C} \tag{32}$$

for $Q = E_{AB,C}\, E_{A,B}$ and a constant $\mathcal{C}$ that only depends on J, R, β and d.

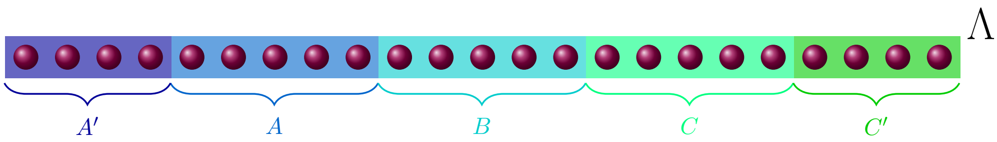


Figure 3: Representation of an interval Λ split into five subintervals $\Lambda = A'ABCC'$.

We defer the proof of Lemma 2.4 to Appendix A. Next, we show that the decay provided by the measure of conditional independence in Theorem 2.3 can be reduced from the total Gibbs state to a large marginal of it. The proof of this lemma is also deferred to Appendix A.

Lemma 2.5. *Let $\Lambda \Subset \mathbb{Z}$ be a finite interval split into five subintervals $\Lambda = A'ABCC'$ as in Figure 3. Then, for $\rho = \rho^\Lambda$ the Gibbs state on Λ, and for every $\beta > 0$, the following holds*

$$\left\| \rho_{AB}\, \rho_B^{-1} \rho_{BC}\, \rho_{ABC}^{-1} - \mathbb{1} \right\| < \mathcal{C}^4 \left\| \rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1} - \mathbb{1} \right\| , \tag{33}$$

for $\mathcal{C} > 1$ depending on R, J and β, and given in Proposition 2.2.

Remark 2.6. *Note that Eq.* (33) *can be interpreted as a modified data-processing inequality for the partial trace for the mixing condition between two positive states η and σ,*

$$\left\| \eta\, \sigma^{-1} - \mathbb{1} \right\| ,$$

with contraction coefficient upper bounded by $\mathcal{C}^4$, identifying $\eta := \rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}$, $\sigma := \rho_{A'ABCC'}$, and taking the partial trace in $A'C'$.

Combining now the findings of Theorem 2.3 and Lemma 2.5, we can conclude the following:

Corollary 2.7. *For Φ a finite-range, translation-invariant interaction over $\mathbb{Z}$, there exists a positive function $\ell \mapsto \tilde{\varepsilon}(\ell)$ with superexponential decay in ℓ, depending only on J, R and β, such that for every $\Lambda \Subset \mathbb{Z}$ split into three subintervals $\Lambda = A'ABCC'$, with $|B| \geq \ell$, for its local Gibbs state $\rho^\Lambda =: \rho_{A'ABCC'}$ it holds that*

$$\left\| \rho_{ABC} \rho_{BC}^{-1} \rho_B \rho_{AB}^{-1} - \mathbb{1} \right\| \leq \tilde{\varepsilon}(\ell) \, . \tag{34}$$

In addition to the above, the authors in [16, Theorem 6.2] showed exponential decay of correlations between spatially separated regions. This will become relevant to us later in the form of local indistinguishability, as per Theorem 2.9.

Proposition 2.8 (Decay of correlations [16, Theorem 6.2])**.** *For Φ a finite-range, translation-invariant interaction over $\mathbb{Z}$ there exist positive constants c, α, depending only on J, R and β such that for every $\Lambda \Subset \mathbb{Z}$ split into three subintervals $\Lambda = ABC$, where B shields A from C and for $O_A \in \mathcal{A}_A$ and $O_C \in \mathcal{A}_C$ it holds that*

$$\left| \operatorname{Tr}_{ABC}[\rho^{ABC} O_A\, O_C] - \operatorname{Tr}_{ABC}[\rho^{ABC} O_A] \operatorname{Tr}_{ABC}[\rho^{ABC} O_C] \right| \leq \|O_A\| \|O_C\| \, c\, e^{-\alpha |B|} \, .$$

The exponential decay of correlations can be used to prove local indistinguishability of Gibbs states, a result that can be found in [16, Proposition 7.1].

Theorem 2.9 (Local indistinguishability [16, Proposition 7.1]). *For Φ a finite-range, translation-invariant interaction over $\mathbb{Z}$, there exist positive constants c', α', depending only on J, R and β, such that for every $\Lambda \Subset \mathbb{Z}$ split into three subintervals $\Lambda = ABC$, where B shields A from C, and for $O_A \in \mathcal{A}_A$ and $O_C \in \mathcal{A}_C$ it holds that*

$$\left|\mathrm{Tr}_{ABC}[\rho^{ABC} O_A] - \mathrm{Tr}_{AB}[\rho^{AB} O_A]\right| \leq \|O_A\|\, c'\, e^{-\alpha'|B|}\,,$$
$$\left|\mathrm{Tr}_{ABC}[\rho^{ABC} O_C] - \mathrm{Tr}_{BC}[\rho^{BC} O_C]\right| \leq \|O_C\|\, c'\, e^{-\alpha'|B|}\,.$$

All these results will be essential for our proof that the purity of a Gibbs state of a translation-invariant, local Hamiltonian in 1D decays exponentially fast with the size of the middle system (cf. Section 3.3).

Remark 2.10. *It is possible to extend all the results to the context of short-range (i.e. exponentially-decaying) interactions, as shown in [64, 22]. We review all these extensions and provide necessary new ones in Appendix B.*

3 Main results

3.1 An upper bound on the DPI for the Belavkin Staszewski entropy

The search for the presence and tightness of data-processing inequality (DPI) in entropy measures, namely that they cannot increase under quantum channels, constitute relevant problems in quantum Shannon theory. For the Umegaki relative entropy, it was originally shown by Petz [65, 66] that saturation of the DPI is equivalent to recoverability of one state in terms of the other, through the so-called *Petz recovery map.* A strengthening of the DPI for the relative entropy, namely lower bounds to the DPI after applying a conditional expectation (or in general, a quantum channel) have been obtained in the past few years [76, 24, 37], starting with the breakthrough of [34]. These results have also been extended to the larger families of standard f-divergences [23, 44]. Results in the converse direction, i.e. upper bounds for the DPI, can be found e.g. in [14].

We can now ask a similar question for the BS-entropy. This quantity also satisfies a DPI, which was proven to be saturated if and only if [12]

$$\widehat{D}(\rho\|\sigma) - \widehat{D}(\mathcal{E}(\rho)\|\,\mathcal{E}(\sigma)) = 0 \Leftrightarrow \sigma = \rho\,\mathcal{E}(\rho)^{-1}\,\mathcal{E}(\sigma) = \mathcal{B}^{\rho}_{\mathcal{E}}(\mathcal{E}(\sigma))\,. \tag{35}$$

The term on the right-hand side above was coined *BS-recovery condition.* In the same paper, a strengthened DPI for the BS-entropy (and any maximal f-divergence) in terms of the distance between a state and its BS-recovery was provided.

Here, we would like to complement this picture by giving an upper bound on the DPI of the BS-entropy. More specifically, we provide an upper bound for $\widehat{D}(X\|Y) - \widehat{D}(\mathcal{E}(X)\|\,\mathcal{E}(Y))$ in terms of distance measures that explicitly feature the recovery condition. We use the usual representation of the logarithm we want to recall here. For $X > 0$, we have that

$$\log(X) = \int_0^\infty \left(\frac{1}{t+1} - \frac{1}{t+X}\right) dt\,. \tag{36}$$

Using this representation we derive the following simple but essential lemma.

Lemma 3.1. *Let $X, Y \in \mathcal{B}(\mathcal{H})$ with $X > 0$, $Y > 0$ and $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ a conditional expectation. We further set $Z = X^{1/2}Y^{-1}X^{1/2}$, $X_{\mathcal{E}} := \mathcal{E}(X)$, $Y_{\mathcal{E}} := \mathcal{E}(Y)$ and finally $Z_{\mathcal{E}} = X_{\mathcal{E}}^{1/2}Y_{\mathcal{E}}^{-1}X_{\mathcal{E}}^{1/2}$. For $T > 0$, we find that*

$$\int_0^T \mathrm{Tr}\left[X_{\mathcal{E}}\frac{1}{t+Z_{\mathcal{E}}} - X\frac{1}{t+Z}\right] dt = \int_0^T \mathrm{Tr}\left[X_{\mathcal{E}}^{1/2}\frac{1}{t+Z_{\mathcal{E}}}X_{\mathcal{E}}^{-1/2}(XY^{-1} - X_{\mathcal{E}}Y_{\mathcal{E}}^{-1})X^{1/2}\frac{1}{t+Z}X^{1/2}\right] dt\,.$$

Proof. We find by cyclicity of the trace and the fact that $\mathcal{E}$ is a conditional expectation that

$$\int_0^T \operatorname{Tr}\left[X_{\mathcal{E}}\frac{1}{1+Z_{\mathcal{E}}} - X\frac{1}{t+Z}\right]dt = \int_0^T \operatorname{Tr}\left[X_{\mathcal{E}}^{1/2}\frac{1}{t+Z_{\mathcal{E}}}X_{\mathcal{E}}^{-1/2}X - X^{1/2}\frac{1}{t+Z}X^{-1/2}X\right]dt\,.$$
$$= \int_0^T \operatorname{Tr}\left[\left(X_{\mathcal{E}}^{1/2}\frac{1}{t+Z_{\mathcal{E}}}X_{\mathcal{E}}^{-1/2} - X^{1/2}\frac{1}{t+Z}X^{-1/2}\right)X\right]dt\,.$$

Recalling that, for any two invertible matrices A and B, one can write $A^{-1}-B^{-1} = A^{-1}(B-A)B^{-1}$, we note that

$$\begin{aligned}
&X_{\mathcal{E}}^{1/2}\frac{1}{t+Z_{\mathcal{E}}}X_{\mathcal{E}}^{-1/2} - X^{1/2}\frac{1}{t+Z}X^{-1/2}\\
&\quad= X_{\mathcal{E}}^{1/2}\frac{1}{t+Z_{\mathcal{E}}}X_{\mathcal{E}}^{-1/2}(X^{1/2}(Z+t)X^{-1/2} - X_{\mathcal{E}}^{1/2}(Z_{\mathcal{E}}+t)X_{\mathcal{E}}^{-1/2})X^{1/2}\frac{1}{t+Z}X^{-1/2}\\
&\quad= X_{\mathcal{E}}^{1/2}\frac{1}{t+Z_{\mathcal{E}}}X_{\mathcal{E}}^{-1/2}(XY^{-1} - X_{\mathcal{E}}Y_{\mathcal{E}}^{-1})X^{1/2}\frac{1}{t+Z}X^{-1/2}
\end{aligned}$$

and inserting it into the above equation gives the claim. □

With this lemma at hand, we are set to prove the upper bound on the BS-entropy DPI.

Theorem 3.2. *Let $X, Y \in \mathcal{B}(\mathcal{H})$ with $X > 0$, $Y > 0$ and $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ a conditional expectation. Then, we have*

$$\begin{aligned}
\widehat{D}(X\|Y) - \widehat{D}(\mathcal{E}(X)\|\,\mathcal{E}(Y)) \leq{}& \left\|X^{-1/2}YX^{-1/2}\right\|\|X\|_1\left\|\mathcal{E}(X)^{1/2}\right\|\left\|\mathcal{E}(X)^{-1/2}\right\|\\
&\cdot\left\|\mathcal{E}(Y)^{-1}\,\mathcal{E}(X)\right\|\left\|XY^{-1}\,\mathcal{E}(Y)\,\mathcal{E}(X)^{-1} - \mathbb{1}\right\|.
\end{aligned} \tag{37}$$

and

$$\begin{aligned}
\widehat{D}(X\|Y) - \widehat{D}(\mathcal{E}(X)\|\,\mathcal{E}(Y)) \leq{}& \left\|X^{-1/2}YX^{-1/2}\right\|\|X\|_1\left\|\mathcal{E}(X)^{1/2}\right\|\left\|\mathcal{E}(X)^{-1/2}\right\|\\
&\cdot\left\|Y^{-1}X\right\|\left\|YX^{-1}\,\mathcal{E}(X)\,\mathcal{E}(Y)^{-1} - \mathbb{1}\right\|.
\end{aligned} \tag{38}$$

Proof. Both inequalities follow from the same estimate through the application of Hölder inequality. To obtain this expression we employ Equation (36) using the notation of Lemma 3.1

$$\widehat{D}(X\|Y) - \widehat{D}(X_{\mathcal{E}}\|Y_{\mathcal{E}}) = \lim_{T\to\infty}\int_0^T \operatorname{Tr}\left[X_{\mathcal{E}}\frac{1}{t+Z_{\mathcal{E}}} - X\frac{1}{t+Z}\right]dt\,.$$

Further, using Lemma 3.1 and Hölder's inequality we can upper bound

$$\begin{aligned}
&\int_0^T \operatorname{Tr}\left[X_{\mathcal{E}}\frac{1}{t+Z_{\mathcal{E}}} - X\frac{1}{t+Z}\right]dt\\
&\quad\leq \int_0^T \left\|X_{\mathcal{E}}^{-1/2}\right\|\frac{1}{t+\left\|Z_{\mathcal{E}}^{-1}\right\|^{-1}}\left\|X_{\mathcal{E}}^{1/2}\right\|\|XY^{-1} - X_{\mathcal{E}}Y_{\mathcal{E}}^{-1}\|\left\|X^{1/2}\right\|_2\frac{1}{t+\|Z^{-1}\|^{-1}}\left\|X^{1/2}\right\|_2 dt\,.
\end{aligned}$$

Note also that $\left\|\frac{1}{t+Z}\right\| = \frac{1}{t+\|Z^{-1}\|^{-1}}$ and analogously for the fraction including $Z_{\mathcal{E}}$. Utilising DPI for $(A,B) \mapsto \left\|A^{1/2}B^{-1}A^{1/2}\right\|$ (see e.g. [78, Proposition 4.7]), we can estimate $\frac{1}{t+\left\|Z_{\mathcal{E}}^{-1}\right\|^{-1}} \leq \frac{1}{t+\|Z^{-1}\|^{-1}}$.

$$\begin{aligned}
&\int_0^T \operatorname{Tr}\left[X_{\mathcal{E}}\frac{1}{t+Z_{\mathcal{E}}} - X\frac{1}{t+Z}\right]dt\\
&\qquad\leq \left\|X_{\mathcal{E}}^{-1/2}\right\|\left\|X_{\mathcal{E}}^{1/2}\right\|\left\|X^{1/2}\right\|_2^2\|XY^{-1} - X_{\mathcal{E}}Y_{\mathcal{E}}^{-1}\|\int_0^T \frac{1}{\left(t+\|Z^{-1}\|^{-1}\right)^2}dt\,.
\end{aligned}$$

Finally, the fact that $\left\|X^{1/2}\right\|_2^2 = \|X\|_1$, integrating both sides, and then taking the limit $T \to \infty$ give

$$\widehat{D}(X\|Y) - \widehat{D}(X_{\mathcal{E}}\|Y_{\mathcal{E}}) \leq \left\|X_{\mathcal{E}}^{-1/2}\right\|\left\|X^{-1/2}YX^{-1/2}\right\|\left\|X_{\mathcal{E}}^{1/2}\right\|\left\|XY^{-1} - X_{\mathcal{E}}Y_{\mathcal{E}}^{-1}\right\|\|X\|_1. \tag{39}$$

□

Remark 3.3. *The above result can readily be extended to quantum channels instead of conditional expectations. The statement and proof of that result can be found in the appendix in Corollary C.1.*

Remark 3.4. *As mentioned in the introduction, the above result complements the findings of [12], providing a comprehensive perspective on the data processing inequality of the BS-entropy. Using the notation of the theorem for quantum states ρ and σ, and the BS-recovery for σ denoted by $\mathcal{B}^{\rho}_{\mathcal{E}}(\mathcal{E}(\sigma)) = \rho\,\mathcal{E}(\rho)^{-1}\,\mathcal{E}(\sigma)$, the following chain of inequalities hold:*

$$\begin{aligned}\left(\frac{\pi}{8}\right)^4 &\left\|\rho^{-1/2}\sigma\rho^{-1/2}\right\|^{-4}\left\|\mathcal{E}(\rho)^{-1}\right\|^{-2}\left\|\mathcal{B}^{\rho}_{\mathcal{E}}(\mathcal{E}(\sigma)) - \sigma\right\|_2^4 \\ &\leq \widehat{D}(\rho\|\sigma) - \widehat{D}(\mathcal{E}(\rho)\|\,\mathcal{E}(\sigma)) \leq \\ &\left\|\rho^{-1/2}\sigma\rho^{-1/2}\right\|\left\|\mathcal{E}(\rho)^{-1}\right\|^{1/2}\left\|\mathcal{E}(\sigma)^{-1}\,\mathcal{E}(\rho)\right\|\left\|\sigma^{-1}\right\|\left\|\mathcal{B}^{\rho}_{\mathcal{E}}(\mathcal{E}(\sigma)) - \sigma\right\|,\end{aligned}$$

where the upper bound follows from Equation (39).

3.2 Superexponential decay of the BS-CMI

In this section, we prove one of our main results: the superexponential decay of the three definitions of the BS-CMI. The proof crucially involves Theorem 3.2, as well as the results from Sec. 2.4. Before we do so, we need a technical lemma concerning norm estimates of functions of Gibbs states of local translation invariant Hamiltonians on a spin chain.

Lemma 3.5. *Let Φ be a finite-range, translation invariant-interaction over $\mathbb{Z}$ and $\Lambda \Subset \mathbb{Z}$ split into three subintervals $\Lambda = ABC$, where B shields A from C, with local Gibbs state $\rho^{\Lambda} = \rho_{ABC}$. Then*

1. $\left\|\rho_A\rho_B\rho_{AB}^{-1}\right\| \leq \mathcal{C}$, $\left\|\rho_{AB}\rho_A^{-1}\rho_B^{-1}\right\| \leq \mathcal{C}$, $\left\|\rho_{ABC}\rho_B^{-1}\right\| \leq \mathcal{C}$,
2. $\left\|\rho_B^{-1}\right\|\|\rho_B\| \leq \mathcal{C}\,e^{\alpha|B|}$,

where the constants $\mathcal{C}, \alpha$ only depend on interaction strength J and range R of Φ. Note that only ρ_{ABC} is the Gibbs state of a local Hamiltonian, while all other states are marginals of that state. Furthermore, these conditions still hold if $\rho_{ABC} = \operatorname{tr}_{A'C'}[\rho_{A'ABCC'}]$ where $\rho_{A'ABCC'}$ is the Gibbs state on the larger system $A'ABCC'$.

Proof. Let us denote the partition function for a given interval by

$$Z_A = \operatorname{Tr}\left[e^{-H_A}\right].$$

Note that the perturbation formula [59, Lemma 3.6]

$$\left|\log\left(\operatorname{Tr}\left[e^{H+P}\right]\right) - \log\left(\operatorname{Tr}\left[e^{H}\right]\right)\right| \leq \|P\|$$

implies that ratios of partition functions are bounded by a constant $\mathcal{C}$ only depending on the interaction strength, range, and inverse temperature

$$\frac{Z_A Z_B}{Z_{AB}}, \frac{Z_{AB}}{Z_A Z_B} \leq \mathcal{C}\ .$$

For the first part, we identify CC' with B'. We use an argument analogous to Corollary 2.7.

$$\begin{aligned}&\|\rho_A\rho_B\rho_{AB}^{-1}\|\\ &\quad=\Big\|\operatorname{tr}_{A'BB'}[\rho^{A'}\otimes\rho^{BB'}E_{A',ABB'}E_{A,BB'}]\operatorname{tr}_{A'AB'}[\rho^{A'A}\otimes\rho^{B'}E_{A'A,BB'}E_{B,B'}]E_{A,B}^{-1}\\ &\qquad\operatorname{tr}_{A'B'}[\rho^{A'}\otimes\rho^{B'}E_{A',ABB'}E_{AB,B'}]^{-1}\Big\|\frac{Z_{A'A}Z_{BB'}}{Z_{A'ABB'}}\,.\end{aligned}$$

Using submultiplicativity of the norm, the first two partial traces are bounded by contractivity of the conditional expectation and Proposition 2.1. The same proposition also bounds $\|E_{A,B}^{-1}\|$. For the inverse partial trace we use Lemma 2.4 for the appropriately identified subsystems. Analogously,

$$\begin{aligned}&\|\rho_{AB}\rho_A^{-1}\rho_B^{-1}\|\\ &\quad=\Big\|\operatorname{tr}_{A'B'}[\rho^{A'}\otimes\rho^{B'}E_{A',ABB'}E_{AB,B'}]E_{A,B}\operatorname{tr}_{A'BB'}[\rho^{A'}\otimes\rho^{BB'}E_{A',ABB'}E_{A,BB'}]^{-1}\\ &\qquad\operatorname{tr}_{A'AB'}[\rho^{A'A}\otimes\rho^{B'}E_{A'A,BB'}E_{B,B'}]^{-1}\Big\|\frac{Z_{A'ABB'}}{Z_{A'A}Z_{BB'}}\,,\end{aligned}$$

and, switching the notation back to a Gibbs state on the original $A'ABCC'$,

$$\begin{aligned}&\|\rho_{ABC}\rho_B^{-1}\|\\ &\quad=\Big\|\operatorname{tr}_{A'C'}[\rho^{A'}\otimes\rho^{C'}E_{A',ABCC'}E_{ABC,C'}]E_{A,BC}E_{B,C}\frac{e^{-H_A}}{Z_A}\frac{e^{-H_C}}{Z_C}\\ &\qquad\operatorname{tr}_{A'ACC'}[\rho^{A'A}\otimes\rho^{CC'}E_{A'A,BCC'}E_{B,CC'}]^{-1}\Big\|\frac{Z_{A'}Z_AZ_CZ_{C'}}{Z_{AA'}Z_{CC'}}\end{aligned}$$

are bounded in the same way, apart from the last equation, which also uses $\|\rho^A\|,\|\rho^C\|\le 1$.

The second point follows from similar estimates, but no partition functions appear due to the cancellation from the two norms.

$$\|\rho_B^{-1}\|\|\rho_B\|=\|(\rho^B)^{-1}\operatorname{tr}_{AC}[E_{A,BC}E_{B,C}\rho^A\otimes\rho^C]^{-1}\|\,\|(\rho^B)\operatorname{tr}_{AC}[E_{A,BC}E_{B,C}\rho^A\otimes\rho^C]\|$$

Again, employing Lemma 2.4 for the norms of inverses of partial traces, we are left with

$$\Big\|(\rho^B)^{-1}\Big\|\|\rho^B\|=\|e^{H_B}\|\|e^{-H_B}\|\le e^{2\|H_B\|}$$

which yields the claim choosing $\alpha=2J$.

□

We are now set to prove the main result of the section.

Theorem 3.6. *For Φ a finite-range, translation-invariant interaction over $\mathbb{Z}$, there exists a positive function $\ell\mapsto\epsilon(\ell)$ with superexponential decay in ℓ, depending only on J, R and β such that for every $\Lambda\Subset\mathbb{Z}$ split into consecutive subintervals $\Lambda=A'ABCC'$, with A' and C' possibly empty, for the marginal on ABC of its local Gibbs state $\operatorname{tr}_{A'C'}[\rho^\Lambda]=\rho_{ABC}$ it holds that*

$$\widehat{I}^x_\rho(A\,;C|B)\le ce^{\alpha|A|}\epsilon(|B|)\quad x\in\{\mathrm{os},\mathrm{ts},\mathrm{rev}\}\,,\tag{40}$$

Here c and α are constants only depending on inverse temperature β, strength J and range R of Φ.

Proof. We begin with the one-sided version and note that we can write

$$\widehat{I}^{\mathrm{os}}_\rho(A\,;C|B) = \widehat{D}(\rho_{ABC}\|\pi_A\otimes\rho_{BC}) - \widehat{D}(\mathcal{E}(\rho_{ABC})\|\,\mathcal{E}(\pi_A\otimes\rho_{BC}))$$

with $\mathcal{E}(\cdot) := \mathrm{tr}_C[\cdot]\otimes\pi_C$ a conditional expectation. Using Theorem 3.2 we obtain

$$\widehat{I}^{\mathrm{os}}_\rho(A\,;C|B) \le \left\|\rho_{ABC}^{-1/2}\rho_{BC}\rho_{ABC}^{-1/2}\right\|\|\rho_{ABC}\|_1\left\|\rho_{AB}^{1/2}\right\|\left\|\rho_{AB}^{-1/2}\right\|\|\rho_B^{-1}\rho_{AB}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\|$$

where we already simplified terms and cancelled constants. The fact that $\|AB\|\le\|BA\|$ for normal AB [11, Proposition IX.1.1] and that for quantum states $\|\rho^{-p}\| = \|\rho^{-1}\|^p$, $p\in[0,\infty)$, gives us

$$\begin{aligned}\widehat{I}^{\mathrm{os}}_\rho(A\,;C|B) &\le \|\rho_{BC}\rho_{ABC}^{-1}\|\left(\|\rho_{AB}^{-1}\|\|\rho_{AB}\|\right)^{1/2}\|\rho_B^{-1}\rho_{AB}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\|\\ &\le (\|\rho_A^{-1}\|\|\rho_A\|)\|\rho_A\rho_{BC}\rho_{ABC}^{-1}\|\left(\|\rho_{AB}^{-1}\|\|\rho_{AB}\|\right)^{1/2}\\ &\qquad\cdot\|\rho_A^{-1}\rho_B^{-1}\rho_{AB}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\|\\ &\le ce^{\alpha(|A|+|B|)}\varepsilon(|B|)\,.\end{aligned}$$

In the last inequality, we used Theorem 2.3 with Lemma 2.5 and Lemma 3.5.

For the two-sided definition, we again use Theorem 3.2 with the same conditional expectation and $X=\rho_{ABC}$ but with $Y=\rho_A\otimes\rho_{BC}$ instead. Applying similar simplifications as before, Theorem 2.3, Lemma 2.5, and Lemma 3.5, we obtain

$$\begin{aligned}\widehat{I}^{\mathrm{ts}}_\rho(A\,;C|B) &\le \|\rho_A\rho_{BC}\rho_{ABC}^{-1}\|\left(\|\rho_{AB}^{-1}\|\|\rho_{AB}\|\right)^{1/2}\|\rho_A^{-1}\rho_B^{-1}\rho_{AB}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\|\\ &\le ce^{\alpha(|A|+|B|)}\varepsilon(|B|)\,.\end{aligned}$$

For the reversed version, we set $X=\pi_A\otimes\rho_{BC}$, $Y=\rho_{ABC}$ with the conditional expectation fixed as before. Employing Theorem 3.2 in the form of the second inequality we obtain after initial simplifications

$$\begin{aligned}\widehat{I}^{\mathrm{rev}}_\rho(A\,;C|B) &\le \left\|\rho_{BC}^{-1/2}\rho_{ABC}\rho_{BC}^{-1/2}\right\|\left(\|\rho_B^{-1}\|\|\rho_B\|\right)^{1/2}\|\rho_{ABC}^{-1}\rho_{BC}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\|\\ &\le \|\rho_{ABC}\rho_A^{-1}\rho_{BC}^{-1}\|\left(\|\rho_B^{-1}\|\|\rho_B\|\right)^{1/2}\|\rho_A^{-1}\|\|\rho_A\|\\ &\qquad\cdot\|\rho_A\rho_{BC}\rho_{ABC}^{-1}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\|\\ &\le ce^{\alpha(|A|+|B|)}\varepsilon(|B|)\,.\end{aligned}$$

The last estimation again follows from Theorem 2.3, Lemma 2.5 and Lemma 3.5.

Ultimately, for all bounds, we absorbed the exponential growth in the B system into the super-exponential decay, yielding the claimed result. □

Note that all the BS-CMI bounds depend on the size of one of the side systems exponentially. At least in the case of the two-sided version, one can relatively easily show a *constant* upper bound on the DPI independent of the dimension, which might hint towards possible improvements. Lastly, we would also like to restate Remark 3.4 and translate it to the context of one-sided BS-CMI to make the inequalities more accessible.

Remark 3.7. *In the case where ρ_{ABC} is a state on a tripartite system (not necessarily with any relation to a Gibbs state), the following chain of inequalities holds:*

$$\begin{aligned}&\left(\frac{\pi}{8}\right)^4\left\|\rho_{ABC}^{-1/2}\rho_{BC}\rho_{ABC}^{-1/2}\right\|^{-4}\|\rho_{AB}^{-1}\|^{-2}\|\rho_B\rho_{AB}^{-1}\rho_{ABC} - \rho_{BC}\|\\ &\qquad\le \widehat{I}^{\mathrm{os}}_\rho(A\,;C|B) \le\\ &\qquad\left\|\rho_{ABC}^{-1/2}\rho_{BC}\rho_{ABC}^{-1/2}\right\|(\|\rho_{AB}^{-1}\|\|\rho_{AB}\|)^{1/2}\|\rho_B^{-1}\rho_{AB}\|\|\rho_{ABC}\rho_{BC}^{-1}\rho_{AB}\rho_B^{-1} - \mathbb{1}\|\,.\end{aligned}$$

Remark 3.8. *An analogue of Theorem 3.6 for exponentially-decaying interactions is provided in Theorem B.5. The main difference is the exponential instead of the superexponential decay with* $|B|$*, which only holds for* β *small enough.*

3.3 Approximate factorisation of the purity

We now consider another possible measure of conditional independence given by the approximate factorisation of the purity. This notion is inspired by [80], where it was shown that for a ρ_Λ prepared by a finite depth circuit with $\Lambda \Subset \mathbb{Z}$ split as $\Lambda = ABC$ it holds that

$$\frac{\mathrm{Tr}_{AB}[\rho_{AB}^2]\,\mathrm{Tr}_{BC}[\rho_{BC}^2]}{\mathrm{Tr}_\Lambda[\rho_\Lambda^2]\,\mathrm{Tr}_B[\rho_B^2]} = 1 \tag{41}$$

whenever $|B| \geq 2\ell - 1$ with ℓ being the depth of the circuit. This may suggest a definition of conditional independence based on a notion of CMI defined in terms of Rényi-2 entropies, as, in analogy with Eq. (1),

$$I_2(A:C|B) := \log \frac{\mathrm{Tr}_\Lambda[\rho_\Lambda^2]\,\mathrm{Tr}_B[\rho_B^2]}{\mathrm{Tr}_{AB}[\rho_{AB}^2]\,\mathrm{Tr}_{BC}[\rho_{BC}^2]} = S_2(\rho_{AB}) + S_2(\rho_{BC}) - S_2(\rho_B) - S_2(\rho_\Lambda)\,, \tag{42}$$

with $S_2(\rho) = -\log \mathrm{Tr}\big[\rho^2\big]$. It should not be confused with the 2-CMI that could be defined from Petz Rényi or sandwiched Rényi divergences. However, notice that this quantity, contrary to $I(A:C|B)$, $\widehat{I}^{\mathrm{os}}_\rho(A\,;C\mid B)$, $\widehat{I}^{\mathrm{ts}}_\rho(A\,;C\mid B)$ and $\widehat{I}^{\mathrm{rev}}_\rho(A\,;C\mid B)$, is not necessarily positive, and likely lacks most other relevant information-theoretic properties.

Nevertheless, motivated by the problem of efficiently learning Rényi entropies the authors of [80] introduce an approximate factorisation condition for this measure, which holds when the previous Eq. (41) fails up to, at most, an exponentially small error in $|B|$. They proved that translation-invariant Matrix Product Density Operators satisfy this property, conjectured it for a larger class of states, and numerically verified it for some relevant models. This approximate factorisation is equivalent to the exponential decay in $|B|$ of Eq. (42). In this section we show this property for any translation-invariant, finite-range Hamiltonian in 1D at any inverse temperature $\beta > 0$.

Proposition 3.9. *Let* Φ *be a finite-range translation-invariant interaction over* $\mathbb{Z}$*. Then, there exist positive constants* c_p, α_p *depending only on the strength* J*, range of the interaction* R *and inverse temperature* $\beta > 0$ *with the following property: For every* $\Lambda \Subset \mathbb{Z}$ *split as* $\Lambda = ABC$*, where* B *shields* A *from* C *(see Figure 4), and for* $\rho_\Lambda := \rho^\Lambda$ *the Gibbs state on* Λ*,*

$$\left| \frac{\mathrm{Tr}_{AB}[\rho_{AB}^2]\,\mathrm{Tr}_{BC}[\rho_{BC}^2]}{\mathrm{Tr}_\Lambda[\rho_\Lambda^2]\,\mathrm{Tr}_B[\rho_B^2]} - 1 \right| \leq c_p e^{-\alpha_p |B|}\,. \tag{43}$$

Proof. Consider $\ell \in \mathbb{N}$ such that $|B| \geq 3\ell$ and split B into B_1, B_2 and B_3 such that $|B_1|, |B_2|, |B_3| \geq \ell$ as in Figure 4. For any $X \Subset \Lambda$, let us denote $Z_X := \mathrm{Tr}_X[e^{-H_X}]$ and by ρ^X the Gibbs state on X, i.e. $\rho^X := e^{-H_X}/Z_X$ (notice the difference with the marginal of the global state ρ_X). Let us further define

$$\lambda_{ABC} := \frac{Z_{ABC} Z_B}{Z_{AB} Z_{BC}}\,. \tag{44}$$

Then, we can rewrite λ_{ABC} in terms of expansionals as

$$\begin{aligned}\lambda_{ABC} &= \frac{\mathrm{Tr}_{ABC}[e^{-H_{ABC}}]\,\mathrm{Tr}_B[e^{-H_B}]}{\mathrm{Tr}_{AB}[e^{-H_{AB}}]\,\mathrm{Tr}_{BC}[e^{-H_{BC}}]} \\ &= \big(\mathrm{Tr}_{ABC}[\rho^{ABC} \underbrace{e^{H_{ABC}} e^{-H_A - H_{BC}}}_{E^{-1*}_{A,BC}}]\big)^{-1} \mathrm{Tr}_{AB}[\rho^{AB} \underbrace{e^{H_{AB}} e^{-H_A - H_B}}_{E^{-1*}_{A,B}}]\,.\end{aligned}$$

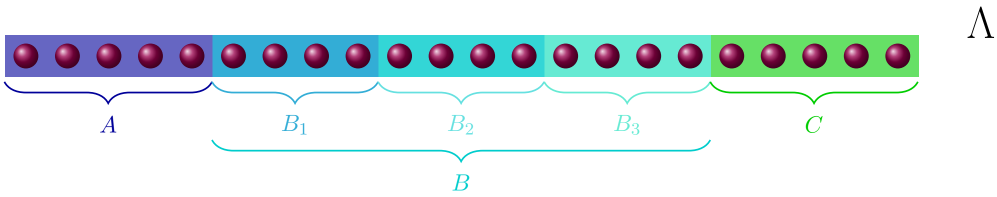


Figure 4: Representation of an interval Λ split into three subintervals $\Lambda = ABC$, with B further split into B_1, B_2 and B_3 such that $|B_1|, |B_2|, |B_3| \geq \ell$.

By Proposition 2.2, we have

$$|\lambda_{ABC}|, |\lambda_{ABC}^{-1}| \leq \mathcal{C}^2 \,, \tag{45}$$

and by Step 2 in the proof of Proposition 8.1 of [16], there exist $\tilde{c}_1, \alpha_1 > 0$ such that

$$|\lambda_{ABC} - 1| \leq \tilde{c}_1 e^{-\alpha_1 \ell} \,. \tag{46}$$

Note that the above two estimates can be used to obtain

$$|\lambda_{ABC}^{-1} - 1| = |\lambda_{ABC}^{-1}||1 - \lambda_{ABC}| \leq \mathcal{C}^2 \, \tilde{c}_1 e^{-\alpha_1 \ell} =: c_1 e^{-\alpha_1 \ell} \tag{47}$$

We will follow similar but more technical steps for

$$\widehat{\lambda}_{ABC}^{-1} := \frac{\mathrm{Tr}_{AB}[\rho_{AB}^2] \, \mathrm{Tr}_{BC}[\rho_{BC}^2]}{\mathrm{Tr}_\Lambda[\rho_\Lambda^2] \, \mathrm{Tr}_B[\rho_B^2]} \,. \tag{48}$$

Let us analyse each of the terms involved independently. For that, we will consider and compare the partition functions and Gibbs states associated to H_Λ and to $2H_\Lambda$, denoting $\widetilde{Z}_X := \mathrm{Tr}_X[e^{-2H_X}]$ and $\widetilde{\rho}_X := e^{-2H_X}/\widetilde{Z}_X$ for $X \Subset \Lambda$.

- For $\mathrm{Tr}_\Lambda[\rho_\Lambda^2]$, we have:

$$\mathrm{Tr}_\Lambda[\rho_\Lambda^2] = \frac{\mathrm{Tr}_{ABC}\left[e^{-2H_{ABC}}\right]}{Z_{ABC}^2} = \frac{\widetilde{Z}_{ABC}}{Z_{ABC}^2} \,.$$

- For $\mathrm{Tr}_{AB}[\rho_{AB}^2]$, we obtain:

$$\begin{aligned}
\mathrm{Tr}_{AB}[\rho_{AB}^2] &= \frac{1}{Z_{ABC}^2} \mathrm{Tr}_{AB}\left[\mathrm{tr}_C[e^{-H_{ABC}}]^2\right] \\
&= \frac{Z_C^2}{Z_{ABC}^2} \mathrm{Tr}_{AB}\Big[e^{-H_{AB}} \,\mathrm{tr}_C\big[\rho^C \underbrace{e^{H_{AB}} e^{H_C} e^{-H_{ABC}}}_{E_{AB,C}^*}\big] \,\mathrm{tr}_C\big[\underbrace{e^{-H_{ABC}} e^{H_C} e^{H_{AB}}}_{E_{AB,C}} \rho^C\big] e^{-H_{AB}}\Big] \\
&= \frac{\widetilde{Z}_{AB} Z_C^2}{Z_{ABC}^2} \mathrm{Tr}_{AB}\Big[\widetilde{\rho}^{AB} \,\mathrm{tr}_C\left[\rho^C E_{AB,C}^*\right] \mathrm{tr}_C\left[E_{AB,C}\rho^C\right]\Big] \,.
\end{aligned}$$

- Similarly for $\mathrm{Tr}_{BC}[\rho_{BC}^2]$, we can write:

$$\mathrm{Tr}_{BC}[\rho_{BC}^2] = \frac{Z_A^2 \widetilde{Z}_{BC}}{Z_{ABC}^2} \mathrm{Tr}_{BC}\Big[\widetilde{\rho}^{BC} \,\mathrm{tr}_A\left[\rho^A E_{A,BC}^*\right] \mathrm{tr}_A\left[E_{A,BC}\rho^A\right]\Big] \,.$$

- Finally, for $\mathrm{Tr}_B[\rho_B^2]$, we have:

$$\begin{aligned}\mathrm{Tr}_B[\rho_B^2] &= \frac{1}{Z_{ABC}^2}\,\mathrm{Tr}_B[\mathrm{tr}_{AC}[e^{-H_{ABC}}]^2] \\ &= \frac{Z_A^2 Z_C^2}{Z_{ABC}^2}\,\mathrm{Tr}_B\Big[e^{-H_B}\,\mathrm{tr}_{AC}\big[\rho^A\otimes\rho^C\,\underbrace{e^{H_A}e^{H_B}e^{-H_{AB}}}_{E^*_{A,B}}\underbrace{e^{H_{AB}}e^{H_C}e^{-H_{ABC}}}_{E^*_{AB,C}}\big] \\ &\qquad\qquad \mathrm{tr}_{AC}\big[\underbrace{e^{-H_{ABC}}e^{H_{AB}}e^{H_C}}_{E_{AB,C}}\underbrace{e^{-H_{AB}}e^{H_A}e^{H_B}}_{E_{A,B}}\rho^A\otimes\rho^C\big]e^{-H_B}\Big] \\ &= \frac{Z_A^2 Z_C^2 \widetilde{Z}_B}{Z_{ABC}^2}\,\mathrm{Tr}_B\Big[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\big[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\big]\,\mathrm{tr}_{AC}\big[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\big]\Big]\,.\end{aligned}$$

Denoting

$$\widetilde{\lambda}_{ABC} := \frac{\widetilde{Z}_{ABC}\widetilde{Z}_B}{\widetilde{Z}_{AB}\widetilde{Z}_{BC}} \tag{49}$$

and replacing all these values in Eq. (48), after noticing that all partition functions Z_X of the original Hamiltonian cancel, we obtain

$$\widehat{\lambda}^{-1}_{ABC} = \widetilde{\lambda}^{-1}_{ABC}\frac{\mathrm{Tr}_{AB}\Big[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\big[\rho^C E^*_{AB,C}\big]\,\mathrm{tr}_C\big[E_{AB,C}\rho^C\big]\Big]\,\mathrm{Tr}_{BC}\Big[\widetilde{\rho}^{BC}\,\mathrm{tr}_A\big[\rho^A E^*_{A,BC}\big]\,\mathrm{tr}_A\big[E_{A,BC}\rho^A\big]\Big]}{\mathrm{Tr}_B\Big[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\big[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\big]\,\mathrm{tr}_{AC}\big[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\big]\Big]}\,.$$

Let us denote the last fraction above by χ_{ABC}. Then, it is clear that

$$\begin{aligned}\left|\frac{\mathrm{Tr}_{AB}[\rho_{AB}^2]\,\mathrm{Tr}_{BC}[\rho_{BC}^2]}{\mathrm{Tr}_\Lambda[\rho_\Lambda^2]\,\mathrm{Tr}_B[\rho_B^2]}-1\right| &= \left|\widetilde{\lambda}^{-1}_{ABC}\chi_{ABC}-1\right| \\ &\leq \left|\widetilde{\lambda}^{-1}_{ABC}-1\right| + \left|\widetilde{\lambda}^{-1}_{ABC}\right|\left|\chi_{ABC}-1\right| \\ &\leq c_1 e^{-\alpha_1\ell} + \mathcal{C}^2\left|\chi_{ABC}-1\right|\,,\end{aligned}$$

where we are using Eq. (45) and Eq. (46) for the partition functions associated to the $2H_\Lambda$. Note that the constants would change slightly, just modifying their dependence from β to $\beta/2$, but we keep the same notation for simplicity. Finally, we bound the last term in the expression above. We will repeat the combination of Proposition 2.1 and Theorem 2.9.

Let us consider a splitting of B into B_1, B_2 and B_3 such that $|B_1|,|B_2|,|B_3|\geq\ell$ (see Figure 4). First, by triangle inequality, Hölder's inequality, point (ii) of Proposition 2.1 and Proposition 2.2, we get

$$\begin{aligned}&\left|\mathrm{Tr}_{AB}\Big[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\big[\rho^C E^*_{AB,C}\big]\,\mathrm{tr}_C\big[E_{AB,C}\rho^C\big]\Big] - \mathrm{Tr}_{AB}\Big[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\big[\rho^C E^*_{B_3,C}\big]\,\mathrm{tr}_C\big[E_{B_3,C}\rho^C\big]\Big]\right| \\ &\quad= \Big|\mathrm{Tr}_{AB}\Big[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\big[\rho^C\big(E^*_{AB,C}-E^*_{B_3,C}\big)\big]\,\mathrm{tr}_C\big[E_{B_3,C}\rho^C\big]\Big] \\ &\qquad + \mathrm{Tr}_{AB}\Big[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\big[\rho^C E^*_{AB,C}\big]\,\mathrm{tr}_C\big[\left(E_{AB,C}-E_{B_3,C}\right)\rho^C\big]\Big]\Big| \\ &\quad\leq \mathcal{C}\big\|E^*_{AB,C}-E^*_{B_3,C}\big\| + \mathcal{C}\|E_{AB,C}-E_{B_3,C}\| \\ &\quad\leq 2\mathcal{C}\delta(\ell)\,.\end{aligned} \tag{50}$$

Similarly for the analogous term tracing out BC,

$$\left|\mathrm{Tr}_{BC}\Big[\widetilde{\rho}^{BC}\,\mathrm{tr}_A\big[\rho^A E^*_{A,BC}\big]\,\mathrm{tr}_A\big[E_{A,BC}\rho^A\big]\Big] - \mathrm{Tr}_{BC}\Big[\widetilde{\rho}^{BC}\,\mathrm{tr}_A\big[\rho^A E^*_{A,B_1}\big]\,\mathrm{tr}_A\big[E_{A,B_1}\rho^A\big]\Big]\right| \leq 2\mathcal{C}\delta(\ell)\,. \tag{51}$$

The term tracing out B is slightly more involved but follows the same lines:

$$\begin{aligned}&\Big|\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\left[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\right]\mathrm{tr}_{AC}\left[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\right]\right]\\&\qquad-\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\left[\rho^A\otimes\rho^C E^*_{A,B_1}\otimes E^*_{B_3,C}\right]\mathrm{tr}_{AC}\left[E_{A,B_1}\otimes E_{B_3,C}\rho^A\otimes\rho^C\right]\right]\Big|\\&=\Big|\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\left[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\right]\mathrm{tr}_{AC}\left[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\right]\right]\\&\qquad-\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{A}\left[\rho^A E^*_{A,B_1}\right]\mathrm{tr}_{A}\left[E_{A,B_1}\rho^A\right]\otimes\mathrm{tr}_C\left[\rho^C E^*_{B_3,C}\right]\mathrm{tr}_C\left[E_{B_3,C}\rho^C\right]\right]\Big|\\&\qquad\le 4\mathcal{C}^3\delta(\ell)\,.\end{aligned}\tag{52}$$

Additionally, note that by Lemma 2.4

$$\begin{aligned}&\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\left[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\right]\mathrm{tr}_{AC}\left[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\right]\right]\\&\qquad\ge\left\|(\mathrm{tr}_{AC}\left[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\right]\mathrm{tr}_{AC}\left[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\right])^{-1}\right\|^{-1}\ge\mathcal{C}^{-2}\end{aligned}\tag{53}$$

Now, combining Eq. (50), Eq. (51), Eq. (52) and Eq. (53), we have

$$\begin{aligned}&|\chi_{ABC}-1|\\&\le\mathcal{C}^2\Big|\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{AC}\left[\rho^A\otimes\rho^C E^*_{A,B}E^*_{AB,C}\right]\mathrm{tr}_{AC}\left[E_{AB,C}E_{A,B}\rho^A\otimes\rho^C\right]\right]\\&\quad-\mathrm{Tr}_{AB}\left[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\left[\rho^C E^*_{AB,C}\right]\mathrm{tr}_C\left[E_{AB,C}\rho^C\right]\right]\mathrm{Tr}_{BC}\left[\widetilde{\rho}^{BC}\,\mathrm{tr}_A\left[\rho^A E^*_{A,BC}\right]\mathrm{tr}_A\left[E_{A,BC}\rho^A\right]\right]\Big|\\&\le\mathcal{C}^2\Big|\mathrm{Tr}_B\left[\widetilde{\rho}^B\,\mathrm{tr}_{A}\left[\rho^A E^*_{A,B_1}\right]\mathrm{tr}_{A}\left[E_{A,B_1}\rho^A\right]\otimes\mathrm{tr}_C\left[\rho^C E^*_{B_3,C}\right]\mathrm{tr}_C\left[E_{B_3,C}\rho^C\right]\right]\\&\quad-\mathrm{Tr}_{AB}\left[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\left[\rho^C E^*_{B_3,C}\right]\mathrm{tr}_C\left[E_{B_3,C}\rho^C\right]\right]\mathrm{Tr}_{BC}\left[\widetilde{\rho}^{BC}\,\mathrm{tr}_A\left[\rho^A E^*_{A,B_1}\right]\mathrm{tr}_A\left[E_{A,B_1}\rho^A\right]\right]\Big|\\&\quad+4\mathcal{C}^5\delta(\ell)+2\mathcal{C}^5\delta(\ell)+2\mathcal{C}^5\delta(\ell)\,.\end{aligned}$$

To conclude, we will estimate the difference in the previous expression using the local indistinguishability of Gibbs states as in Theorem 2.9. Indeed, note that

$$\Big|\mathrm{Tr}_{BC}\left[\widetilde{\rho}^{BC}\,\mathrm{tr}_A\left[\rho^A E^*_{A,B_1}\right]\mathrm{tr}_A\left[E_{A,B_1}\rho^A\right]\right]-\mathrm{Tr}_B\left[\widetilde{\rho}^{B}\,\mathrm{tr}_A\left[\rho^A E^*_{A,B_1}\right]\mathrm{tr}_A\left[E_{A,B_1}\rho^A\right]\right]\Big|\le\mathcal{C}^2\,c_2\,e^{-\alpha_2\ell}\,,$$

as well as

$$\Big|\mathrm{Tr}_{AB}\left[\widetilde{\rho}^{AB}\,\mathrm{tr}_C\left[\rho^C E^*_{B_3,C}\right]\mathrm{tr}_C\left[E_{B_3,C}\rho^C\right]\right]-\mathrm{Tr}_B\left[\widetilde{\rho}^{B}\,\mathrm{tr}_C\left[\rho^C E^*_{B_3,C}\right]\mathrm{tr}_C\left[E_{B_3,C}\rho^C\right]\right]\Big|\le\mathcal{C}^2\,c_2\,e^{-\alpha_2\ell}\,,$$

for certain constants $c_2,\alpha_2>0$. Therefore, denoting

$$\xi_{B_1}:=\mathrm{tr}_A\left[\rho^A E^*_{A,B_1}\right]\mathrm{tr}_A\left[E_{A,B_1}\rho^A\right]\,,\quad\xi_{B_3}:=\mathrm{tr}_C\left[\rho^C E^*_{B_3,C}\right]\mathrm{tr}_C\left[E_{B_3,C}\rho^C\right]\,,$$

we have

$$\begin{aligned}&\left|\mathrm{Tr}_B\left[\widetilde{\rho}^B\xi_{B_1}\otimes\xi_{B_3}\right]-\mathrm{Tr}_{AB}\left[\widetilde{\rho}^{AB}\xi_{B_3}\right]\mathrm{Tr}_{BC}\left[\widetilde{\rho}^{BC}\xi_{B_1}\right]\right|\\&\qquad\le\left|\mathrm{Tr}_B\left[\widetilde{\rho}^B\xi_{B_1}\otimes\xi_{B_3}\right]-\mathrm{Tr}_{B}\left[\widetilde{\rho}^{B}\xi_{B_3}\right]\mathrm{Tr}_{B}\left[\widetilde{\rho}^{B}\xi_{B_1}\right]\right|+2\mathcal{C}^4\,c_2\,e^{-\alpha_2\ell}\,,\end{aligned}$$

and we conclude the proof by using exponential decay of correlations as in Proposition 2.8 giving us an upper estimate $\mathcal{C}^4\,c_3e^{-\alpha_3\ell}$. Combining all of the above results finally gives the claim:

$$|\widehat{\lambda}^{-1}_{ABC}-1|\le c_1e^{-\alpha_1\ell}+8\,\mathcal{C}^7\,\delta(\ell)+\mathcal{C}^8(2c_2e^{-\alpha_2\ell}+c_3e^{-\alpha_3\ell})\,.$$

□

Remark 3.10. *Note that Proposition 3.9 also holds for exponentially-decaying interactions, with the same proof, just by adapting to that case the technical tools employed. This is the content of Proposition B.6.*

4 Applications

This section is devoted to applications of the main results from the previous section in the context of MPO approximations and learning of Gibbs states. In Section 4.1, we show that a set of information-theoretic criteria, most importantly the decay of the BS-CMI, imply an efficient MPO representation of a state, which by the previous section exists for one-dimensional Gibbs states. Using the explicit form of this reconstruction in Section 4.2, we show that this representation can be learned by tomography of small marginals. In Section 4.3, we outline the scheme for estimating the global purity given Proposition 3.9.

4.1 Positive MPO approximations from recovery maps

In this section, we provide a sequential reconstruction using a symmetric recovery map for the BS-CMI, newly introduced in this manuscript. The motivation for this map arises from Eq. (35), where we recalled that the DPI for the BS-entropy saturates if, and only if, each state can be recovered from the other by the so-called (asymmetric) BS-recovery condition. This condition, albeit appealing and useful for applications in the context of Gibbs states (cf. Theorem 3.6), is operationally flawed because of its lack of positivity, let alone Hermiticity. The desire to have a related positive map encourages us to define the following *symmetric recovery map* for a particular case in which both states are defined in a tripartite space and the conditional expectation considered is a partial trace:

$$\mathcal{R}(X) = \rho_B^{1/2}(\rho_B^{-1/2}\rho_{AB}\rho_B^{-1/2})^{1/2}\rho_B^{-1/2}X\rho_B^{-1/2}(\rho_B^{-1/2}\rho_{AB}\rho_B^{-1/2})^{1/2}\rho_B^{1/2} \,. \tag{54}$$

Surprisingly, this is a completely positive linear map with the same fixed points as the BS-recovery condition in a tripartite space, which is a trace-preserving linear map, but not even positive [13]. The definition of this map is not arbitrary, since it follows from the combination of some bounds obtained in [12] and [24], as we will see in the Lemma below. Interestingly, it yields a single-shot recovery error bound that involves the inverse BS-CMI, the lowest eigenvalue of marginals and the maximal mutual information.

Lemma 4.1. *Given a tripartite state ρ_{ABC}, the following bound on the distance between the state and the recovery map from Eq.* (54) *holds*

$$\widehat{I}_\rho^{\mathrm{rev}}(A\,;C|B) = \widehat{D}(\pi_A\otimes\rho_{BC}\|\rho_{ABC}) - \widehat{D}(\pi_A\otimes\rho_B\|\rho_{AB}) \tag{55}$$

$$\geq \left(\frac{\pi}{8}\right)^4 \|\Gamma\|^{-2}\|\,\mathcal{R}(\rho_{BC}) - \rho_{ABC}\|_1^4 \,, \tag{56}$$

where $\Gamma = \rho_{BC}^{-1/2}\rho_{ABC}\rho_{BC}^{-1/2}$.

Remark 4.2. *Note that $\mathcal{R}(X)$ is completely positive but not trace-preserving, so it is not a quantum channel. Several choices of recovery map for the BS-entropy are possible, see [12] for an alternative definition. The one in the above Lemma, despite its more complicated form, is necessary for proving the bound on the Lipschitz constants below and has not been considered before to the best of our knowledge.*

Proof of Lemma 4.1. We use a particular case of the strengthened data-processing inequality for the BS-entropy [12, Theorem 5.3], where we choose the channel to be the conditional expectation $\mathcal{E}(\cdot) = \mathrm{tr}_C[\cdot]\otimes\pi_C$, $\sigma = \pi_A\otimes\rho_{BC}$ for the first argument, and $\rho = \rho_{ABC}$ for the second (note the different convention for the naming of arguments). Employing now [12, Theorem 5.3] gives the bound

$$\begin{aligned}\widehat{I}^{\mathrm{rev}}(A\,;C|B) &= \widehat{D}(\pi_A \otimes \rho_{BC} \| \rho_{ABC}) - \widehat{D}(\pi_A \otimes \rho_B \| \rho_{AB}) \\ &\geq \left(\frac{\pi}{4}\right)^4 \|\Gamma\|^{-2} \|\rho_{BC}^{1/2} \rho_B^{-1/2} \Gamma_{\mathcal{E}}^{1/2} \rho_B^{1/2} - \Gamma^{1/2} \rho_{BC}^{1/2}\|_2^4 \,,\end{aligned} \tag{57}$$

where $\Gamma = \rho_{BC}^{-1/2} \rho_{ABC} \rho_{BC}^{-1/2}$ and $\Gamma_{\mathcal{E}} = \rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2}$. Noticing that

$$\begin{aligned}\mathrm{Tr}[\mathcal{R}(\rho_{BC})] &= \mathrm{Tr}_{ABC}[\rho_B^{1/2} (\rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2})^{1/2} \rho_B^{-1/2} \rho_{BC} \rho_B^{-1/2} (\rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2})^{1/2} \rho_B^{1/2}] \\ &= \mathrm{Tr}_{AB}[\rho_B^{1/2} (\rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2}) \rho_B^{1/2}] = 1\end{aligned}$$

in addition to $\mathrm{Tr}_{ABC}[\rho_{ABC}] = 1$ allows us to employ [24, Lemma 2.2], which states

$$\|X^* X - Y^* Y\|_1 \leq 2 \|X - Y\|_2$$

for operators X, Y such that $\mathrm{Tr}[X^* X] = \mathrm{Tr}[Y^* Y] = 1$. Considering here

$$X = \rho_{BC}^{1/2} \rho_B^{-1/2} \Gamma_{\mathcal{E}}^{1/2} \rho_B^{1/2} \,, \quad Y = \Gamma^{1/2} \rho_{BC}^{1/2} \,,$$

this, in turn, gives a lower bound to Eq. (57) and a strengthened DPI for the inverse BS-CMI involving the distance of ρ_{ABC} to $\mathcal{R}(\rho_{BC})$ in 1 norm, i.e. the claim of Lemma 4.1. □

As we will see in the next pages, the recovery map introduced in Eq. (54), when applied iteratively, provides a set of information-theoretic criteria for the existence of MPO descriptions of an arbitrary state. For that, let us consider a sequence of recovery maps that can reconstruct a quantum state on a spin chain. We assume an appropriately subdivided chain on subintervals $A_1, \dots, A_N$ as in Figure 5.

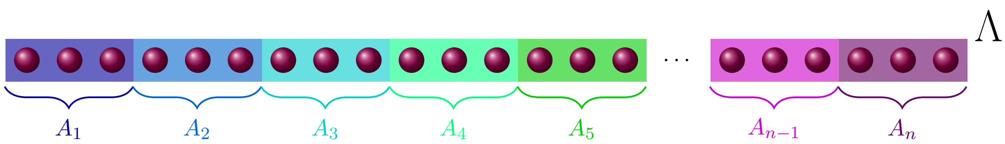


Figure 5: Representation of an interval Λ split into multiple subintervals $\Lambda = A_1 A_2 \dots A_n$.

Their size is chosen so that the inverted BS-CMI is sufficiently small, and will be chosen in the next subsection. For simplicity, let us denote by $\rho_i := \rho_{A_i} = \mathrm{tr}_{\Lambda \setminus A_i}[\rho]$ and by $\rho_{i:j} := \rho_{A_i, A_{i+1}, \dots, A_j} = \mathrm{tr}_{\Lambda \setminus (A_i \cup \dots \cup A_j)}[\rho]$.

A standard argument to extend a recovery result from a single recovery map to a chain is to use the contractivity of the recovery map to show that the errors behave additively. This is not possible in our setting, as the non-trace-preserving maps are not contractive. However, we will be able to overcome this technical difficulty by proving a Lipschitz bound independently of the level of concatenation. More specifically, in Lemma 4.3 we show that the errors only suffer a constant amplification that does not grow exponentially in the length of the recovery chain.

We denote the individual recovery maps $\mathcal{R}_i : \mathcal{B}(\mathcal{H}_{A_i}) \to \mathcal{B}(\mathcal{H}_{A_i A_{i+1}})$, $i = 1, \dots, N-1$ as

$$\mathcal{R}_i(X) = \rho_i^{1/2} (\rho_i^{-1/2} \rho_{i:i+1} \rho_i^{-1/2})^{1/2} \rho_i^{-1/2} X \rho_i^{-1/2} (\rho_i^{-1/2} \rho_{i:i+1} \rho_i^{-1/2})^{1/2} \rho_i^{1/2} \,. \tag{58}$$

For simplicity, we will abbreviate the map when adding on a larger chain, and replace $\mathbb{1}^{\otimes i-1} \otimes \mathcal{R}_i$ by $\mathcal{R}_i$.

Lemma 4.3. *Let $1 \leq j < N$, be natural numbers and $X \in \mathcal{B}(\mathcal{H}_{A_j})$ be positive semidefinite. There is a constant independent of the number of concatenated maps that bounds the Lipschitz constant of the concatenated map:*

$$\left\| \left(\bigcirc_{i=j}^{N-1} \mathcal{R}_i \right)(X) \right\|_1 \leq \left\| \rho_j^{-1/2} X \rho_j^{-1/2} \right\|_1 \leq \|\rho_j^{-1}\| \|X\|_1 .$$

This also holds if the $\rho_i, \rho_{i:i+1}$ are not consistent marginals of a fixed global state, as long as they are local positive states and $\mathrm{tr}_j[\rho_{j:j+1}] = \rho_{j+1}$.

Proof. The map $\bigcirc_{i=j}^{N-1} \mathcal{R}_i(\cdot)$ preserves positivity being a concatenation of positive maps. Hence by the operator inequality $0 \leq X \leq \|\rho_j^{-1/2} X \rho_j^{-1/2}\| \rho_j$, we readily observe that:

$$\left\| \bigcirc_{i=j}^{N-1} \mathcal{R}_i(X) \right\|_1 \leq \left\| \rho_j^{-1/2} X \rho_j^{-1/2} \right\| \left\| \bigcirc_{i=j}^{N-1} \mathcal{R}_i(\rho_j) \right\|_1 .$$

Further inspection reveals that $\mathcal{R}_i(\rho_i) = \rho_{i:i+1}$ and $[\mathrm{tr}_i, \mathcal{R}_k] = 0$ for $k > i$. Leveraging these observations, we can immediately conclude:

$$\begin{aligned} \left\| \bigcirc_{i=j}^{N-1} \mathcal{R}_i(\rho_j) \right\|_1 &= \mathrm{Tr}\left[\bigcirc_{i=j}^{N-1} \mathcal{R}_i(\rho_j) \right] = \mathrm{Tr}\left[\mathrm{tr}_j [\bigcirc_{i=j}^{N-1} \mathcal{R}_i(\rho_j)] \right] \\ &= \mathrm{Tr}\left[\mathrm{tr}_j [\bigcirc_{i=j+1}^{N-1} \mathcal{R}_i(\rho_{j:j+1})] \right] \\ &= \mathrm{Tr}\left[\bigcirc_{i=j+1}^{N-1} \mathcal{R}_i(\mathrm{tr}_j[\rho_{j:j+1}]) \right] = \mathrm{Tr}\left[\bigcirc_{i=j+1}^{N-1} \mathcal{R}_i(\rho_{j+1}) \right] . \end{aligned}$$

Iterating this process yields $\left\| \bigcirc_{i=j}^{N-1} \mathcal{R}_i(\rho_j) \right\|_1 = \mathrm{Tr}[\rho_N] = 1$. Applying Hölder's inequality and the hierarchy of Schatten p-norms to simplify $\left\| \rho_j^{-1/2} X \rho_j^{-1/2} \right\|$ finally confirms the desired result. □

We now combine Lemma 4.3 with Lemma 4.1 to decompose the overall recovery error, obtaining a bound on the recovery error based on entropic quantities. This is the main technical result of the section.

Theorem 4.4. *For a multipartite quantum state $\rho_{1:N}$, the recovery error on the chain between the concatenated recovery map and the original state is bounded by*

$$\begin{aligned} &\left\| \left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right)(\rho_1) - \rho_{1:N} \right\|_1 \\ &\quad \leq \frac{16(N-1)}{\pi} \sup_i \left\| \rho_i^{-1} \right\| \exp(I_\infty(A_1 \dots A_{i-1} : A_i)/2) \widehat{I}_\rho^{\mathrm{rev}}(A_i : A_1 \dots A_{i-2} | A_{i-1})^{1/4} . \end{aligned} \tag{59}$$

Note that the recovery map representation $\left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right)(\rho_1)$ is also a matrix product operator representation with bond dimension $D = \dim(A_i)^3$.

Proof. We denote the overall recovery error by ε

$$\varepsilon = \|\mathcal{R}_{N-1}(\dots \mathcal{R}_1(\rho_1) \dots) - \rho_{1:N}\|_1 = \left\| \left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right)(\rho_1) - \rho_{1:N} \right\|_1 . \tag{60}$$

First, we split the total error ε into the individual errors for each recovery step, where every error is still amplified by the subsequent recovery maps:

$$\begin{aligned} \varepsilon &\leq \sum_{i=1}^{N-1} \left\| \left(\bigcirc_{j=i+1}^{N-1} \mathcal{R}_j \right)(\rho_{1\dots i+1} - \mathcal{R}_i(\rho_{1\dots i})) \right\|_1 \\ &\leq \sum_{i=1}^{N-1} \left\| \left(\bigcirc_{j=i+1}^{N-1} \mathcal{R}_j \right)(X^+_{A_{i+1}}) \right\|_1 + \left\| \left(\bigcirc_{j=i+1}^{N-1} \mathcal{R}_j \right)(X^-_{A_{i+1}}) \right\|_1 \end{aligned} \tag{61}$$

Here, $X_i = \rho_{1:i+1} - \mathcal{R}_i(\rho_{A_1 \ldots A_i})$ with decomposition into positive and negative parts $X_i = X_i^+ - X_i^-$ and $X_{A_i}^{\pm} = \mathrm{Tr}_{A_1 \ldots A_{i-1}}[X_i^{\pm}]$ such that the last norm is taken on systems $A_i \ldots A_N$ only. Furthermore, this means that

$$\|X_{A_i}^{\pm}\|_1 = \|X_i^{\pm}\|_1 \leq \|X_{A_i}\|_1.$$

Finally, combining Eqs. (60) and (61), and using Lemma 4.3 we can upper bound the error as

$$\begin{aligned}\varepsilon &= \left\| \left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right) (\rho_1) - \rho_{1:N} \right\|_1 \\ &\leq 2(N-1) \sup_i \|\rho_i^{-1}\| \frac{8}{\pi} \sqrt{\|\Gamma_i\|} \widehat{I}_\rho^{\mathrm{rev}}(A_i : A_1 \ldots A_{i-2} | A_{i-1})^{1/4} \,,\end{aligned} \tag{62}$$

where in the last inequality we are using that, by Lemma 4.1,

$$\|X_{A_i}\|_1 \leq \frac{8}{\pi} \sqrt{\|\Gamma_i\|} \widehat{I}_\rho^{\mathrm{rev}}(A_i : A_1 \ldots A_{i-2} | A_{i-1})^{1/4} \,,$$

and we introduced $\Gamma_i = \rho_{1:i-1}^{-1/2} \rho_{1:i} \rho_{1:i-1}^{-1/2}$. To obtain the final form of the result, let us note that all the quantities involved can be rewritten in terms of divergences. The factor Γ_i is related to the maximal Rényi mutual information as follows:

$$\begin{aligned}\|\Gamma_i\| &\leq \left\| \rho_i^{1/2} \right\|^2 \left\| (\rho_{1:i-1} \otimes \rho_i)^{-1/2} \rho_{1:i} (\rho_{1:i-1} \otimes \rho_i)^{-1/2} \right\| \\ &\leq \exp(I_\infty(A_1 \ldots A_{i-1} : A_i)) \,.\end{aligned}$$

This concludes the proof. □

We now consider the implications of Theorem 4.4 for MPO representations of Gibbs states. Note that due to the expected exponential growth of $\left\|\rho_i^{-1}\right\|$ in Eq. (59) with the size of the subsystem, at least an exponential decay of $\widehat{I}_\rho^{\mathrm{rev}}(A\,; C|B)$ with sufficient rate, for any adjacent three systems A, B, C, is needed to employ the above result. In the case of thermal states, however, Theorem 3.6 guarantees a superexponential rate, which dominates over the other prefactors.

Corollary 4.5. *For an n-site marginal of a Gibbs state in one dimension with translation-invariant interaction Φ and a given accuracy ε, there is an MPO representation of bond dimension*

$$D = \exp\left(2 \log(d) \, \mathcal{C}_1 \, \frac{\log(n/\varepsilon) + \mathcal{C}_2}{\log(\log(n/\varepsilon))} \right), \tag{63}$$

where $\mathcal{C}_1, \mathcal{C}_2$ are constants depending on J, R, and β.

The MPO representation is given by the above construction, i.e.,

$$\left\| \left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right) (\rho_1) - \rho_{1:N} \right\| \leq \varepsilon$$

choosing the A_i as consecutive regions of at least l spins

$$|A_i| = l \geq \mathcal{C}_1 \, \frac{\log(n/\varepsilon) + \mathcal{C}_2}{\log(\log(n/\varepsilon))} \,.$$

Remark 4.6. *Due to the uniformity of the result in Theorem 3.6, this result applies equally to a local Gibbs state, i.e. the Gibbs state of the Hamiltonian on n sites, as to the marginal of a Gibbs state on a larger or even infinite system.*

Proof of Corollary 4.5. Here, we follow a similar procedure to that of [33, Corollary 3.3]. First, recall that Theorem 3.6 provides the needed decay for the BS-CMI in the case of thermal states. Furthermore, in this setting, by Lemma 3.5, the growth of $\left\|\rho_i^{-1}\right\|$ is exponential in $|A_i|$. Finally, the maximal mutual information obeys an area law for thermal states [71, Theorem 2], namely

$$I_\infty(A:B) \leq \mathcal{C}$$

for any adjacent intervals A, B and a constant $\mathcal{C}$ that only depends on β, J and R.

We now proceed by partitioning our system into subsystems of size $|A_i| = l$ to be determined later. Combining all these results we achieve the desired accuracy if

$$\varepsilon^4 \leq \mathcal{C}_1 \, \frac{n}{l} \frac{\mathcal{C}_2^{\lfloor l/2 \rfloor + 1}}{(\lfloor \lfloor l/2 \rfloor / R \rfloor + 1)!} \, .$$

Note that the factor $\exp(\alpha(|A_i| + |A_{i+1}|))$ for every i is included in $\mathcal{C}_2^{\lfloor l/2 \rfloor + 1}$. Using Stirling's approximation we find that a sufficient condition for this is given by

$$\log\left(\frac{n\,\mathcal{C}_1}{\varepsilon^4}\right) \leq -\left(\frac{l}{2}+1\right)\log(\mathcal{C}_2) + \left(\frac{l}{2R} - \frac{1}{R}\right)\left(\log\left(\frac{l}{2R} - \frac{1}{R}\right) - 1\right) \, .$$

While this inequality cannot be inverted analytically, a further relaxation based on a dual description of the convex function $x \mapsto x\log(x)$ yields the sufficient condition

$$2R \frac{\log\left(\frac{n\,\mathcal{C}_1 \mathcal{C}_2^2}{\varepsilon^4}\right) + a}{\log(a) - R\log(\mathcal{C}_2)} + 1 \leq l$$

for any $a > 0$. Choosing $a = \log(n/\varepsilon)\mathcal{C}_2^R$ and again combining all constants the claim follows. □

Remark 4.7. *Corollary 4.5 can be tailored to accommodate exponentially-decaying interactions, thereby yielding an MPO approximation for a Gibbs state in this context, which to the best of our knowledge is the first of its kind. The bond dimension is slightly worse in that case though, as it arises from the decay of the BS-CMI, which is only exponential in $|B|$ in this case, as opposed to the superexponential behaviour in the finite-range case. This is the content of Corollary B.8.*

4.2 Reconstructing the positive MPO from local tomography

Since Theorem 4.4 provides a way to reconstruct an entire state ρ from the smaller marginals, it also offers a reconstruction of classical representations of ρ, in the form of a MPO, from tomographic data on small regions. What is required from the tomographic data is to produce a classical representation of the marginals with small tomography errors, which can be done with standard results [39, 63]. To prove that our final MPO reconstruction of ρ is accurate, we thus need to show that the reconstruction process, which happens at the level of classical post-processing, is stable to the small errors in the local tomography. Since in our setting the Hamiltonian is translation invariant, the cost of learning a state from measurements is limited to the cost of learning the state in a single region (although the precision might depend on the final size of the reconstruction).

Our starting point is a measurement scheme for general quantum states that we will apply to the marginals of the Gibbs state. Since we do not aim to give constants in our runtime explicitly, any polynomial time measurement scheme is sufficient for our purposes. Sample optimal measurement schemes for arbitrary quantum states with quadratic sample complexity in the inverse error and dimension have been developed in [39, 63]. The situation regarding the classical runtime of these schemes is however unclear. Nevertheless, the following result is an immediate consequence of measuring each coefficient of the density matrix individually to sufficient precision and using the equivalence of matrix norms up to dimensional factors.

Lemma 4.8 (consequence of [39, 63])**.** *There exists a measurement scheme that produces a classical representation of a quantum state $\hat{\rho}$ such that with probability $\geq 1 - c$*

$$\|\rho - \hat{\rho}\|_1 \leq \delta \,, \tag{64}$$

with several samples and classical runtime bounded by

$$n_s \leq C_s \operatorname{poly}\left(d, \frac{1}{\delta}\right) \log(1/c)\,,$$
$$t \leq C_t \operatorname{poly}\left(d, \frac{1}{\delta}\right) \log(1/c)\,.$$

With this, we can now state the cost of local tomography that we require to produce the accurate MPO representation of the whole state from the previous section. The proof essentially shows how the errors in the marginals from Lemma 4.8 propagate in the final error of the MPO approximation, which allows us to estimate the final costs of samples and classical runtime.

Theorem 4.9. *Let us consider a n-site marginal of a Gibbs state in one dimension with translation-invariant interaction, and define the approximate reconstruction map as in Eq.* (58)*,*

$$\hat{\mathcal{R}}_i(X) = (\hat{\rho}_i)^{1/2}((\hat{\rho}_i)^{-1/2}\hat{\rho}_{i:i+1}(\hat{\rho}_i)^{-1/2})^{1/2}(\hat{\rho}_i)^{-1/2}X(\hat{\rho}_i)^{-1/2}((\hat{\rho}_i)^{-1/2}\hat{\rho}_{i:i+1}(\hat{\rho}_i)^{-1/2})^{1/2}(\hat{\rho}_i)^{1/2}\,, \tag{65}$$

where each $\hat{\rho}_i$ is an approximation of the corresponding marginal of ρ. Then, for a given accuracy ε, the MPO representation $\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1)$ has bond dimension as in Eq. (63) *and is such that, with probability $\geq 1 - c$,*

$$\left\|\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1) - \rho_{1:N}\right\|_1 \leq \varepsilon\,.$$

The sample complexity and classical post-processing time for finding this MPO are $\operatorname{poly}(n/\varepsilon)\log(1/c)$.

Proof. We assume the repartitioned subsystems $A_1, \dots, A_N$ each of size $|A_i| = l$ as of Corollary 4.5. We need to show that the accuracy of the reconstruction map $\mathcal{R}_i(X)$ is robust to small errors in the estimated marginals $\rho_i, \rho_{i:i+1}$. We repeatedly apply those approximate reconstructions, and the error induced by each is shown in Eq. (62).

To construct a good approximation to $\left(\bigcirc_{i=1}^{N-1}\mathcal{R}_i\right)(\rho_1)$ we require estimates of the marginals $\hat{\rho}_i$ and $\hat{\rho}_{i:i+1}$, which we take to be such that with probability $\geq 1 - c$

$$\|\rho_{i:i+1} - \hat{\rho}_{i:i+1}\|_1 \leq \delta, \tag{66}$$

for which we need $\operatorname{poly}\left(d^l, \delta^{-1}\right)\log(N/c)$ samples. Then, for consistency, we take the single-region marginals as the reductions $\hat{\rho}_{i+1} = \operatorname{tr}_i[\hat{\rho}_{i:i+1}]$. To account for the possibility of failure in the tomography scheme, we choose a confidence bound of c/N for each marginal such that, by a union bound, all marginals fulfil Eq. (66).

To obtain an accurate approximation to the recovery map, we need our estimates to have well-behaved inverses. Let us note that according to Lemma 3.5

$$\left\|\rho_{i:i+1}^{-1}\right\| \leq \mathcal{C}\, e^{\alpha l}\,.$$

Thereby, explicitly requiring

$$\delta \leq \frac{1}{2\mathcal{C}} e^{-\alpha l} \leq \frac{1}{2} \left\|\rho_{i:i+1}^{-1}\right\|^{-1} , \tag{67}$$

we obtain the equivalent bounds for the approximate marginals

$$\left\|\rho_{i:i+1}^{-1}\right\| , \left\|\rho_i^{-1}\right\| , \left\|(\hat{\rho})_{i:i+1}^{-1}\right\| , \left\|(\hat{\rho})_i^{-1}\right\| \leq 2\,\mathcal{C}\, e^{\alpha l} .$$

With this and the definition of $\hat{\mathcal{R}}_i$ from Eq. (65), our target MPO approximation is $\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1)$. To control the approximation error, we repeatedly apply the triangle inequality so that

$$\left\|\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1) - \left(\bigcirc_{i=1}^{N-1}\mathcal{R}_i\right)(\rho_1)\right\|_1 \leq \sum_{j=1}^{N-1}\left\|\left(\bigcirc_{i=j+1}^{N-1}\mathcal{R}_i\right)\left(\mathcal{R}_j - \hat{\mathcal{R}}_j\right)\left(\bigcirc_{i=1}^{j-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1)\right\|_1 + \left\|\left(\bigcirc_{i=1}^{N-1}\mathcal{R}_i\right)(\hat{\rho}_1 - \rho_1)\right\|_1 .$$

Applying Lemma 4.3 to each term in the sum, and Eq. (66) we have that

$$\left\|\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1) - \left(\bigcirc_{i=1}^{N-1}\mathcal{R}_i\right)(\rho_1)\right\|_1 \leq \sum_{j=1}^{N-1}\left\|\rho_{j+1}^{-1}\right\|\left\|\left(\mathcal{R}_j - \hat{\mathcal{R}}_j\right)\left(\bigcirc_{i=1}^{j-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1)\right\|_1 + 2\delta\left\|\rho_1^{-1}\right\|.$$

To bound the summands $\left\|\left(\mathcal{R}_j - \mathcal{R}'_j\right)(X)\right\|_1$, we require the following two continuity bounds for the square root and its inverse, as proven in [25],

$$\left\|\sqrt{\rho} - \sqrt{\hat{\rho}}\right\| \leq \max\{\left\|\rho^{-1}\right\|, \left\|(\hat{\rho})^{-1}\right\|\} \cdot \sqrt{8}\|\rho - \hat{\rho}\|,$$

$$\left\|\rho^{-1/2} - (\hat{\rho})^{-1/2}\right\| \leq \max\{\left\|\rho^{-1}\right\|, \left\|(\hat{\rho})^{-1}\right\|\} \cdot \left\|\rho^{-1/2}\right\|\left\|(\hat{\rho})^{-1/2}\right\| \cdot \sqrt{8}\|\rho - \hat{\rho}\|.$$

Repeated applications of these and the triangle inequality, and using that $\|\cdot\| \leq \|\cdot\|_1$ then yield

$$\left\|\left(\mathcal{R}_j - \hat{\mathcal{R}}_j\right)(X)\right\|_1 \leq C\delta\|X\|_1\left(\|{\rho_j}^{-1}\|\|(\hat{\rho}_j)^{-1}\|\right)^D$$

for some numerical constants C, D. Again, notice that by Lemma 4.3,

$$\left\|\left(\bigcirc_{i=1}^{j-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1)\right\|_1 \leq 1.$$

Putting everything together, and introducing new numerical constants C', D', we thus obtain

$$\left\|\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1) - \left(\bigcirc_{i=1}^{N-1}\mathcal{R}_i\right)(\rho_1)\right\|_1 \leq C' N\delta d^{\alpha D' l}.$$

This means (recalling Eq. (67)) that if we set

$$\delta = \min\left\{\frac{\varepsilon d^{-\alpha D' l}}{C' N}, \frac{1}{2\mathcal{C}} e^{-\alpha l}\right\}$$

by Corollary 4.5 and the triangle inequality,

$$\left\|\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1) - \rho_{1:N}\right\|_1 \leq 2\varepsilon.$$

Given the choice of δ, the sample complexity is bounded by poly $\left(d^l, N/\varepsilon\right)\log(1/c)$. The time complexity, apart from the contribution due to the tomography itself, consists of elementary matrix functions and multiplications which also run in time poly $\left(d^l, N\right)$. Noting the sublogarithmic dependence of l on n/ε in Corollary 4.5 and that $N \leq n$, the result follows. □

Remark 4.10. *Compared to the previous section which gave a* subpolynomial *dependence in system size and inverse error, we only achieve polynomial dependence here. For the inverse error, this is due to the polynomial dependence in the tomography scheme and cannot be avoided. However, for the system size, the situation is slightly different. Instead of explicitly giving all N channels it can be noted that for* infinite *translation-invariant systems, they are all identical and only measuring one of the marginals and reconstructing one of the channels can be done in subpolynomial time in n. Similarly, for* finite *translation-invariant only the sample complexity can be improved to subpolynomial in n by measuring all subsystems simultaneously for each sample. However, writing out the MPO for the n systems takes linear time in n by definition again in both cases.*

4.3 Efficient estimation of the global purity

Beyond MPO reconstructions of the entire state, the results on the factorization of the purity from Section 3.3 allow us to efficiently estimate the purity of the global state $\rho_{1:N}$ from local purity estimations.

For completeness, we now reproduce a simplified version of the approximation scheme of the purity of 1D Gibbs states in [80], where a more detailed analysis of the polynomial scaling of resources and prefactors can be found. Our main contribution is Proposition 3.9, which constitutes the missing ingredient for a fully rigorous polynomial bound on the sample complexity of their estimation algorithm. Instead, in [80], the exponential decay in Eq. (43) was only studied numerically. To simplify the discussion, we do not consider the randomized measurement toolbox [30], which potentially yields a slightly more favourable scaling and more amenable classical post-processing, and instead use the result from Lemma 4.8.

Theorem 4.11. *Under the conditions of Proposition 3.9, there is an algorithm that outputs an estimate $\hat{P}_2(\rho_{1:N})$ such that*

$$\left|\frac{\hat{P}_2(\rho_{1:N})}{\mathrm{Tr}\left[\rho_{1:N}^2\right]} - 1\right| \leq \varepsilon\,, \tag{68}$$

with a number of samples and classical post-processing cost given by $\mathrm{poly}(n/\varepsilon)$*.*

Proof. Consider adjacent marginals of the state $\rho_{1:N}$ on regions of size l and adjacent pairs $\rho_{i:i+1}$. An estimate for the purity is given by

$$P_2(\rho_{1:N}) = \frac{\prod_{j=1}^{N-1}\mathrm{Tr}_{j:j+1}[\rho_{j:j+1}^2]}{\prod_{j=2}^{N-1}\mathrm{Tr}_j[\rho_j^2]}\,.$$

Then, from an iterated application of Proposition 3.9, and the straightforward estimate for scalars x, y,

$$|xy-1| \leq |x||y-1| + |x-1|\,,$$

we have that (see also Lemma 5 in [80])

$$\left|\frac{P_2(\rho_{1:N})}{\mathrm{Tr}\left[\rho_{1:N}^2\right]} - 1\right| \leq \left(1+\kappa e^{-\alpha l}\right)^N - 1\,.$$

Therefore, with the choice $l = \mathcal{O}\left(\log N/\varepsilon\right)$ we achieve error ε.

The numbers $\mathrm{Tr}_{j:j+1}[\rho_{j:j+1}^2]$ and $\mathrm{Tr}_j[\rho_j^2]$ can be obtained to multiplicative error δ from $\mathrm{poly}(d^l, \delta^{-1})$ local measurements and the scheme in Lemma 4.8, since

$$\frac{\mathrm{Tr}_j[\rho_j^2]}{\mathrm{Tr}_j[\hat{\rho}_j^2]} - 1 \leq \mathrm{Tr}_j[\hat{\rho}_j^2]^{-1}\|\rho_j - \hat{\rho}_j\|_1 \leq d^l\delta\,,$$

and thus our estimation of the global purity is

$$\hat{P}_2(\rho_{1:N}) = \frac{\prod_{j=1}^{N-1} \mathrm{Tr}_{j:j+1}[\hat{\rho}_{j:j+1}^2]}{\prod_{j=2}^{N-1} \mathrm{Tr}_j[\hat{\rho}_j^2]} .$$

With a similar argument as above, this is such that

$$\left| \frac{\hat{P}_2(\rho_{1:N})}{P_2(\rho_{1:N})} - 1 \right| \leq \left(1 + d^{2l}\delta\right)^{2N} - 1 .$$

Therefore, choosing $d^l\delta = e^{-\Omega(l)} = \mathcal{O}(1/\mathrm{poly}(N/\varepsilon))$ we again obtain the approximation error ε. The result follows from the estimates on the sample and computational cost in Lemma 4.8, and noticing that $n \geq N$. □

Notice that it is also possible to estimate the purity via the estimation of the MPO representation in Theorem 4.9, since

$$\left|\mathrm{Tr}\left[\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1)^2\right] - \mathrm{Tr}\left[\rho_{1:N}^2\right]\right| \leq \left\|\left(\bigcirc_{i=1}^{N-1}\hat{\mathcal{R}}_i\right)(\hat{\rho}_1) - \rho_{1:N}\right\|_1 .$$

This, however, is a much worse estimate than Eq. (68), since the purity is typically an exponentially small number. Thus, additive approximations (as opposed to multiplicative ones) are much less meaningful.

Acknowledgements. AC, PG and AR-d-A acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 470903074 - TRR 352. AC, PG and AR-d-A also acknowledge funding by the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State Governments. SOS acknowledges support from the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/W524141/1. AMA acknowledges support from the Spanish Agencia Estatal de Investigación through the grants "IFT Centro de Excelencia Severo Ochoa CEX2020-001007-S" and "Ramón y Cajal RyC2021-031610-I", financed by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. This project was funded within the QuantERA II Programme which has received funding from the EU's H2020 research and innovation programme under the GA No 101017733. AR-d-A acknowledges financial support by the Ministry for Digital Transformation and of Civil Service of the Spanish Government through the QUANTUM ENIA project call – Quantum Spain project, and by the European Union through the Recovery, Transformation and Resilience Plan – NextGenerationEU within the framework of the Digital Spain 2026 Agenda.

Data Availability: Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.
Conflict of interests: The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Á. M. Alhambra and J. I. Cirac. Locally Accurate Tensor Networks for Thermal States and Time Evolution. *PRX Quantum*, 2(4):040331, Nov. 2021. doi:10.1103/PRXQuantum.2.040331. arXiv:2106.00710. 7

[2] A. Anshu, I. Arad, and D. Gosset. An area law for 2d frustration-free spin systems. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 12–18, Rome Italy, June 2022. ACM. doi:10.1145/3519935.3519962. arXiv:2103.02492. 2

[3] A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar. Sample-efficient learning of interacting quantum systems. *Nature Physics*, 17(8):931–935, Aug. 2021. doi:10.1038/s41567-021-01232-0. arXiv:2004.07266. 4, 7

[4] A. Anshu, S. Bab Hadiashar, R. Jain, A. Nayak, and D. Touchette. One-Shot Quantum State Redistribution and Quantum Markov Chains. *IEEE Transactions on Information Theory*, 69(9):5788–5804, Sept. 2023. doi:10.1109/TIT.2023.3271316. arXiv:2104.08753. 3

[5] I. Arad, R. Firanko, and R. Jain. Area law for the maximally mixed ground state in degenerate 1D gapped systems, Oct. 2023. arXiv:2310.19028. 2

[6] H. Araki. Gibbs states of a one dimensional quantum lattice. *Communications in Mathematical Physics*, 14(2):120–157, June 1969. doi:10.1007/BF01645134. 2, 3, 5, 11

[7] E. Bairey, I. Arad, and N. H. Lindner. Learning a Local Hamiltonian from Local Measurements. *Physical Review Letters*, 122(2):020504, Jan. 2019. doi:10.1103/PhysRevLett.122.020504. arXiv:1807.04564. 4

[8] A. Bakshi, A. Liu, A. Moitra, and E. Tang. Learning quantum hamiltonians at any temperature in polynomial time. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 1470–1477, New York, NY, USA, 2024. Association for Computing Machinery. 4, 7

[9] V. P. Belavkin and P. Staszewski. C*-algebraic generalization of relative entropy and entropy. *Annales de l'I.H.P. Physique théorique*, 37(1):51–58, 1982. 3, 4, 9

[10] M. Berta, K. P. Seshadreesan, and M. M. Wilde. Rényi generalizations of the conditional quantum mutual information. *Journal of Mathematical Physics*, 56(2):022205, Feb. 2015. doi:10.1063/1.4908102. arXiv:1403.6102. 3

[11] R. Bhatia. *Matrix analysis.* Number 169 in Graduate texts in mathematics. Springer, New York, 1997. 18

[12] A. Bluhm and Á. Capel. A strengthened data processing inequality for the Belavkin–Staszewski relative entropy. *Reviews in Mathematical Physics*, 32(02):2050005, Mar. 2020. doi:10.1142/S0129055X20500051. arXiv:1904.10768. 4, 5, 14, 16, 23, 24, 40, 45

[13] A. Bluhm, A. Capel, P. Costa Rico, and A. Jencova. Belavkin Staszewski Quantum Markov Chains. *in preparation*, 2024. 23

[14] A. Bluhm, Á. Capel, P. Gondolf, and A. Pérez-Hernández. Continuity of Quantum Entropic Quantities via Almost Convexity. *IEEE Transactions on Information Theory*, 69(9):5869–5901, Sept. 2023. doi:10.1109/TIT.2023.3277892. arXiv:2208.00922. 9, 14

[15] A. Bluhm, Á. Capel, P. Gondolf, and A. Pérez-Hernández. General Continuity Bounds for Quantum Relative Entropies. In *2023 IEEE International Symposium on Information Theory (ISIT)*, pages 162–167, Taipei, Taiwan, June 2023. IEEE. doi:10.1109/ISIT54713.2023.10206734. arXiv:2305.10140. 9

[16] A. Bluhm, Á. Capel, and A. Pérez-Hernández. Exponential decay of mutual information for Gibbs states of local Hamiltonians. *Quantum*, 6:650, Feb. 2022. doi:10.22331/q-2022-02-10-650. arXiv:2104.04419. 2, 3, 4, 5, 9, 11, 12, 13, 14, 20, 38, 39, 42

[17] A. Bluhm, A. Capel, and A. Pérez-Hernández. Strong decay of correlations for Gibbs states in any dimension, Jan. 2024. arXiv:2401.10147. 3, 6, 41, 44

[18] F. G. S. L. Brandão, T. S. Cubitt, A. Lucia, S. Michalakis, and D. Pérez-García. Area law for fixed points of rapidly mixing dissipative quantum systems. *Journal of Mathematical Physics*, 56(10):102202, Oct. 2015. doi:10.1063/1.4932612. arXiv:1505.02776. 2

[19] F. G. S. L. Brandão and M. Horodecki. Exponential Decay of Correlations Implies Area Law. *Communications in Mathematical Physics*, 333(2):761–798, Jan. 2015. doi:10.1007/s00220-014-2213-8. arXiv:1206.2947. 2

[20] F. G. S. L. Brandão and M. J. Kastoryano. Finite Correlation Length Implies Efficient Preparation of Quantum Thermal States. *Communications in Mathematical Physics*, 365(1):1–16, Jan. 2019. doi:10.1007/s00220-018-3150-8. arXiv:1609.07877. 3, 4

[21] W. Brown and D. Poulin. Quantum Markov Networks and Commuting Hamiltonians, June 2012. arXiv:1206.0755. 3

[22] Á. Capel, M. Moscolari, S. Teufel, and T. Wessel. From decay of correlations to locality and stability of the Gibbs state, Feb. 2023. arXiv:2310.09182. 3, 6, 14, 41, 43

[23] E. A. Carlen and A. Vershynina. Recovery and the Data Processing Inequality for Quasi-Entropies. *IEEE Transactions on Information Theory*, 64(10):6929–6938, Oct. 2018. doi:10.1109/TIT.2018.2812038. arXiv:1710.08080. 14

[24] E. A. Carlen and A. Vershynina. Recovery map stability for the data processing inequality. *Journal of Physics A: Mathematical and Theoretical*, 53(3):035204, Jan. 2020. doi:10.1088/1751-8121/ab5ab7. arXiv:1710.02409. 14, 23, 24

[25] A. S. Cavaretta and L. Smithies. Lipschitz-type bounds for the map $A \to |A|$ on $L(H)$. *Linear Algebra and its Applications*, 360:231–235, Feb. 2003. doi:10.1016/S0024-3795(02)00453-6. 29

[26] P. Clifford and J. M. Hammersley. Markov fields on finite graphs and lattices. https://ora.ox.ac.uk/objects/uuid:4ea849da-1511-4578-bb88-6a8d02f457a6, 1971. 3

[27] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu. Efficient quantum state tomography. *Nature Communications*, 1(1):149, Dec. 2010. doi:10.1038/ncomms1147. arXiv:1101.4366. 4

[28] N. Datta. Min- and Max-Relative Entropies and a New Entanglement Monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, June 2009. doi:10.1109/TIT.2009.2018325. arXiv:0803.2770. 9

[29] I. Devetak and J. Yard. Exact Cost of Redistributing Multipartite Quantum States. *Physical Review Letters*, 100(23):230501, June 2008. doi:10.1103/PhysRevLett.100.230501. arXiv:quant-ph/0612050. 3

[30] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller. The randomized measurement toolbox. *Nature Reviews Physics*, 5(1):9–24, Dec. 2022. doi:10.1038/s42254-022-00535-2. arXiv:2203.11374. 30

[31] K. Fang and H. Fawzi. Geometric Rényi Divergence and its Applications in Quantum Channel Capacities. *Communications in Mathematical Physics*, 384(3):1615–1677, June 2021. doi:10.1007/s00220-021-04064-4. arXiv:1909.05758. 9

[32] M. Fanizza, N. Galke, J. Lumbreras, C. Rouzé, and A. Winter. Learning finitely correlated states: stability of the spectral reconstruction, Dec. 2023. arXiv:2312.07516. 4, 7

[33] H. Fawzi, O. Fawzi, and S. O. Scalet. A subpolynomial-time algorithm for the free energy of one-dimensional quantum systems in the thermodynamic limit. *Quantum*, 7:1011, May 2023. doi:10.22331/q-2023-05-22-1011. arXiv:2209.14989. 27

[34] O. Fawzi and R. Renner. Quantum Conditional Mutual Information and Approximate Markov Chains. *Communications in Mathematical Physics*, 340(2):575–611, Dec. 2015. doi:10.1007/s00220-015-2466-x. arXiv:1410.0664. 3, 7, 14

[35] R. Firanko, M. Goldstein, and I. Arad. Area law for steady states of detailed-balance local lindbladians. *Journal of Mathematical Physics*, 65(5), May 2024. 2

[36] J. Fröhlich and D. Ueltschi. Some properties of correlations of quantum lattice systems in thermal equilibrium. *Journal of Mathematical Physics*, 56(5):053302, May 2015. doi:10.1063/1.4921305. arXiv:1412.2534. 2, 3

[37] L. Gao and M. M. Wilde. Recoverability for optimized quantum f-divergences. *Journal of Physics A: Mathematical and Theoretical*, 54(38):385302, Sept. 2021. doi:10.1088/1751-8121/ac1dc2. arXiv:2008.01668. 14

[38] J. Guth Jarkovský, A. Molnár, N. Schuch, and J. I. Cirac. Efficient Description of Many-Body Systems with Matrix Product Density Operators. *PRX Quantum*, 1(1):010304, Sept. 2020. doi:10.1103/PRXQuantum.1.010304. arXiv:2003.12418. 2

[39] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, pages 1–1, 2017. doi:10.1109/TIT.2017.2719044. arXiv:1508.01797. 27, 28

[40] J. Haah, R. Kothari, and E. Tang. Optimal learning of quantum Hamiltonians from high-temperature Gibbs states. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 135–146, Denver, CO, USA, Oct. 2022. IEEE. doi:10.1109/FOCS54457.2022.00020. arXiv:2108.04842. 4, 7

[41] M. B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024–P08024, Aug. 2007. doi:10.1088/1742-5468/2007/08/P08024. arXiv:0705.2024. 2, 7

[42] M. B. Hastings and T. Koma. Spectral Gap and Exponential Decay of Correlations. *Communications in Mathematical Physics*, 265(3):781–804, Aug. 2006. doi:10.1007/s00220-006-0030-4. arXiv:math-ph/0507008. 2

[43] P. Hayden, R. Jozsa, D. Petz, and A. Winter. Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality. *Communications in Mathematical Physics*, 246(2):359–374, Apr. 2004. doi:10.1007/s00220-004-1049-z. arXiv:quant-ph/0304007. 3

[44] F. Hiai and M. Mosonyi. Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023, Aug. 2017. doi:10.1142/S0129055X17500234. arXiv:1604.03089. 9, 14

[45] M. K. Joshi, C. Kokail, R. Van Bijnen, F. Kranzl, T. V. Zache, R. Blatt, C. F. Roos, and P. Zoller. Exploring large-scale entanglement in quantum simulation. *Nature*, 624(7992):539–544, Dec. 2023. doi:10.1038/s41586-023-06768-0. arXiv:2306.00057. 4

[46] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter. Universal Recovery Maps and Approximate Sufficiency of Quantum Relative Entropy. *Annales Henri Poincaré*, 19(10):2955–2978, Oct. 2018. doi:10.1007/s00023-018-0716-0. arXiv:1509.07127. 3, 7

[47] K. Kato and F. G. S. L. Brandão. Quantum Approximate Markov Chains are Thermal. *Communications in Mathematical Physics*, 370(1):117–149, Aug. 2019. doi:10.1007/s00220-019-03485-6. arXiv:1609.06636. 3, 4

[48] A. Kitaev and J. Preskill. Topological Entanglement Entropy. *Physical Review Letters*, 96(11):110404, Mar. 2006. doi:10.1103/PhysRevLett.96.110404. arXiv:hep-th/0510092. 3

[49] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert. Locality of Temperature. *Physical Review X*, 4(3):031019, July 2014. doi:10.1103/PhysRevX.4.031019. arXiv:1309.0816. 2

[50] J. Kochanowski, A. M. Alhambra, A. Capel, and C. Rouzé. Rapid thermalization of dissipative many-body dynamics of commuting hamiltonians, 2024. 3

[51] C. Kokail, R. Van Bijnen, A. Elben, B. Vermersch, and P. Zoller. Entanglement Hamiltonian tomography in quantum simulation. *Nature Physics*, 17(8):936–942, Aug. 2021. doi:10.1038/s41567-021-01260-w. arXiv:2009.09000. 4

[52] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, Mar. 1951. doi:10.1214/aoms/1177729694. 8

[53] T. Kuwahara. Clustering of conditional mutual information and quantum markov structure at arbitrary temperatures, 2024. arXiv:2407.05835. 3, 4

[54] T. Kuwahara, Á. M. Alhambra, and A. Anshu. Improved Thermal Area Law and Quasilinear Time Algorithm for Quantum Gibbs States. *Physical Review X*, 11(1):011047, Mar. 2021. doi:10.1103/PhysRevX.11.011047. arXiv:2007.11174. 2, 7

[55] T. Kuwahara, K. Kato, and F. G. S. L. Brandão. Clustering of Conditional Mutual Information for Quantum Gibbs States above a Threshold Temperature. *Physical Review Letters*, 124(22):220601, June 2020. doi:10.1103/PhysRevLett.124.220601. arXiv:1910.09425. 4

[56] T. Kuwahara and K. Saito. Exponential Clustering of Bipartite Quantum Entanglement at Arbitrary Temperatures. *Physical Review X*, 12(2):021022, Apr. 2022. doi:10.1103/PhysRevX.12.021022. arXiv:2108.12209. 3

[57] Z. Landau, U. Vazirani, and T. Vidick. A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians. *Nature Physics*, 11(7):566–569, July 2015. doi:10.1038/nphys3345. arXiv:1307.5143. 2, 7

[58] M. Leifer and D. Poulin. Quantum Graphical Models and Belief Propagation. *Annals of Physics*, 323(8):1899–1946, Aug. 2008. doi:10.1016/j.aop.2007.10.001. arXiv:0708.1337. 3

[59] M. Lenci and L. Rey-Bellet. Large Deviations in Quantum Lattice Systems: One-Phase Region. *Journal of Statistical Physics*, 119(3):715–746, May 2005. doi:10.1007/s10955-005-3015-3. arXiv:math-ph/0406065. 16

[60] E. H. Lieb and M. B. Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14(12):1938–1941, Dec. 1973. doi:10.1063/1.1666274. 3

[61] K. Matsumoto. A New Quantum Version of f-Divergence. In M. Ozawa, J. Butterfield, H. Halvorson, M. Rédei, Y. Kitajima, and F. Buscemi, editors, *Reality and Measurement in Algebraic Quantum Theory*, volume 261, pages 229–273. Springer Singapore, Singapore, 2018. doi:10.1007/978-981-13-2487-1_10. arXiv:1311.4722. 9

[62] A. Molnár, N. Schuch, F. Verstraete, and J. I. Cirac. Approximating Gibbs states of local Hamiltonians efficiently with projected entangled pair states. *Physical Review B*, 91(4):045138, Jan. 2015. doi:10.1103/PhysRevB.91.045138. arXiv:1406.2973. 7

[63] R. O'Donnell and J. Wright. Efficient quantum tomography. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 899–912, Cambridge MA USA, June 2016. ACM. doi:10.1145/2897518.2897544. arXiv:1508.01907. 27, 28

[64] D. Pérez-García and A. Pérez-Hernández. Locality Estimates for Complex Time Evolution in 1D. *Communications in Mathematical Physics*, 399(2):929–970, Apr. 2023. doi:10.1007/s00220-022-04573-w. arXiv:2004.10516. 6, 11, 14, 41, 43

[65] D. Petz. Sufficiency of channels over von Neumann algebras. *The Quarterly Journal of Mathematics*, 39(1):97–108, 1988. doi:10.1093/qmath/39.1.97. 14

[66] D. Petz. Monotonicity of quantum relative entropy revisited. *Reviews in Mathematical Physics*, 15(01):79–91, Mar. 2003. doi:10.1142/S0129055X03001576. arXiv:quant-ph/0209053. 14

[67] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer. Entropy, Entanglement, and Area: Analytical Results for Harmonic Lattice Systems. *Physical Review Letters*, 94(6):060503, Feb. 2005. doi:10.1103/PhysRevLett.94.060503. arXiv:quant-ph/0405142. 2

[68] Z. Qin, C. Jameson, Z. Gong, M. B. Wakin, and Z. Zhu. Quantum state tomography for matrix product density operators. *IEEE Transactions on Information Theory*, 70(7):5030–5056, July 2024. 7

[69] C. Rouzé and D. Stilck França. Learning quantum many-body systems from a few copies. *Quantum*, 8:1319, Apr. 2024. 4

[70] C. Rouzé, D. Stilck França, E. Onorati, and J. D. Watson. Efficient learning of ground and thermal states within phases of matter. *Nature Communications*, 15(1), Sept. 2024. 3, 4

[71] S. O. Scalet, Á. M. Alhambra, G. Styliaris, and J. I. Cirac. Computable Rényi mutual information: Area laws and correlations. *Quantum*, 5:541, Sept. 2021. doi:10.22331/q-2021-09-14-541. arXiv:2103.01709. 2, 6, 9, 27

[72] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac. Entropy Scaling and Simulability by Matrix Product States. *Physical Review Letters*, 100(3):030504, Jan. 2008. doi:10.1103/PhysRevLett.100.030504. arXiv:0705.0292. 7

[73] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, July 1948. doi:10.1002/j.1538-7305.1948.tb01338.x. 8

[74] B. Shi, K. Kato, and I. H. Kim. Fusion rules from entanglement. *Annals of Physics*, 418:168164, July 2020. doi:10.1016/j.aop.2020.168164. arXiv:1906.09376. 3

[75] D. Sutter. Approximate Quantum Markov Chains. In *Approximate Quantum Markov Chains*, volume 28, pages 75–100. Springer International Publishing, Cham, 2018. doi:10.1007/978-3-319-78732-9_5. arXiv:1802.05477. 3

[76] D. Sutter, M. Berta, and M. Tomamichel. Multivariate Trace Inequalities. *Communications in Mathematical Physics*, 352(1):37–58, May 2017. doi:10.1007/s00220-016-2778-5. arXiv:1604.03023. 7, 14

[77] P. Svetlichnyy and T. A. B. Kennedy. Decay of quantum conditional mutual information for purely generated finitely correlated states. *Journal of Mathematical Physics*, 63(7):072201, July 2022. doi:10.1063/5.0085358. arXiv:2210.09387. 3

[78] M. Tomamichel. *Quantum Information Processing with Finite Resources*, volume 5 of *SpringerBriefs in Mathematical Physics*. Springer International Publishing, Cham, 2016. doi:10.1007/978-3-319-21891-5. arXiv:1504.00233. 15

[79] H. Umegaki. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Mathematical Journal*, 14(2), Jan. 1962. doi:10.2996/kmj/1138844604. 8

[80] B. Vermersch, M. Ljubotina, J. I. Cirac, P. Zoller, M. Serbyn, and L. Piroli. Many-body entropies and entanglement from polynomially many local measurements. *Phys. Rev. X*, 14:031035, Aug 2024. doi:10.1103/PhysRevX.14.031035. arXiv:2311.08108. 3, 5, 19, 30, 44

[81] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. *Physical Review B*, 73(9):094423, Mar. 2006. doi:10.1103/PhysRevB.73.094423. arXiv:cond-mat/0505140. 2, 7

[82] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems. *Physical Review Letters*, 93(20):207204, Nov. 2004. doi:10.1103/PhysRevLett.93.207204. arXiv:cond-mat/0406426. 7

[83] M. M. Wilde. Recoverability for holevo's just-as-good fidelity. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 2331–2335, 2018. 45

[84] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac. Area Laws in Quantum Systems: Mutual Information and Correlations. *Physical Review Letters*, 100(7):070502, Feb. 2008. doi:10.1103/PhysRevLett.100.070502. arXiv:0704.3906. 2

[85] Y. Zhai, B. Yang, and Z. Xi. Belavkin–Staszewski Relative Entropy, Conditional Entropy, and Mutual Information. *Entropy*, 24(6):837, June 2022. doi:10.3390/e24060837. 9

A Proofs of approximate factorization of Gibbs states in 1D

This appendix is devoted to the proofs of the new results appearing in Section 2.4, in connection to the approximate factorisation of Gibbs states in 1D.

A.1 Proof of Lemma 2.4

Let us recall that we want to prove the following inequality:

$$\left\|\mathrm{tr}_{AC}[\rho^A \otimes \rho^C E_{AB,C}\, E_{A,B}]^{-1}\right\| \leq \mathcal{C}\,. \tag{69}$$

Before providing its proof, as a technical tool we introduce the following measure of locality

$$\|T\|_I := \inf\{\|T-P\| : P \in \mathcal{A}_I\}$$

and the following Lemma, whose proof can be found in [16].

Lemma A.1. *[16, Lemma 4.1] Let T, T' be two observables. Then*

$$\|TT'\|_I \leq 2(\|T\|_I\|T'\| + \|T\|\|T'\|_I)$$

and if T is positive definite

$$\|T^{-1}\|_I \leq 2\|T^{-1}\|^2\|T\|_I.$$

For T an observable on a bipartite system AB, ρ_A any state on A, $I \subset AB$, and $I' = I \cap B$, we also have

$$\|\,\mathrm{tr}_A[\rho_A T]\|_{I'} \leq \|T\|_I$$

Proof. The proof of the last point has been omitted in the reference, but follows straightforwardly from the contractiveness of the operator norm. By compactness, let $P \in \mathcal{A}_I$ be an operator such that $\|T\|_I = \|T-P\|$. Then,

$$\|\,\mathrm{tr}_A[\rho_A T]\|_{I'} \leq \|\,\mathrm{tr}_A[\rho_A T] - \mathrm{tr}_A[\rho_A P]\| \leq \|T-P\| = \|T\|_I.$$

□

Proof of Lemma 2.4. The proof of this lemma closely follows the arguments of the third point of [16, Corollary 4.4], however, due to subtle differences in the statement we reprove it here. The original strategy used that the operator Q is almost local in a small interval (after a partial trace and identification of neighbouring intervals), a preliminary that we replace with Q being approximately supported on two small *but distant* intervals. First, note that

$$\begin{aligned}\mathrm{tr}_{AC}[\rho^A \otimes \rho^C Q] &= \mathrm{tr}_{AC}[\rho^A \otimes \rho^C e^{-H_{ABC}} e^{H_A+H_B+H_C}] \\ &= e^{\frac{1}{2}H_B}\,\mathrm{tr}_{AC}[\rho^A \otimes \rho^C F_{A,B}F_{AB,C}F^*_{AB,C}F^*_{A,B}]e^{-\frac{1}{2}H_B} \\ &= e^{\frac{1}{2}H_B}\,\mathrm{tr}_{AC}[\rho^A \otimes \rho^C F]e^{-\frac{1}{2}H_B}\,,\end{aligned}$$

with $F_{X,Y} = e^{-\frac{1}{2}H_{XY}}e^{\frac{1}{2}(H_X+H_Y)}$ and $F = F_{A,B}F_{AB,C}F^*_{AB,C}F^*_{A,B} \geq 0$. By Proposition 2.1, we know that each factor $F_{X,Y}, F^{-1}_{X,Y}$ is bounded in norm by $\mathcal{G}$ (for $\Phi/2$), and thus $\|F\|, \left\|F^{-1}\right\| \leq \mathcal{G}^4$.

Let us define the intervals I_n as neighbourhoods of the cuts AB and BC and I'_n as I_n reduced to B. More precisely if $l.$ and $u.$ denote the lower and upper boundaries of an interval, we define $I_n = \{u_A-n, \ldots, l_B+n\} \cup \{u_B-n, \ldots, l_C+n\}$. Our objective now is to minimize the impact of F by approximating it solely in terms of local operators on these sets. Subsequently, we will demonstrate that by performing a partial trace, these operators are further reduced to act non-trivially solely on $I'_n = B \cap I_n$.

Using Lemma A.1 and Proposition 2.1 we can derive

$$\begin{aligned} \|F\|_{I_n} &= \left\|F_{A,B}F_{AB,C}F^*_{AB,C}F^*_{A,B}\right\|_{I_n} \\ &\leq 4\,\mathcal{G}^3\left(\|F_{A,B}\|_{I_n} + \|F_{AB,C}\|_{I_n} + \|F^*_{AB,C}\|_{I_n} + \|F^*_{A,B}\|_{I_n}\right) \\ &\leq 16\,\mathcal{G}^3\,\frac{\mathcal{G}^n}{(\lfloor n/R\rfloor+1)!}. \end{aligned}$$

Using Lemma A.1 and the fact that the map $\mathrm{tr}_{AC}[\rho^A \otimes \rho^C \cdot]$ is positive unital we have

$$\begin{aligned} \|F'\|_{I'_n} &\leq \|F\|_{I_n} \leq 16\,\mathcal{G}^3\,\frac{\mathcal{G}^n}{(\lfloor n/R\rfloor+1)!} \\ \|F'^{-1}\| &\leq \|F^{-1}\| \leq \mathcal{G}^4, \end{aligned}$$

where $F' = \mathrm{tr}_{AC}[\rho^A \otimes \rho^C F]$. Employing Lemma A.1 again

$$\|F'^{-1}\|_{I'_n} \leq 2\|F'^{-1}\|^2\|F'\|_{I'_n} \leq 32\,\mathcal{G}^{11}\,\frac{\mathcal{G}^n}{(\lfloor n/R\rfloor+1)!}.$$

The next step is completely analogous to [16, Lemma 4.2] with the only deviation being that [16, Proposition 3.2], involving uniform estimates for the complex time evolution of any observable, does not apply directly.

Let us denote by P_n a collection of I'_n-local operators fulfilling $\|F'^{-1} - P_n\| = \|F'^{-1}\|_{I'_n}$ and $S_1 = P_1$, $S_n = P_n - P_{n-1}$. This implies the bounds

$$\|S_1\| \leq \|F'^{-1}\|_{I'_1} + \|F'^{-1}\| \qquad \text{and} \qquad \|S_n\| \leq \|F'^{-1}\|_{I'_n} + \|F'^{-1}\|_{I'_{n-1}}.$$

Note that for $n \geq \lfloor |B|/2 \rfloor + 1$, $P_n = F'^{-1}$ and thereby $S_{n+1} = 0$. We start with a decomposition

$$F'^{-1} = \sum_{n=1}^{\lfloor |B|/2\rfloor+1} S_n\,,$$

where we stop the local decomposition once the two distinct intervals start overlapping. Our goal is to bound

$$e^{\frac{1}{2}H_B}F'^{-1}e^{-\frac{1}{2}H_B}$$

by individually bounding the norms of the imaginary time-evolved S_n. To this end, let us note that there exists an orthogonal decomposition for all $n \leq \lfloor |B|/2 \rfloor$

$$S_n = \sum_{i=1}^{d^{4n}} S^l_{n,i} \otimes S^r_{n,i}\,.$$

Such a decomposition further possesses the property that $\|S^l_{n,i} \otimes S^r_{n,i}\| \leq d^{2n}\|S_n\|$, which is derived from the orthogonality of the $S^l_{n,i} \otimes S^r_{n,i}$ and the inequality

$$\|\cdot\| \leq \|\cdot\|_2 \leq \sqrt{d'}\|\cdot\|\,.$$

Here, d' is the dimension of the underlying Hilbert space. We can now use that the time evolution is a group automorphism on each term individually and that we have [12, Proposition 3.2] allowing us to estimate

$$\begin{aligned}\left\|e^{\frac{1}{2}H_B} S^l_{n,i} \otimes S^r_{n,i} e^{-\frac{1}{2}H_B}\right\| &\leq \left\|e^{\frac{1}{2}H_B} S^l_{n,i} e^{-\frac{1}{2}H_B}\right\| \left\|e^{\frac{1}{2}H_B} S^r_{n,i} e^{-\frac{1}{2}H_B}\right\| \\ &\leq \mathcal{G}^{2n}\, d^{2n} \|S_n\|\end{aligned}$$

At last for $n = \lfloor |B|/2 \rfloor + 1$, we simply regard R_n as an B-local operator, so the directly applying [12, Proposition 3.2] yields

$$\left\|e^{\frac{1}{2}H_B} S_n e^{-\frac{1}{2}H_B}\right\| \leq \mathcal{G}^{|B|}\, \|S_n\| \, .$$

Putting everything together, we have

$$\left\|e^{\frac{1}{2}H_B} F'^{-1} e^{-\frac{1}{2}H_B}\right\| \leq 32 d^6\, \mathcal{C}^2 (\mathcal{G}^{12} + \mathcal{G}^4) + \left(32\, \mathcal{G}^{11} \sum_{n=2}^{\lfloor |B/2| \rfloor + 1} \mathcal{G}^{2n}\, d^{6n} \frac{\mathcal{G}^n}{(\lfloor n/R \rfloor + 1)!}\right)$$

and the statement follows by combining constants. □

A.2 Proof of Lemma 2.5

Next, we prove that, for a Gibbs state $\rho^{A'ABCC'}$ on a finite chain $A'ABCC'$, the following holds.

$$\left\|\rho_{AB}\, \rho_B^{-1} \rho_{BC}\, \rho_{ABC}^{-1} - \mathbb{1}\right\|_\infty < \mathcal{C}^4\, \left\|\rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1} - \mathbb{1}\right\|_\infty \, , \tag{70}$$

Proof. We absorb β in the interaction, so the argument will be made for $\beta = 1$. Let us rewrite

$$\begin{aligned}&\rho_{AB}\, \rho_B^{-1} \rho_{BC}\, \rho_{ABC}^{-1} \\ &\quad = \mathrm{tr}_{A'C'}\left[\rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\right]\, \mathrm{tr}_{A'C'}\left[\rho_{A'ABCC'}\right]^{-1} \\ &\quad = \mathrm{tr}_{A'C'}\left[\rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1}\, \rho_{A'ABCC'}\right]\, \mathrm{tr}_{A'C'}\left[\rho_{A'ABCC'}\right]^{-1} \\ &\quad = \mathrm{tr}_{A'C'}\left[\rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1}\, e^{-H_{A'ABCC'}}\right]\, \mathrm{tr}_{A'C'}\left[e^{-H_{A'ABCC'}}\right]^{-1} \\ &\quad = \mathrm{tr}_{A'C'}\left[\rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1}\, e^{-H_{A'ABCC'}}\, e^{H_{ABC}}\right]\, \mathrm{tr}_{A'C'}\left[e^{-H_{A'ABCC'}}\, e^{H_{ABC}}\right]^{-1} \, .\end{aligned}$$

Recalling that

$$E_{A'ABC,C'} := e^{-H_{A'ABCC'}}\, e^{H_{A'ABC} + H_{C'}} \, , \quad E_{A',ABC} := e^{-H_{A'ABC}}\, e^{H_{A'} + H_{ABC}} \, ,$$

then, from the above identity, we deduce that

$$\begin{aligned}&\rho_{AB}\, \rho_B^{-1} \rho_{BC}\, \rho_{ABC}^{-1} \\ &= \mathrm{tr}_{A'C'}\left[\rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1}\, E_{A'ABC,C'}\, E_{A',ABC}\, e^{-H_{A'} - H_{C'}}\right]\, \mathrm{tr}_{A'C'}\left[E_{A'ABC,C'}\, E_{A',ABC}\, e^{-H_{A'} - H_{C'}}\right]^{-1} \\ &= \mathrm{tr}_{A'C'}\left[\rho^{A'} \otimes \rho^{C'}\, \rho_{A'AB}\, \rho_B^{-1} \rho_{BCC'}\, \rho_{A'ABCC'}^{-1}\, E_{A'ABC,C'}\, E_{A',ABC}\right]\, \mathrm{tr}_{A'C'}\left[\rho^{A'} \otimes \rho^{C'}\, E_{A'ABC,C'}\, E_{A',ABC}\right]^{-1}\end{aligned}$$

Thus

$$\begin{aligned}&\rho_{AB}\, \rho_B^{-1} \rho_{BC}\, \rho_{ABC}^{-1} - \mathbb{1} \\ &\quad = \mathrm{tr}_{A'C'}\left[\rho^{A'} \otimes \rho^{C'}\, Q\, E_{A'ABC,C'}\, E_{A',ABC}\right]\, \mathrm{tr}_{A'C'}\left[\rho^{A'} \otimes \rho^{C'}\, E_{A'ABC,C'}\, E_{A',ABC}\right]^{-1} ,\end{aligned}$$

where

$$Q := \rho_{A'AB}\,\rho_B^{-1}\rho_{BCC'}\,\rho_{A'ABCC'}^{-1} - \mathbb{1}\ .$$

By Equation (27), we have

$$\left\|\mathrm{tr}_{A'C'}\left[\rho^{A'}\otimes\rho^{C'}\,Q\,E_{A'ABC,C'}\,E_{A',ABC}\right]\right\| \leq \left\|E_{A'ABC,C'}\right\|\left\|E_{A',ABC}\right\|\|Q\| \leq \mathcal{C}^2\|Q\|\,,$$

and for the term with the inverse we have an upper bound of $\mathcal{C}^2$ by Lemma 2.4. Therefore,

$$\left\|\rho_{AB}\,\rho_B^{-1}\rho_{BC}\,\rho_{ABC}^{-1} - \mathbb{1}\right\| \leq \mathcal{C}^4\,\|Q\|\,.$$

□

B Extension of results to exponentially-decaying interactions

In this appendix, we extend the main results of this manuscript to the framework of exponentially-decaying interactions. All results presented in Section 2.4 admit extensions to this setting, albeit with some small modifications in the estimates. Here, we collect the generalisations of all these technical results in the context of exponentially-decaying interactions and either refer to a source for proof or give the proof here. Next, we use them to generalise the main results of this manuscript (Theorem 3.6, Theorem 4.11 and Theorem 4.9) to exponentially-decaying interactions. To increase readability, we will only highlight those parts of the proofs that differ from the finite-range counterpart.

Consider an interaction on $\Sigma \subseteq \mathbb{Z}$. To define exponentially-decaying interactions, we introduce for each $\lambda > 0$ the following notation

$$\Omega_n := \sup_{x\in\Sigma}\sum\{\|\Phi(X)\| : X \ni x, \mathrm{diam}(X) \geq n\}$$
$$\|\Phi\|_\lambda := \sum_{n\geq 0}\Omega_n e^{\lambda n} \in [0,\infty]\,.$$

We say that Φ is *exponentially decaying*, if there exists $\lambda > 0$ such that $\|\Phi\|_\lambda < \infty$. Note that there is a subtle difference in the definition of $\|\Phi\|_\lambda$ in the papers from which we extract the results presented and utilised below. The above definition is the one from [64] which is conformal with $|||\cdot|||_\lambda$ from [17] and [22] in the sense that, $|||\Phi|||_\lambda \leq \|\Phi\|_\lambda \leq |||\Phi|||_{\lambda-\varepsilon}$ for all $\varepsilon > 0$.[1] Hence by requiring $\|\Phi\|_\lambda < \infty$ we immediately get that $|||\Phi|||_\lambda < \infty$ is satisfied for the same λ and we further can replace appearances of the latter with the first. Note further that the finite range interactions are exponentially decaying. Given a finite interval $\Lambda \Subset \Sigma$ split into X and Y, the expansional of a Hamiltonian on X, Y is defined analogously to Eq. (24), and the estimates of Proposition 2.1 extend similarly (see [64]). The important difference is that (ii) presents exponential decay with ℓ, rather than superexponential and that the result only holds up to some critical temperature $\beta_c(\lambda) > 0$. Here we recall the simplified formulation of [22, Lemma 31].

Proposition B.1 ([64, Theorem 3.1])**.** *Let Φ be an exponentially-decaying interaction which is further translation invariant and let $\beta < \beta_c(\lambda)$. Then the following hold:*

(i) There is an absolute constant $\tilde{\mathcal{G}} > 1$ depending only on λ and β such that, for any finite interval $\Lambda = XY \Subset \mathbb{Z}$ split into two subintervals X and Y, we have:

$$\|E_{X,Y}\|\,,\ \left\|E_{X,Y}^{-1}\right\| \ \leq\ \tilde{\mathcal{G}}\,.$$

[1] For proof of the relation, the reader can consult [22, Section 9.1].

(ii) There are positive constants $\mathcal{K}, \alpha > 0$ depending on λ and β such that if we add two intervals $\widetilde{X}$ and $\widetilde{Y}$ adjacent to X and Y, respectively, so that we get a larger interval $\widetilde{X}XY\widetilde{Y}$, then

$$\left\| E_{X,Y}^{-1} - E_{\widetilde{X}X,Y\widetilde{Y}}^{-1} \right\| , \left\| E_{X,Y} - E_{\widetilde{X}X,Y\widetilde{Y}} \right\| \leq \mathcal{K} e^{-\alpha \ell} .$$

for any $\ell \in \mathbb{N}$ such that $\ell \leq |X| , |Y|$.

Next, note that Lemma A.1 is independent of the range of the interactions. Note further that Proposition 2.2 relies on estimates for expansionals and certain contractions thereof which allows us to conclude its proof for exponentially-decaying interactions directly from Lemma A.1 and point (i) of Proposition B.1. The statement can be found in Proposition B.2. By similar arguments, we conclude Lemma 2.4 in the exponentially-decaying setting.

Proposition B.2. *Under the conditions above, there is an absolute constant $\tilde{\mathcal{C}} > 1$ depending only on λ and β such that*

$$\| \operatorname{tr}_B[\rho^B Q] \|, \| \operatorname{tr}_B[\rho^B Q]^{-1} \| \leq \tilde{\mathcal{C}} , \qquad Q \in \{ E_{B,C}^* , E_{B,C} , E_{A,B}^* , E_{A,B} \} , \tag{71}$$

$$\| \operatorname{tr}_{AB}[\rho^{AB} Q] \| , \| \operatorname{tr}_{AB}[\rho^{AB} Q]^{-1} \| \leq \tilde{\mathcal{C}} , \qquad Q \in \{ E_{A,B}^{*\,-1} , E_{A,B}^{-1} \} , \tag{72}$$

$$\| \operatorname{tr}_B \left[\rho^B E_{A,B}^* E_{AB,C}^* \right] \| , \| \operatorname{tr}_B \left[\rho^B E_{A,B}^* E_{AB,C}^* \right]^{-1} \| \leq \tilde{\mathcal{C}} , \tag{73}$$

$$\| \operatorname{tr}_{AC}[\rho^A \otimes \rho^C E_{AB,C} \, E_{A,BC}]^{-1} \| \leq \tilde{\mathcal{C}} . \tag{74}$$

The next result constitutes a generalisation of [16, Theorem 5.1] to the setting of exponentially-decaying interactions. Its proof completely follows that of its finite-range counterpart, and thus we only provide a very brief sketch of it, highlighting the differences.

Theorem B.3. *With Φ on $\mathbb{Z}$ an exponentially-decaying interaction that is further translation-invariant and with $\beta < \beta_c(\lambda)$, there are positive constants $\tilde{K}, \tilde{\alpha} > 0$ depending only on λ and β such that for every $\Lambda \Subset \mathbb{Z}$ split into three subintervals $\Lambda = ABC$, where B shields A from C, for its local Gibbs state $\rho^\Lambda =: \rho_{ABC}$ it holds that*

$$\| \rho_{ABC} \rho_{BC}^{-1} \rho_B \rho_{AB}^{-1} - \mathbb{1} \| \leq \tilde{K} e^{-\tilde{\alpha} |B|} . \tag{75}$$

Sketch of the proof. From [16, Eq. (19)], we can rewrite

$$\rho_{ABC} \rho_{BC}^{-1} \rho_B \rho_{AB}^{-1} = E_{A,BC} \operatorname{tr}_A[E_{A,BC} \rho^A]^{-1} \operatorname{tr}_A[\tilde{E}_{A,BC} \rho^A] \tilde{E}_{A,BC}^{-1} ,$$

with

$$\tilde{E}_{A,BC} = \operatorname{tr}_C[e^{-H_{ABC}}] \operatorname{tr}_C[e^{-H_{BC}}]^{-1} e^{H_A} .$$

Now, the conclusion of the proof is a combination of two statements:

1. Statement: The four factors in the right-hand side of the previous expression and their inverses are uniformly bounded. This is a consequence of Proposition B.1 and Proposition B.2. The universal constant is given by $\tilde{\mathcal{G}} \tilde{\mathcal{C}}^2$.

2. Statement: The following inequalities hold,

$$\left\| \tilde{E}_{A,BC} - E_{A,BC} \right\| \leq \tilde{K}' e^{-\tilde{\alpha}' |B|} ,$$
$$\left\| \operatorname{tr}_A[\tilde{E}_{A,BC} \rho^A] - \operatorname{tr}_A[E_{A,BC} \rho^A] \right\| \leq \tilde{K}' e^{-\tilde{\alpha}' |B|} .$$

The second inequality follows from the first one by contractivity. For the first one, we split the terms in the same way as in the proof of [16], obtaining $\tilde{K}' = 4 \mathcal{C} \, \mathcal{K} \, e^{-\alpha}$ and $\tilde{\alpha}' = \alpha / 2$.

We conclude by taking $\tilde{\alpha} = \tilde{\alpha}'$ and $\tilde{K} = 8\tilde{\mathcal{C}}^7\tilde{\mathcal{G}}^3\mathcal{K}e^{-\alpha}$. □

Next, one can immediately conclude Lemma 2.5 for exponentially-decaying interactions as of Eq. (74). This, jointly with Theorem B.3, allows us to deduce a generalisation of Theorem B.3 to the case where the Gibbs state is given on $\Lambda = A'ABCC'$ and but we only compare its marginals on ABC, AB, B and BC, respectively. Namely, we can show that, in this context, the following inequality holds:

$$\left\|\rho_{ABC}\rho_{BC}^{-1}\rho_B\rho_{AB}^{-1} - \mathbb{1}\right\| \leq \tilde{C}^4\tilde{K}e^{-\tilde{\alpha}|B|}\,. \tag{76}$$

To conclude the preliminaries for exponentially-decaying interactions, let us recall that the exponential decay of correlations and local indistinguishability hold at any positive temperature for the considered setting. The proofs of these results can be found in [22, Theorem 26] and [22, Theorem 28], respectively, as a consequence of [64, Theorem 4.4].

Proposition B.4 (Decay of correlations and local indistinguishability)**.** *For Φ an exponentially-decaying, translation invariant interaction over $\mathbb{Z}$, the following inequalities hold for universal constants $c, c', \gamma, \gamma' > 0$ depending only on λ and β:*

$$\left|\mathrm{Tr}_{ABC}[\rho^{ABC}O_A\,O_C] - \mathrm{Tr}_{ABC}[\rho^{ABC}O_A]\,\mathrm{Tr}_{ABC}[\rho^{ABC}O_C]\right| \leq \|O_A\|\|O_C\|\,c\,e^{-\alpha|B|}\,,$$
$$\left|\mathrm{Tr}_{ABC}[\rho^{ABC}O_A] - \mathrm{Tr}_{AB}[\rho^{AB}O_A]\right| \leq \|O_A\|\,c'\,e^{-\gamma'|B|}\,,$$
$$\left|\mathrm{Tr}_{ABC}[\rho^{ABC}O_C] - \mathrm{Tr}_{BC}[\rho^{BC}O_C]\right| \leq \|O_C\|\,c'\,e^{-\gamma'|B|}\,.$$

B.1 Exponential decay of the BS-CMI for exponentially-decaying interactions

First, note that most estimates from Lemma 3.5 follow straightforwardly in the context of exponentially-decaying interactions, and only the last one differs. For that, we have as before

$$\left\|\rho_B^{-1}\right\|\|\rho_B\| = \left\|(\rho^B)^{-1}\,\mathrm{tr}_{AC}[E_{A,BC}E_{B,C}\rho^A\rho^C]^{-1}\right\|\left\|\rho^B\,\mathrm{tr}_{AC}[E_{A,BC}E_{B,C}\rho^A\rho^C]\right\|\,,$$

with the difference that now

$$\left\|(\rho^B)^{-1}\right\|\left\|\rho^B\right\| \leq e^{2\|H_B\|} \leq e^{2|B|\beta\|\Phi\|_\lambda}\,.$$

Having this in mind, we can show the following bounds for the decays of the various versions of the BS-CMI by replacing the respective bounds in the proof of Theorem 3.6.

Theorem B.5. *Let Φ an exponentially-decaying, translation invariant interaction over $\mathbb{Z}$ and $\Lambda \Subset \mathbb{Z}$ any interval split into consecutive parts $\Lambda = A'ABCC'$, with A' and C' possibly empty. Then for the marginal on ABC of its local Gibbs state $\mathrm{tr}_{A'C'}[\rho^\Lambda] = \rho_{ABC}$ there exist a critical temperature $\widetilde{\beta}_c(\lambda)$ universal constants $c, \alpha, \gamma > 0$ independent of Λ and only dependent on λ and β such that*

$$\widehat{I}^x_\rho(A\,;C|B) \leq \tilde{c}e^{\tilde{\alpha}|A|}e^{-\gamma|B|} \quad x \in \{\mathrm{os}, \mathrm{ts}, \mathrm{rev}\}\,. \tag{77}$$

Proof. The bounds themselves are completely analogous to the proof in Theorem 3.6, replacing the superexponentially decaying function $\varepsilon(|B|)$ by $\tilde{K}e^{-\tilde{\alpha}|B|}$ from Theorem B.3. Now whenever $\tilde{\alpha} > 2\beta\|\Phi\|_\lambda$, we obtain a positive decay rate $\gamma = \tilde{\alpha} - 2\beta\|\Phi\|_\lambda$. Since $\tilde{\alpha} = \lambda - 2\beta\Omega_0$ (see [64, Section 2.3.2]) we can find $\widetilde{\beta}_c(\lambda)$ such that $\gamma > 0$ for $\beta < \widetilde{\beta}_c(\lambda)$. □

B.2 Efficient estimation of the global purity for exponentially-decaying interactions

The objective of this subsection is to develop the findings about the efficient evaluation of the overall purity of a Gibbs state on a translation-invariant spin chain in the context of interactions that decay exponentially. We will apply the necessary technical results to this context to achieve this. However, most results have similar proofs to the finite-range case, so we will skip those parts for simplicity and only highlight the differences.

First, we can prove the following proposition: the generalisation of Proposition 3.9.

Proposition B.6. *Let Φ be an exponentially-decaying, translation-invariant interaction. Then, for $\beta < \beta_c(\lambda)$ there exist positive constants $\tilde{c}_p, \tilde{\alpha}_p$ depending only on λ and β such that, for every $\Lambda \Subset \mathbb{Z}$ split as $\Lambda = ABC$ and for $\rho_\Lambda := \rho^\Lambda$ the Gibbs state on Λ, the following holds.*

$$\left| \frac{\operatorname{Tr}_{AB}[\rho_{AB}^2] \operatorname{Tr}_{BC}[\rho_{BC}^2]}{\operatorname{Tr}_{\Lambda}[\rho_{\Lambda}^2] \operatorname{Tr}_{B}[\rho_{B}^2]} - 1 \right| \leq \tilde{c}_p e^{-\tilde{\alpha}_p |B|} . \tag{78}$$

Sketch of the proof. The proof is analogous to that of Proposition 3.9. Let us recall that we can estimate the left-hand side of Eq. (78) by:

$$\left| \widetilde{\lambda}_{ABC}^{-1} - 1 \right| + \left| \widetilde{\lambda}_{ABC}^{-1} \right| \left| \chi_{ABC} - 1 \right| ,$$

where the terms are defined completely analogously to the finite-range case, but we keep the same notation for simplicity. The estimates of $\left| \widetilde{\lambda}_{ABC}^{-1} - 1 \right|$ and $\left| \widetilde{\lambda}_{ABC}^{-1} \right|$ were shown in [17, Lemma 6.1] to hold exactly in the same way as for finite range. In the finite range case we used a combination of estimates on Araki's expansionals, decay of correlations and local indistinguishability to conclude the decay of $|\chi_{ABC} - 1|$. All these properties also hold for exponentially-decaying interactions as a consequence of Proposition B.1 and Proposition B.4, hence we can conclude the proof. □

Next, note that the global purity of a Gibbs state split into many regions can be estimated efficiently from this proposition in the same way as in Theorem 4.11. Since the proof is the same as for Theorem 4.11, and follows from [80], we omit it here.

Theorem B.7. *Under the conditions of Proposition B.6, there is an algorithm that outputs an estimate $\hat{P}_2(\rho_{1:N})$ such that*

$$\left| \frac{\hat{P}_2(\rho_{1:N})}{\operatorname{Tr}\left[\rho_{1:N}^2\right]} - 1 \right| \leq \varepsilon, \tag{79}$$

with a number of samples and classical post-processing cost given by $\operatorname{poly}(n/\varepsilon)$.

B.3 Learning of Gibbs states via MPO approximations for exponentially-decaying interactions

In analogy with the previous subsection, the recovery map employed in the reconstruction of the Gibbs state from its marginals for the MPO approximation is defined in the same way for exponentially-decaying interactions as for the finite-range ones. Therefore, the estimate in Theorem 4.4 for the MPO approximation is still valid, and we omit its explicit formulation in this context for simplicity. The only change occurs in the analogue of Corollary 4.5, since the bond dimension is inherited from the decay of the BS-CMI, and thus it is expected to be slightly worse in this case. Indeed, as a consequence of Theorem B.5, we obtain the following.

Corollary B.8. *For a n-site marginal of a Gibbs state in one dimension with translation-invariant, exponentially-decaying interaction Φ and a given accuracy ε, there is an MPO representation of bond dimension*

$$D = \exp\Big(2\log(d)\tilde{\mathcal{C}}_1 \log(n/\varepsilon)\Big), \tag{80}$$

where $\tilde{\mathcal{C}}_1$ is a constant depending on λ and β and $\tilde{\alpha} > 0$ is obtained from Theorem B.5. The MPO representation is given by

$$\left\| \left(\bigcirc_{i=1}^{N-1} \mathcal{R}_i \right)(\rho_1) - \rho_{1:N} \right\| \leq \varepsilon$$

choosing the A_i as consecutive regions of at least l spins

$$|A_i| = l \geq \mathcal{C}_1 \log(n/\varepsilon)\,.$$

C Lifting the upper bound on the DPI of the BS-entropy to channels

Note that just by restriction of the original Hilbert space to the common support of X and Y, one can obtain the same result as Theorem 3.2 for the case that X and Y are not full-rank but share the same kernel. In that case, the inverses are replaced by Moore-Penrose pseudo-inverses. With this at hand, we can now prove the upper bound on the DPI for the BS-entropy for quantum channels instead of conditional expectations.

Corollary C.1. *For $X, Y \in \mathcal{B}(\mathcal{H}_A)$ having the same support and $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ a quantum channel, we find that*

$$\begin{aligned} \widehat{D}(X\|Y) - \widehat{D}(\mathcal{T}(X)\|\,\mathcal{T}(Y)) \leq &\left\|X^{-1/2}YX^{-1/2}\right\| \|X\|_1 \left\|\mathcal{T}(X)^{1/2}\right\| \left\|\mathcal{T}(X)^{-1/2}\right\| \\ &\cdot \left\|\mathcal{T}(Y)^{-1}\,\mathcal{T}(X)\right\| \left\|XY^{-1}\,\mathcal{T}^*(\mathcal{T}(Y)\,\mathcal{T}(X)^{-1}) - \mathbb{1}\right\| \end{aligned} \tag{81}$$

where inverses are replaced by Moore-Penrose pseudo-inverses in case the operators are not full-rank.

Proof. Following similar arguments as in [83, 12], we note that by Stinespring's dilation theorem, we can write every quantum channel as a composition:

$$\mathcal{T}(\cdot) = \mathrm{tr}_C[V \cdot V^*] \tag{82}$$

where $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_C$ is an isometry and $\mathcal{H}_C$ an auxiliary space. Using the isometric invariance of the BS-entropy and its additivity under tensor products, we get

$$\widehat{D}(X\|Y) - \widehat{D}(\mathcal{T}(X)\|\,\mathcal{T}(Y)) = \widehat{D}(X_V\|Y_V) - \widehat{D}(\pi_C \otimes \mathrm{tr}_C[X_V]\|\pi_C \otimes \mathrm{tr}_C[Y_V])\,,$$

where we set $X_V = VXV^*$ and $Y_V = VYV^*$. Both of the latter operators agree in their support, hence with the considerations brought forward at the beginning of this appendix we utilise Theorem 3.2 to obtain

$$\begin{aligned} \widehat{D}(X\|Y) - \widehat{D}(\mathcal{T}(X)\|\,\mathcal{T}(Y)) \leq &\left\|X_V^{-1/2}Y_VX_V^{-1/2}\right\| \|X_V\|_1 \left\|\mathcal{T}(X)^{1/2}\right\| \left\|\mathcal{T}(X)^{-1/2}\right\| \\ &\cdot \left\|\mathcal{T}(Y)^{-1}\,\mathcal{T}(X)\right\| \left\|X_VY_V^{-1}(\mathcal{T}(Y)\,\mathcal{T}(X)^{-1}) \otimes \mathbb{1}_C - \mathbb{1}\right\| \end{aligned}$$

where we already cancelled constants and used isometric invariance of the Schatten p-norms and Eq. (82) to simplify the expression. Using again isometric invariance on the remaining terms and identifying $\mathcal{T}^*$ confirms the claim. □

Quasi-optimal sampling from Gibbs states via non-commutative optimal transport metrics

Ángela Capel[†1,2], Paul Gondolf∗[‡1], Jan Kochanowski[§3], and Cambyse Rouzé[¶3]

[1]Fachbereich Mathematik, Universität Tübingen, 72076 Tübingen, Germany
[2]Dep. of App. Mathematics and Theo. Physics, University of Cambridge, United Kingdom
[3]Inria, Télécom Paris - LTCI, Institut Polytechnique de Paris, 91120 Palaiseau, France

July 9, 2025

Abstract

We study the problem of sampling from and preparing quantum Gibbs states of local commuting Hamiltonians on hypercubic lattices of arbitrary dimension. We prove that any such Gibbs state which satisfies a clustering condition that we coin decay of matrix-valued quantum conditional mutual information (MCMI) can be quasi-optimally prepared on a quantum computer. We do this by controlling the mixing time of the corresponding Davies evolution in a normalized quantum Wasserstein distance of order one. To the best of our knowledge, this is the first time that such a non-commutative transport metric has been used in the study of quantum dynamics, and the first time quasi-rapid mixing is implied by solely an explicit clustering condition. Our result is based on a weak approximate tensorization and a weak modified logarithmic Sobolev inequality for such systems, as well as a new general weak transport cost inequality. If we furthermore assume a constraint on the local gap of the thermalising dynamics, we obtain rapid mixing in trace distance for interactions beyond the range of two, thereby extending the state-of-the-art results that only cover the nearest neighbour case. We conclude by showing that systems that admit effective local Hamiltonians, like quantum CSS codes at high temperature, satisfy this MCMI decay and can thus be efficiently prepared and sampled from.

[†]ac2722@cam.ac.uk
[‡]paul.gondolf@uni-tuebingen.de
[§]jan.kochanowski@inria.fr
[¶]rouzecambyse@gmail.com
∗ corresponding author
authors listed in alphabetical order

Contents

1 Introduction 1

2 Main results 2
2.1 Notation and preliminaries . . . 4
2.2 Uniform decay of the MCMI . . . 5
2.3 Entropy factorization and a weak AT . . . 6
2.4 Quasi-rapid mixing of Lipschitz observables . . . 8
2.5 Rapid mixing under polynomial local gap . . . 10

3 A comparison with quantum Gibbs sampling 11

4 Outlook 12

5 Acknowledgments 12

A Theoretical Framework 20
A.1 Basic notation . . . 20
A.2 Mixing, rapid mixing, and logarithmic Sobolev inequalities for open quantum systems 22
A.3 Local commuting Hamiltonian and Davies generators on lattices . . . 23
A.4 Decay of correlations in Gibbs states . . . 25
A.4.1 Correlation measures and the MCMI . . . 26
A.4.2 MCMI decay from complete analyticity in classical systems . . . 27
A.4.3 The effective Hamiltonians as tool to proof correlation decay . . . 28

B Proof techniques - extended 29
B.1 A general weak entropy factorization . . . 29
B.2 A coarse-graining of the hypercubic lattice . . . 30
B.3 A weak approximate tensorization for Davies channels . . . 35
B.4 A weak transport cost inequality for Davies channels . . . 36
B.5 A MLSI alike inequality for local Davies semigroups at every temperature . . . 38
B.6 Weak MLSI for the global Davies semigroup . . . 39

C Main results - extended 42
C.1 W_1-mixing from MCMI-decay . . . 42
C.2 Rapid mixing from MCMI-decay and polynomial local gap . . . 44
C.3 Quasi-Optimal Gibbs state preparation from MCMI-decay . . . 46

D Examples 47
D.1 MCMI-decay from strong effective Hamiltonian . . . 47
D.2 MCMI-decay from commuting marginals at high temperature . . . 48

1 Introduction

The problem of Gibbs state preparation is fundamental in statistical mechanics. Given a Hamiltonian H on a quantum system and an inverse temperature $\beta < \infty$, the Gibbs state $\sigma^\beta = e^{-\beta H} / \operatorname{Tr}\left[e^{-\beta H}\right]$ describes the properties of the system at its thermal equilibrium [Alh23]. Thus, Gibbs states and their fundamental properties are essential for the study of quantum systems, for example in the contexts of simulation of many-body systems [MSVC15], their use as topological quantum memories [LCP13] or their thermalization processes [RGE12, MAMW15].

Gibbs sampling has been for a long time a cornerstone in statistical mechanics. It is an extremely relevant problem for the fields of physics and computer science, both in the context of classical and quantum mechanics, in part due to its applications in multiple scenarios [vAGGdW20, HS83]. Classical Markov Chain Monte Carlo (MCMC) methods constitute one of the canonical tools for sampling from classical Gibbs states of spin systems [LPW08]. These methods have been demonstrated to be efficient at sufficiently high temperatures [Mar99] and are widely considered efficient in practice [BGJM11]. However, there is no clear consensus on how to best extend such algorithms to quantum Gibbs sampling. Many relevant classical algorithms for Gibbs sampling have appeared in the past few years [CHPZ24, AJK+21a, AJK+21b, AMS22]. Here we take a further step and aim at proving the efficiency of quantum extensions thereof.

A good Gibbs sampler is expected to prepare the Gibbs state in polynomial time. In the past few years, many quantum algorithms inspired by the classical Monte Carlo have been proposed in [TOV+11, RWW23, WT23], among others, but for long without any provable guarantee or only validated under strong theoretical assumptions [SM21, CB21]. This has drastically changed with the very recent appearance of [CKG23, CKBG23], where a quantum algorithm to prepare quantum Gibbs states was proposed and subsequently shown to be efficient in [RFA24a]. This has inspired an important collection of works in the context of quantum Gibbs sampling, with algorithms not only based on dissipation [CKBG23, GCDK24, BCG+24, RWW23, DLL24], but also in some other techniques [GSLW19, JI24, KSM+24]. A detailed discussion on some relevant literature for quantum Gibbs sampling is presented in Section 3.

In this paper, we investigate quantum Gibbs sampling with Davies generators [Dav79, Dav76] associated with local commuting Hamiltonians. Local commuting Hamiltonians hereby constitute a class of systems that non-trivially extends beyond the classical regime [BH24, HJ24, AKV18, AE11] and include highly entangled systems like CSS codes (e.g. the Toric code) and quantum double models [Kit03, CS96, Ste96]. Davies Lindbladians constitute the canonical objects to model the Markovian dissipation of a quantum system weakly coupled with an infinite-dimensional thermal bath and can be regarded as the natural quantum analogue of Glauber dynamics [Mar99]. Hence, they represent a natural tool for quantum Gibbs sampling. Many recent efforts to develop efficient Gibbs samplers have been undertaken in the context of commuting Hamiltonians, such as for 1D translation-invariant systems at any positive temperature [BCG+23, BCG+24, KACR24], and high-dimensional systems at high-enough temperature [CRF20, KACR24]. In high dimensions at low temperature, efficient Gibbs samplers are only known to exist for Kitaev's quantum double models in 2D [AFH09, DLLZ24, LPGPH23, KLCT16].

The efficiency of Gibbs samplers is determined by the speed of convergence, or mixing, of the corresponding Lindbladian towards its thermal equilibrium, estimated through the notion of *mixing time*. Generically, the algorithmic sample complexity can be upper bounded by the system size times its mixing time [LW22, RWW23]. Both classically and quantumly, the most frequent way of estimating this mixing time is through the spectral gap of the Lindbladian [Bar17, KB16]. A positive spectral gap provides an upper bound on the mixing time that scales linearly with the system size, in a regime known in the literature as *fast mixing*, or *poly-time mixing* [HJ24], but this is generally believed to be an overestimation. A more precise estimate follows from the existence of a $\Omega(1/\operatorname{poly}\log N)$ modified logarithmic Sobolev inequality [KT13], which implies a mixing time that scales polylogarithmically with the system size, a regime known as *rapid mixing*. For continous time local reversible Glauber dynamics on bounded-degree graphs, it has been shown that the mixing time obeys $\Omega(\log N)$ [HS07, Theorem 4.1] while in [CLV23] the bound was achieved for a subclass of models, establishing the optimality of rapid mixing. In this regime, the algorithmic sample complexity scales linearly with system size, up to a polylogarithmic correction, a property commonly referred to as *optimal sampling*.

In accordance with this nomenclature we use this manuscript the term quasi-rapid mixing to describe a mixing time that scales quasi-logarithmically with system size, specifically $\mathcal{O}(\exp\{\operatorname{poly}\log\log\})$. Such a mixing time implies hence an algorithmic complexity that remains linear up to quasi-logarithmic corrections—significantly better than the quadratic scaling one would obtain from a naive gap assumption. In line with the literature, we refer to this preparation complexity as quasi-optimal [HS07]. In terms of scaling, however, it is much closer to the linear scaling—optimal for preparation circuit complexity—than to any polynomial growth. Similarly, a quasi-logarithmic mixing time can be considered quasi-optimal in the sense that it is far closer to the polylogarithmic scaling, which is optimal, than to any polynomial scaling.

Whereas some quantum Gibbs samplers associated with non-commuting Hamiltonians were previously only known to exhibit fast mixing [RFA24a], recent results have established that rapid mixing can also occur in this setting under the existence of Lieb-Robinson bounds [RFA24b]. Until now, examples of rapid mixing had been primarily found in the context of commuting Hamiltonians [CRF20, BCG$^+$23, BCG$^+$24, KACR24]. This new development expands the scope of rapid mixing beyond commuting systems, though our focus in this manuscript remains on the latter.

2 Main results

In this work, we demonstrate the quasi-rapid mixing in normalized Wasserstein distance of Davies generators associated with local commuting Hamiltonians, assuming that their Gibbs state satisfies a specific clustering condition. By additionally assuming a polynomially decaying gap of the local Davies generator, we strengthen our result to rapid mixing in trace distance.

To our knowledge, this is the first instance where quasi-rapid mixing is derived solely from a static and explicit notion of decay of correlations in the steady-state or Gibbs state. Furthermore, while mixing times based on normalized Wasserstein distances have been studied in the classical setting [AMS22], we believe this is the first time that mixing with respect to normalized quantum Wasserstein distances has been considered in the quantum setting, thereby further connecting quantum optimal transport theory and Gibbs sampling [DPR22, RD19, DPMTL21, DPP24, BF24].

We introduce the *matrix-valued quantum conditional mutual information (MCMI)* in Definition 2.3, whose decay is more general than that of both the covariance and the conditional mutual information—the former shown to be sufficient for rapid thermalization in one-dimensional lattices [KACR24]. Our first main result requires only the uniform decay of the MCMI across arbitrary 4-partitions of the underlying lattice:

Theorem 2.1 (Quasi-optimal preparation of MCMI-decaying Gibbs states, informal) *Let σ be a Gibbs state of a local commuting Hamiltonian that satisfies uniform exponential decay of its MCMI. Then there exists a quantum circuit with circuit complexity and runtime*

$$\mathcal{O}\left(N, \text{quasi-log}(N), \text{quasi-poly}\left(\frac{1}{\epsilon}\right)\right) = o(N^{1+\delta})$$

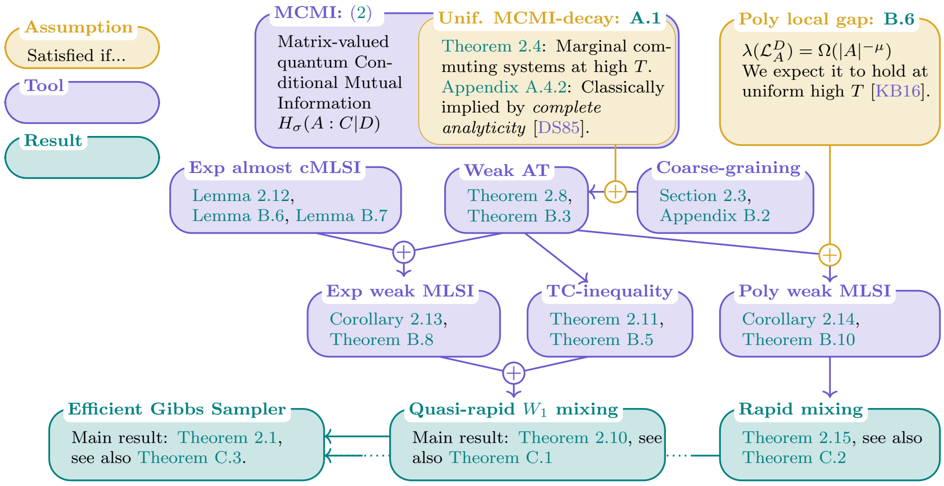


Figure 1: Logical structure of the results in this article. In orange, we represent the assumptions considered; in purple, the main technical lemmas, of independent interest; and in green, the main technical results regarding the mixing times of the evolution and their application in the context of Gibbs sampling. Note that the main result is implied by either the quasi-rapid W_1 mixing, or the rapid mixing.

that prepares a state ρ that is ϵ-close in normalized W_1 distance to σ, for any $\delta > 0$.

Corollary 2.2 *Let σ be a Gibbs state of a commuting Hamiltonian, which satisfies the marginal commuting property, see Appendix A.4.3, at uniformly high temperature can be prepared quasi-optimally, i.e. there exists a quantum circuit with complexity and runtime*

$$\mathcal{O}\left(N, \text{quasi-log}(N), \text{quasi-poly}\left(\frac{1}{\epsilon}\right)\right) = o(N^{1+\delta})$$

that prepares a state ρ that is ϵ-close in normalized W_1 distance to σ, for any $\delta > 0$.

To establish the quasi-rapid mixing time and achieve efficient sampling, we prove a suitable *weak modified logarithmic Sobolev inequality* (weak MLSI) for the Davies evolution corresponding to a local commuting Hamiltonian (see Corollary 2.13). This yields exponential decay of the relative entropy between any time-evolved initial state and the Gibbs state, with a prefactor depending polynomially on the system size. Following recent work [CRF20, BCG$^+$24, KACR24], we employ a 'divide and conquer' strategy, using the uniform decay of the MCMI for the global Gibbs state and leveraging the locality of the Davies generators to estimate global mixing in terms of local mixing plus an additive error term.

More precisely, we employ the uniform decay of the MCMI to prove a general entropy factorization (Lemma 2.6), which we extend to an approximate tensorization on $\mathbb{Z}^D$ (Theorem 2.8). This result reduces the relative entropy of two states supported on the global lattice to a sum of conditional relative entropies in smaller subregions, with an additive error controlled by the decay of the MCMI. This is the crucial ingredient to prove the weak MLSI (Corollary 2.13). Finally, to apply this weak MLSI to our Wasserstein distance-based mixing time, we derive a general *weak transport cost* inequality (Theorem 2.11).

Notably, unlike previous results on MLSIs of Davies dynamics in many-body quantum spin systems [BCG+23, BCG+24, KACR24], our main result (Theorem 2.10, see also Theorem C.1) shows that quasi-rapid mixing is implied solely by the decay of the MCMI, without additional assumptions on quantities like the local gap. This is significant because it is the first time that a dynamical property as strong as quasi-rapid mixing can be derived for physically relevant quantum spin systems solely from a static condition on the system's equilibrium (the Gibbs state), without further weaker dynamical assumptions.

Moreover, under further assumptions—specifically, an at most polynomially decaying gap of the local Davies generators, which we expect to hold at uniformly high-temperature —we also show rapid mixing of these dynamics with respect to the trace distance (see Theorem 2.15).

In the following sections, we provide a detailed overview of these results and the required lemmas, while omitting most technical details and proofs. A formal introduction to the necessary mathematical concepts, along with the full proofs, can be found in Appendix A, Appendix B, and Appendix C. We conclude with Appendix D, where we present examples of systems for which the required decay of correlations—specifically, the decay of the MCMI—holds, making our main results applicable.

For the readers' convenience, we collect all main results, important technical lemmas and assumptions in Figure 1.

2.1 Notation and preliminaries

This work considers quantum spin systems on finite hypercubic lattices $\Lambda \subset \mathbb{Z}^D$. The finite-dimensional Hilbert space of said systems is $\mathcal{H}_\Lambda := \bigotimes_{x\in\Lambda}\mathcal{H}_x$, where each site contains a qudit $\mathcal{H}_x \simeq \mathbb{C}^d$. The operator norm on the set of bounded operators $\mathcal{B}(\mathcal{H})$ is denoted as $\|\cdot\|_\infty$ and the trace norm as $\|\cdot\|_1$. The set of all quantum states, i.e. non-negative, trace-normalized operators on $\mathcal{H}$, is denoted as $\mathcal{S}(\mathcal{H})$. The partial trace is $\operatorname{tr}_A : \mathcal{B}(\mathcal{H}_\Lambda) \to \mathcal{B}(\mathcal{H}_{\bar{A}})$, where $\bar{A} := \Lambda \setminus A$ is the complement region of $A \subset \Lambda$. A $(\kappa, r)-$local, commuting, $J-$bounded Hamiltonian $H_\Lambda \in \mathcal{B}(\mathcal{H})$ is a self-adjoint operator $H_\Lambda = \sum_{A\subseteq\Lambda} h_A$, such that $\|h_A\|_\infty \le J$, $h_A = 0$, whenever $|A| > \kappa$, or $\operatorname{diam}(A) > r$, and for each $A, B \subset \Lambda$, $[h_A, h_B] = 0$, where $[X,Y] := XY - YX$ is the commutator of two operators $X, Y \in \mathcal{B}(\mathcal{H})$. We define the boundary of a sublattice $R \subseteq \Lambda$, tied to the connectivity of the Hamiltonian as $\partial R := \{k \in \Lambda : \exists A \subseteq \Lambda, h_A \neq 0, A \cap R \neq \emptyset, k \in A\}$ and write $R\partial := R \cup \partial R$. For a sublattice $R \subseteq \Lambda$ we define the *local Hamiltonian* as $H_R := \sum_{A\subseteq R} h_A$ and set the *growth constant* g of the system (Λ, H_Λ) as $g := \max_{k\in\Lambda} |\{A \subseteq \Lambda : h_A \neq 0, k \in A\}|$. Hence it satisfies $\|H_R\|_\infty \le gJ|R|$. The local, resp. global Gibbs state at inverse temperature $\beta > 0$, $R \subset \Lambda$, resp. $R = \Lambda$ is given by $\sigma^R := \frac{e^{-\beta H_R}}{\operatorname{Tr}[e^{-\beta H_R}]}$ and its reduced state is defined as $\sigma_R := \operatorname{tr}_{\bar{R}}[\sigma^\Lambda]$. The relative entropy between a state ρ and a full rank state $\sigma > 0$ is defined as $D(\rho\|\sigma) := \operatorname{tr}[\rho(\log\rho - \log\sigma)]$. The generator of the semigroup dynamic we are going to investigate in this paper is called the Davies semigroup, which for a fixed local commuting Hamiltonian H_Λ and inverse temperature β is given by

$$\mathcal{L}^D_\Lambda(\rho) := \sum_{k\in\Lambda}\mathcal{L}^D_k \quad \text{with} \quad \mathcal{L}^D_k(\rho) := \sum_{\alpha,\omega}\chi^{\beta,\omega}_{\alpha,k}\left(S^\omega_{\alpha,k}\rho S^{\omega,\dagger}_{\alpha,k} - \frac{1}{2}\left\{\rho, S^{\omega,\dagger}_{\alpha,k}S^\omega_{\alpha,k}\right\}\right). \tag{1}$$

The $S^\omega_{\alpha,k}$ depend on H_Λ through $e^{itH_\Lambda}S_{\alpha,k}e^{-itH_\Lambda} = \sum_\omega e^{it\omega}S^\omega_{\alpha,k}$ for all $t \in \mathbb{R}$, where $\{S_{\alpha,k}\}_\alpha$ labels a set of self-adjoint operators supported on k that form a Kraus-decomposition of the partial trace tr_k. The prefactors $\chi^{\beta,\omega}_{\alpha,k}$ satisfy the KMS condition $\chi^{\beta,-\omega}_{\alpha,k} = e^{-\beta\omega}\chi^{\beta,\omega}_{\alpha,k}$ and we assume them to be uniformly bounded from above and below as $0 < \chi^\beta_{\min} \le \chi^{\beta,\omega}_{\alpha,k} \le \chi^\beta_{\max}$. We define the local Davies generators as $\mathcal{L}_A^D := \sum_{k\in A}\mathcal{L}_k^D$ and note that for all $A \subseteq \Lambda$ the generated dynamics converge to projecting channels, we denote by $E_A = \lim_{t\to\infty} e^{t\,\mathcal{L}_A}$.

2.2 Uniform decay of the MCMI

In classical spin systems, there is a strong connection between static correlation properties in the thermal or steady state—such as spectral independence [ALG20] or decay of correlations [Mar99]—and the mixing time of the thermalizing dynamics. Specifically, Glauber dynamics, which models the thermalization of a discrete spin system, rapidly approaches the equilibrium Gibbs measure if, and only if, correlations between spatially separated regions decay exponentially with the distance between them. A similar connection has been established in the quantum case, where notions of decay of correlations have been shown to be closely related to mixing [BCPH24, CRF20, KACR24]. This work builds on these previous results by introducing the *matrix-valued quantum conditional mutual information (MCMI)* as a correlation measure closely tied to the mixing properties of the so-called Davies dynamics. This quantity, previously considered in the study of quantum many-body systems [KKB20], is closely related to quantum conditional mutual information and the mixing condition studied in [BCG$^+$24, BCPH22, BCPH24], where the mixing property was identified as a condition to imply rapid mixing of Davies dynamics on 1D lattices.

Definition 2.3 (Matrix valued quantum conditional mutual information (MCMI)) Given a quantum state $\sigma \in \mathcal{S}(\mathcal{H}_{ABCD})$ on some tensor product Hilbert space $\mathcal{H}_{ABCD} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$, then its MCMI is defined as

$$\mathbf{H}_\sigma(A : C|D) := \log\sigma_{ACD} + \log\sigma_D - \log\sigma_{AD} - \log\sigma_{CD}, \tag{2}$$

and we will frequently write $H_\sigma(A : C|D) := \|\mathbf{H}_\sigma(A : C|D)\|_\infty$ for its operator norm.

It is a general notion of mixing that controls both:

1. the conditional mutual information $I_\sigma(A : C|D) := \mathrm{tr}[\sigma\mathbf{H}_\sigma(A : C|D)] \le H_\sigma(A : C|D)$;
2. the mutual information and thus also covariance $I_\sigma(A : C) := \mathrm{tr}[\sigma_{AC}\mathbf{H}_\sigma(A : C|\emptyset)] \le H_\sigma(A : C|\emptyset)$

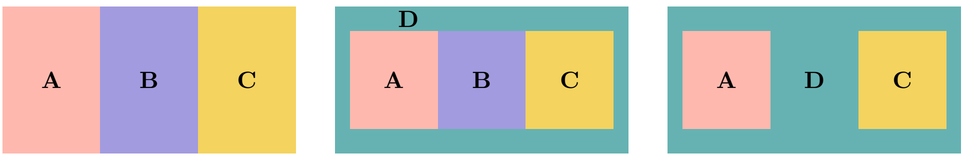


Figure 2: A lattice Λ is partitioned into four distinct regions, such that A and C are separated by either B or D or both. B is a system that is averaged over (through a partial trace) while D we condition on.

For σ a Gibbs state of a local Hamiltonian at inverse temperature β, the norm of the MCMI is believed to decay exponentially in $\mathrm{dist}(A, C)$ at high-enough temperature. This was, erroneously, thought proven in [KKB20] for some time. Although we do not resolve this open question, we prove that the conjecture holds for *marginal commuting* Hamiltonians: We say that $H_\Lambda = \sum_{A\subseteq\Lambda} h_A$ is *marginal commuting* (see [BCPH24, Definition 3.5]) if the algebra $\mathcal{A}$ generated by $\{h_A\}_{A\subset\Lambda}$ is commuting and closed under arbitrary partial traces, i.e. for all $A \subseteq \Lambda$, $I_A \otimes \mathrm{tr}_A[\mathcal{A}] \subseteq \mathcal{A}$. This condition, at high-enough temperature, is a particular case of applicability of the following theorem:

Theorem 2.4 (Uniform decay of the MCMI) *Let σ be the Gibbs state to some (κ, r)-local and J-bounded Hamiltonian on a graph Λ that has growth constant g, then there exist some constants $K, \xi > 0$ such that for any partition $\Lambda = A \sqcup B \sqcup C \sqcup D$*

$$H_\sigma(A : C|D) \leq K|\Lambda| \exp(-\operatorname{dist}(A, C)/\xi), \tag{3}$$

if H_Λ admits a strong local effective Hamiltonian in the sense of Appendix A.4.3, see also Definition 3.1 of [BCPH24]. In this case, we say that the state σ satisfies uniform decay of the MCMI.

Remark 1. In [BCPH24], it was shown that all marginal commuting systems at suitably high temperature satisfy this condition and hence, by the theorem above, also uniform MCMI decay, see Appendix D.2 and Theorem D.2. These notably include all quantum CSS codes, and all classical Hamiltonians with a Gibbs state that satisfies complete analyticity [Ces01, DS85, DS87], see Appendix A.4.2.

2.3 Entropy factorization and a weak AT

Statements on (strong) approximate tensorization (AT) relate conditional relative entropies on lattices to ones on subsystems with multiplicative error terms that quantify the closeness of the fixed point to being a product state. Such statements are central in both the classical and quantum setting when proving MLSIs and efficient bounds on mixing times of certain dynamics. The issue, however, is that, unlike in the classical setting, so far no general iterable approximate tensorization statement is known in the quantum setting. The ones that are known are sufficient when considering the fixed point to be a tensor product [CLPG18b], systems on 1-dimensional lattices [BCG+23, BCG+24], or nearest-neighbour interactions [CRF20, KACR24] but do not allow to go beyond these settings. In this work, we circumvent this problem by considering a *weak approximate tensorization (weak AT)* which will suffice for the quasi-rapid mixing of the Davies dynamics in W_1-distance.

To state our weak AT we first want to introduce the concept of conditional relative entropy which will serve as a useful tractable proxy for the Davies conditional relative entropy we ultimately wish to study. For more detail see Appendix A.3 and (48).

Definition 2.5 (Conditional relative entropy [CLPG18b]) Given a lattice Λ the conditional relative entropy on $A \subset \Lambda$ between two states $\rho, \sigma \in \mathcal{S}(\mathcal{H}_\Lambda)$ is defined as

$$D_A(\rho\|\sigma) := D(\rho\|\sigma) - D(\rho_{\bar{A}}\|\sigma_{\bar{A}}). \tag{4}$$

We are now able to show the following novel relation between the conditional relative entropies:

Lemma 2.6 (Weak entropy factorization) *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$ and $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ with $\sigma > 0$, then*

$$D_{ABC}(\rho\|\sigma) \leq D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma) + \|\mathbf{H}_\sigma(A : C|D)\|_\infty . \tag{5}$$

The proof only requires the data-processing-inequality and Hölder inequality for Schatten-p norms and can be found in Lemma B.1. Unlike the previously known entropy factorizations [CLPG18b, CRF20], our new results allow us to have a system D that we carry as a conditioning through the estimate which is key in the iteration of this statement, yielding a weak AT statement for lattices of arbitrary dimension later on. Note that classically (i.e. when all involved states commute) one can derive a multiplicative correction, i.e. a strong *entropy factorization* of the form

$$(1 - \|\exp(\mathbf{H}_\sigma(A : C|D)) - I\|_\infty) D_{ABC}(\rho\|\sigma) \leq D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma) . \tag{6}$$

Establishing a similar relation in the quantum setting would allow us to lift all weak inequalities to their strong counterparts making for example the requirement of uniform polynomial local gap as a prior for rapid mixing obsolete (c.f. Theorem 2.15). However, establishing such an inequality for quantum systems where $D \neq \emptyset$ remains still open.

To extend the above entropy factorization to a (weak) approximate tensorization statement on a lattice $\Lambda_L := [\![-L, L]\!]^D \subset \mathbb{Z}^D$, we require a suitable *coarse-graining* of Λ w.r.t. which we split the conditional relative entropy. For $D = 1, 2$ such were given in [BK18]. Effectively a suitable coarse-graining is a family of subsets $\{C_{a,i}\}_{a,i}$ that covers Λ, are all suitably local, and have overlaps between neighbouring levels a.

Lemma 2.7 *For any $D \in \mathbb{N}$, there exists a suitable coarse-graining. Explicitly, given $r \leq k, c, \ell \leq 2L+1$ such that $\ell > 2D(k+c)$, there exists a coarse-graining $\{C_{a,i}\}_{a\in[\![D]\!],i}$ of $\Lambda_L = [\![-L, L]\!]^D$ made up of $D+1$ different levels indexed by a, such that*

1. *$\bigcup_{a,i} C_{a,i} = \Lambda$ and each site $x \in \Lambda$ is included in at most $D+1$ sets $C_{a,i}$,*
2. *$\{C_{a,i}\}_i$ is a collection of mutually disjoint subsets $\forall a \in [\![D]\!]$,*
3. *$|C_{a,i}| \leq \ell^D$,*
4. *Each level a has a suitable overlap c with the level coming before and after.*

An explicit construction of the covering may be found in Appendix B.2 with its key properties summarized in Lemma B.2 and a figure detailing the construction for the case $D = 2$ in Figure 3.

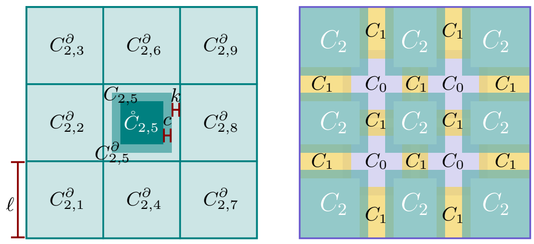


Figure 3: An example of the coarse-graining of a $D = 2$ dimensional lattice. On the left the largest regions in the top level $a = 2$ of the hierarchy with side length ℓ. On the right the different overlapping levels, labeled by $a = 2, 1, 0$, in different colors. Each of the local cells $C_{a,i}$ has size $|C_{a,i}| \leq \ell^2$.

With the entropy factorisation as well as the lattice coarse-graining in place, we are now able to prove one of the core results of this paper, namely the weak approximate tensorization of Gibbs states of local commuting Hamiltonians that exhibit uniform decay of the MCMI.

Theorem 2.8 (Weak approximate tensorization) *Given a suitable coarse-graining of the lattices Λ with constants (k, c, l) as described above, and corresponding Davies channels $\{E_{C_{a,i}} : a \in [\![0, D]\!], i \in \mathcal{I}_a\}$ corresponding to the Gibbs state σ at temperature β of a local commuting Hamiltonian, the following inequality holds for all $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$:*

$$D(\rho\|\sigma) \leq \sum_{a=0}^{D} \sum_{i_a\in\mathcal{I}_a} D(\rho\|E_{C_{a,i_a}}(\rho)) + \sum_{a=1}^{D} \zeta_a(\sigma), \qquad \zeta_a(\sigma) := \|\mathbf{H}_\sigma(X_a : Z_a|W_a)\|_\infty\,, \tag{7}$$

where $W_a \sqcup X_a \sqcup Y_a \sqcup Z_a =: \overline{C^a} \sqcup \mathring{C}_a \sqcup (C_a\backslash\mathring{C}_a) \sqcup (C^a\backslash C_a)$ with $d(X_a, Z_a) = c \geq r$, for $a \in [\![1, D]\!]$. The explicit definition of these sets can be found in Appendix B.2. If σ further satisfies uniform decay of MCMI with constants K and ξ, it holds for any $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$ that

$$D(\rho\|\sigma) \leq \sum_{a=0}^{D} \sum_{i_a\in\mathcal{I}_a} D(\rho\|E_{C_{a,i_a}}(\rho)) + D2^D K|\Lambda|e^{-c/\xi}\,. \tag{8}$$

We refer to (7) *as a* $(1, c_2)$-weak AT *w.r.t. the coarse-graining* $\{C_{a,i}\}_{a,i}$ *where*

$$c_2 = \sum_{a=1}^{D} \zeta_a(\sigma) \leq K 2^D K |\Lambda| e^{-\frac{c}{\xi}} .$$

The proof works by induction over a using the entropy factorization Lemma 2.6 and the coarse-graining introduced before, separating one level of the hierarchy in every step. It may be found in Appendix B.6.

2.4 Quasi-rapid mixing of Lipschitz observables

In this section, we prove the quasi-rapid mixing of the Davies dynamics w.r.t the quantum Wasserstein distance of order 1. We recall its definition.

Definition 2.9 (W_1-distance from [DPMTL21]) For two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H}_\Lambda)$, the quantum Wasserstein distance of order 1 is given as the dual norm to a Lipschitz norm.

$$\|\rho - \sigma\|_{W_1} := \max_{H \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}_\mathcal{I}), \|H\|_L \leq 1} \operatorname{tr}[H(\rho - \sigma)] , \tag{9}$$

where the Lipschitz-norm of $H \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}_\Lambda)$ is defined as

$$\|H\|_L := 2 \max_{k \in \Lambda} \min_{\tilde{H} \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}_{\Lambda \setminus \{k\}})} \left\| H - I_k \otimes \tilde{H} \right\|_\infty . \tag{10}$$

One can show that for two states $\rho, \sigma \in \mathcal{S}(\mathcal{H}_\Lambda)$, it satisfies $\frac{1}{2}\|\rho - \sigma\|_1 \leq \|\rho - \sigma\|_{W_1} \leq |\Lambda| \|\rho - \sigma\|_1$, and hence it is an extensive norm that indicates how well two states differ locally [DPMTL21].

Quite naturally we define the ϵ-mixing time for a primitive QMS with fixed point σ w.r.t this normalized quantum Wasserstein distance as

$$t_{\mathrm{mix}}^{W_1}(\epsilon) := \inf \left\{ t \geq 0 : \forall \rho \in \mathcal{S}(\mathcal{H}_\Lambda), \left\| e^{t \mathcal{L}_\Lambda}(\rho) - \sigma \right\|_{W_1} \leq |\Lambda| \epsilon \right\} . \tag{11}$$

Note that mixing times based on normalized Wasserstein distances have also been considered in recent classical Gibbs sampling, see e.g. [AMS22] in which the authors prove optimal sampling via algorithmic stochastic localization for certain spin-glass models at high temperature in normalized Wasserstein-2 distance.

The rest of this section is dedicated to proving the first main result of this work:

Theorem 2.10 (Quasi rapid Wasserstein mixing) *Let* $\{\mathcal{L}^D_{\Lambda_L}\}_L$ *be a family of Davies Lindbladians corresponding to a family of* (κ, r)*-local, commuting,* J*-bounded Hamiltonians* $\{H_{\Lambda_L}\}_L$*, each of which has growth constant* g*. Consider* $\epsilon > 0$ *and denote by* $N = |\Lambda_L|$ *the size of the lattice. Then, if the invariant states of* $\mathcal{L}^D_{\Lambda_L}$*, i.e. the Gibbs states* σ^{Λ_L}*, all satisfy uniform exponential decay of its matrix-valued quantum conditional mutual information (MCMI) as defined in* (2)*, the semigroups* $\{e^{t \mathcal{L}^D_{\Lambda_L}}\}_{t \geq 0}$ *satisfy*

$$t_{\mathrm{mix}}^{W_1}(\epsilon) = \text{quasi-poly}\left(\frac{1}{\epsilon^2} \operatorname{poly} \log \frac{N}{\epsilon^2} \right) = \text{quasi-poly}(\epsilon^{-1})_{\epsilon \to 0} \text{quasi-log}(N)_{N \to \infty} . \tag{12}$$

For a precise upper bound on $t_{\mathrm{mix}}^{W_1}(\epsilon)$*, see Theorem C.1.*

In the following, we give an overview of the main strategy and tools required to prove this result. The proof may be found in Appendix C.1 and consists of the following two ingredients.

1. Relating the W_1 distance to the relative entropy. We do this via a general *weak transport cost* (weak TC) inequality implied by the weak AT (7). This is an inequality relating the W_1 distance between any state and a fixed point of a QMS that satisfies a (weak) AT to their conditional relative entropy.

2. Showing sufficient decay of the relative entropy: We prove a (α, ϵ) weak MLSI, in which the cMLSI constant verifies $\alpha(\epsilon)^{-1} = \mathcal{O}(\exp \operatorname{poly}\log \frac{N}{\epsilon})$. This requires the combination of the (weak) AT with a cMLSI alike estimate relating the Davies conditional relative entropy $D(\rho \| E_A(\rho))$ to the entropy production of $\mathcal{L}^D_{A\partial}$ at an exponential cost in $|A\partial|$.

Let us begin with the statement of the weak transport cost inequality for Gibbs states of local commuting Hamiltonians:

Theorem 2.11 (Weak transport cost inequality) *Let σ be the Gibbs state at inverse temperature β of a local commuting Hamiltonian on a lattice Λ that satisfies a $(1, c_2)$-weak AT w.r.t. some coarse-graining $\{A_i\}_{i=1}^{n_A}$ of Λ, like in (7). Then it satisfies a (b_1, b_2)-*weak transport cost *inequality with $b_1 = 2\sqrt{2n_A}\max_i |A_i\partial|$ and $b_2 = |\Lambda|\sqrt{2c_2}$, i.e.:*

$$\|\rho - \sigma\|_{W^1} \leq \max_i 2\sqrt{2}|A_i\partial|\sqrt{n_A D(\rho\|\sigma)} + |\Lambda|\sqrt{2c_2} \tag{13}$$

The proof of this result may be found in Theorem B.5 of Appendix B.4. The second step estimates the Davies conditional relative entropy on some local region A with the entropy production on a slightly larger region:

Lemma 2.12 *Let $A \subseteq \Lambda$ and let $E_A := \lim_{t\to\infty} e^{t\mathcal{L}^D_A}$ be the conditional expectation of a local (on A) Davies generator at inverse temperature β corresponding to a local, commuting, $J-$bounded Hamiltonian that on the lattice Λ has growth constant g, then*

$$D(\rho\|E_A(\rho)) \leq e^{2gJ(1+2\beta|A\partial\partial|)}(\chi^0_{\min})^{-1}\operatorname{EP}_{\mathcal{L}^D_{A\partial}}(\rho), \tag{14}$$

where $\chi^0_{\min} > 0$ is the constant from (1).

This is not quite a bound on the cMLSI constant of the QMS generated by $\mathcal{L}_A$, because the entropy production on the RHS is the one of $\mathcal{L}_{A\partial}$ and not $\mathcal{L}_A$. However, this bound, together with the weak AT in Theorem 2.8 will suffice to obtain a suitable global MLSI constant. The idea behind the proof of this lemma (c.f. Appendix B.5) is to relate both the Davies conditional relative entropy and the entropy production of the Davies generators at inverse temperature β to the ones at 0. Then at 0 inverse temperature, the Davies semigroup is closely related to the depolarizing one allowing us to connect the two only by slightly enlarging the domain of the generator in the entropy production (c.f. Appendix B.6). A direct consequence of this Lemma, (7), and a standard Grönwall argument is the following weak MLSI.

Corollary 2.13 (Quasi-poly weak MLSI, Theorem B.8) *In the context of Theorem 2.10 it holds that*

$$D(e^{t\mathcal{L}^D_\Lambda}(\rho)\|\sigma) \leq e^{-\alpha(\epsilon)t}D(\rho\|\sigma) + \epsilon \quad \textit{where} \quad \frac{1}{\alpha(\epsilon)} = \mathcal{O}\left(\exp\left(\operatorname{poly}\log\frac{N}{\epsilon}\right)\right).$$

The proof of Theorem 2.10 now combines the above quasi-poly weak MLSI Corollary 2.13 with the weak transport cost inequality of Theorem 2.11. For the detailed proof see Theorem B.8.

2.5 Rapid mixing under polynomial local gap

Having stronger assumptions on the gap of the local Davies generators, i.e. faster local mixing unsurprisingly yields tighter mixing bounds. The gap of the local Davies generator on $A \subset \Lambda$ is defined as the distance between the largest and second to largest eigenvalue of $\mathcal{L}_A^D$. A formal definition is given in (27). In [GR22, RD19] it is shown to lower bound the local cMLSI constant of $\mathcal{L}_A$ as $\alpha_c(\mathcal{L}_A) \geq \Omega\left(\frac{\lambda(\mathcal{L}_A)}{|A|}\right)$. Assuming a polynomial local gap with degree $\mu \in \mathbb{N}_0$ amounts to

$$\lambda(\mathcal{L}_A) \geq \Omega(|A|^{-\mu}) \quad \text{and hence} \quad \alpha_c(\mathcal{L}_A) \geq \Omega(|A|^{-1-\mu}) \quad \forall A \subset \Lambda, \tag{15}$$

which is equivalent to assuming a tightened version of Lemma 2.12 (or Lemma B.9). Note that technically an exact gap is not necessary and a result where the region of the generator is enlarged would suffice (analogous to Lemma 2.12), however, we will here and in the following always assume a gap for simplicity. With this assumption on the gap we can improve the weak cMLSI as follows:

Corollary 2.14 (Polylogarithmic weak MLSI) *In the notation of Theorem 2.10 under the additional assumption of a local gap that is at most polynomially decaying with degree $\mu \in \mathbb{N}_0$ it holds that*

$$D(e^{t\,\mathcal{L}_\Lambda^D}(\rho)\|\sigma) \leq e^{-\alpha(\epsilon)t}D(\rho\|\sigma) + \epsilon, \tag{16}$$

where $\frac{1}{\alpha(\epsilon)} = \mathcal{O}\left(\left(\log\frac{N}{\epsilon}\right)^{D(1+\mu)}\right)$.

This weak MLSI combined with either Pinsker's inequality or Theorem 2.11 directly gives the following strengthened mixing time bounds.

Theorem 2.15 (Mixing under poly gap) *Let $\{\mathcal{L}_{\Lambda_L}^D\}_L$ be a family of Davies Lindbladians corresponding to a family of (κ, r)-local, commuting, J-bounded Hamiltonians $\{H_{\Lambda_L}\}_L$ with uniform growth constant g. Assume that the local gap is at most polynomially decreasing with degree μ. Consider $\epsilon > 0$ and denote by $N = |\Lambda_L|$ the size of the lattice. Then, if the invariant state of $\mathcal{L}_{\Lambda_L}^D$, i.e. the Gibbs state σ^{Λ_L}, satisfies uniform exponential decay of its matrix-valued quantum conditional mutual information (MCMI) as defined in (2), the semigroup $\{e^{t\,\mathcal{L}_{\Lambda_L}^D}\}_{t\geq 0}$ satisfy*

$$t_{mix}^1(\epsilon) = \mathcal{O}\left(\left(\log\frac{N}{\epsilon^2}\right)^{1+D(1+\mu)}\right)_{\substack{N\to\infty\\ \epsilon\to 0}}, \qquad t_{mix}^{W_1}(\epsilon) = \mathcal{O}\left(\left(\log\left(\frac{1}{\epsilon^2}\log\frac{N}{\epsilon^2}\right)\right)^{1+D(1+\mu)}\right)_{\substack{N\to\infty\\ \epsilon\to 0}}.$$

For the non-asymptotic expressions for these mixing times see Theorem C.2.

Polynomial assumptions on the local gap, have been used before to show rapid thermalization of Davies dynamics, e.g. for commuting quantum systems in 1D [BCG+24] under a condition equivalent to non-conditional MCMI-decay. In [KACR24] this was superseded, however, removing the required assumption on the local gap.

Remark 2. Note that instead of a uniform polynomial local gap, one can also require very high temperature, i.e. $\beta \sim \frac{1}{|\Lambda|}$ leading to a constant correction in Lemma 2.12, giving a result analogous to the one above only relying on the uniform decay of the MCMI.

To now get the main result Theorem 2.1 we can use any Lindblad simulation circuit to simulate $e^{t_{\mathrm{mix}}^{W_1}(\varepsilon)\,\mathcal{L}_\Lambda^D}$, which by the above, on any input state is close to σ in normalized W_1 distance. Using the one from [CKBG23, Theorem III.2] and our above mixing time bounds yields directly Theorem 2.1, though any Lindblad simulation circuit, e.g. also [LW22] may be used. We call this algorithm *quasi-optimal* since it scales as $\mathcal{O}(N)$ up to quasi-logarithmic corrections, following the convention to denote sampling algorithms which scale as $\mathcal{O}(N)$ up to logarithmic corrections as 'optimal'. For more details see Theorem C.3. A direct consequence of Theorem 2.10 and Theorem 2.15 is the following main result.

Corollary 2.16 *Let σ be a Gibbs state of a marginal commuting (see Appendix A.4.3), $(\kappa, r)-$local, $J-$bounded Hamiltonian on Λ_L at a uniformly suitably high temperature. Then the corresponding Davies semi group has a W_1-mixing time bounded by*

$$t_{\text{mix}}^{W_1}(\epsilon) = \text{quasi-poly}(\epsilon^{-1})_{\epsilon \to 0} \text{quasi-log}(N)_{N \to \infty} \tag{17}$$

and if additionally the gap of the local Davies generators is polynomial bounded, then

$$t_{mix}^{1}(\epsilon) = \text{poly}\log\left(N, \frac{1}{\varepsilon}\right), \qquad t_{mix}^{W_1}(\epsilon) = \text{poly}\log\left(\frac{1}{\epsilon}, \log N\right). \tag{18}$$

Such systems include notably all quantum CSS codes and classical Gibbs states that satisfy complete analyticity [Ces01, DS85, DS87], see Appendix A.4.2.

3 A comparison with quantum Gibbs sampling

Here we further investigate the connections of the current submission with the vast recent literature in quantum Gibbs sampling.

In the presence of a local commuting Hamiltonian, the natural candidate way to design quantum Gibbs samplers is employing a Davies generator, which is local in this case. In [KB16], the authors showed that there is an equivalence between the existence of a positive spectral gap for the Lindbladian (which yields fast mixing) and a certain form of strong clustering in the Gibbs state of the Hamiltonian. This generalizes to the quantum setting the works [SZ92a, SZ92b]. However, this is not enough to yield rapid mixing. [BCG+23, BCG+24] showed the existence of a positive (logarithmically-decaying with the system size) MLSI for translation-invariant 1D commuting local Hamiltonians at any positive temperature, and thus rapid mixing. This result was improved to constant MLSI in [KACR24], where rapid mixing was also shown for nearest neighbour Hamiltonians on hypercubic lattices at high-enough temperature, as well as for b-ary trees with small correlation length at high-enough temperature. This improves upon [CRF20], where another efficient Gibbs sampler was provided for the former setting in terms of the so-called Schmidt generators, based on [BV05].

The very recent papers [JI24, GCDK24] present quantum generalizations of Glauber and Metropolis dynamics, inspired by the breakthrough [TOV+11]. Whereas [JI24] uses quantum phase estimation (QPE) in its algorithm, [GCDK24] is built on the Lindbladian introduced in [CKG23, CKBG23], which can be regarded as a modification of Davies generator so that it is quasi-local for non-commuting Hamiltonians, in contrast to the usual Davies. A family of quantum Gibbs samplers satisfying the KMS detailed balance condition, which in particular includes the construction from [CKBG23], was presented in [DLL24]. Moreover, the concurrent [RFA24a, BLMT24] show the existence of a positive spectral gap, and thus fast mixing, for the former Lindbladians at high-enough temperature, yielding efficiency for the Gibbs samplers. In [RFA24b], this result was further extended to establish rapid mixing under the existence of Lieb-Robinson bounds for the underlying Hamiltonian. All these works appeared after long-term efforts to implement Davies generators for non-commuting Hamiltonians in the most accurate way possible, as an exact implementation is impossible. [WT23] dealt with this problem by assuming a rounding promise on the Hamiltonian, which was subsequently removed in [RWW23] by using randomized rounding.

Additionally to all these approaches based on Davies (or similar) generators, there are some others based on Grover approaches [PW09, CS17], quantum imaginary time evolution [MST+19] or quantum Singular Value Transforms [GSLW19]. Some other approaches based on dissipation are [ZBC23], where the quantum Gibbs sampler is designed with simple local update rules, or [FLT24], which is constructed from an energy functional inspired by the hypocoercivity of (classical) kinetic theory. Other works related to preparation of quantum Gibbs states are [KSM+24, HMS20, FFS24], among many others.

4 Outlook

The findings of this manuscript open up a series of natural questions in the contexts of entropy factorizations, mixing times, and decay of correlations on Gibbs states. As mentioned after Lemma 2.6, a strong entropy factorization is known to hold classically, see (6). Such a result presents an upper bound for the conditional relative entropy in ABC in terms of the sum of two conditional relative entropies in AB and BC, respectively, and a multiplicative error term, without the need for an additional additive error term. We strongly believe that something analogous should also hold in the quantum regime in quite some generality; however, a proof for this is still missing. Nevertheless, such a result is at least known to hold in the case that the second state is a tensor product [CLPG18b]. Such a strong entropy factorization would directly lead to 'strong' versions of the MLSI and TC inequality, with which one could directly conclude rapid mixing from the decay of MCMI by allowing decomposition of the lattice into local regions of constant size, following the argument of e.g. [Mar99, CRF20, KACR24].

Refocusing on the main approach developed in this paper, based on the existence of a weak AT, the ingredients required to conclude rapid mixing are decay of the MCMI and a suitable local gap. One would hope that a polynomially-bounded local gap, or more precisely an almost local gap (i.e. one can allow for constant size enlargements of the regions in the generator analogously to Lemma 2.12) should hold for the local Davies generators, assuming that the global Gibbs state satisfies some clustering condition, such as decay of the MCMI, or that it stems from a local CSS code. Proving this specifically for 2D Toric codes at every positive temperature, which we expect to hold, would imply an exponential improvement in the mixing time recently proven in [DLLZ24], inspired in the original [AFH09], and subsequently extended to 2D quantum double models in [KLCT16] and [LPGPH23].

Finally, another relevant line of research is that of the MCMI decay for more general models, i.e. general (commuting) Hamiltonians, not only marginal commuting ones, which should be of independent interest, given the connections of the MCMI to the notions of conditional mutual information and mutual information, and their range of applicability. Furthermore, it would be interesting to see if the methods translate also to the non-commuting setting directly relating an information-theoretic measure - the MCMI - to the speed of the convergence of the samplers, e.g. the one in [CKBG23].

5 Acknowledgments

J.K. and P.G. thank Samuel Scalet for helpful discussions regarding effective Hamiltonians, resulting in a much-simplified presentation of Theorem D.2. Both also would like to thank Sebastian Stengele for disclosing the marginal commuting property of CSS codes to them and for insights into this topic. A.C. and P.G. acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 470903074 - TRR 352. A.C. and P.G. also acknowledge funding by the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State Governments. This project was funded within the QuantERA II Programme which has received funding from the EU's H2020 research and innovation programme under the GA No 101017733. C.R. acknowledge financial support from the ANR project QTraj (ANR-20-CE40-0024-01) of the French National Research Agency (ANR) as well as France 2030 under the French National Research Agency award number "ANR-22-PNCQ-0002".

References

[AE11] Dorit Aharonov and Lior Eldar. On the complexity of commuting local hamiltonians, and tight conditions for topological order in such systems. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 334–343. IEEE, October 2011. doi:10.1109/focs.2011.58.

[AFH09] R Alicki, M Fannes, and M Horodecki. On thermalization in kitaev's 2d model. *Journal of Physics A: Mathematical and Theoretical*, 42(6):065303, January 2009. doi:10.1088/1751-8113/42/6/065303.

[AJK+21a] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence i: Modified log-sobolev inequalities for fractionally log-concave distributions and high-temperature ising models, 2021. doi:10.48550/ARXIV.2106.04105.

[AJK+21b] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence ii: Optimal sampling and concentration via restricted modified log-sobolev inequalities, 2021. doi:10.48550/ARXIV.2111.03247.

[AKV18] Dorit Aharonov, Oded Kenneth, and Itamar Vigdorovich. On the complexity of two dimensional commuting local hamiltonians. In Stacey Jeffery, editor, *13th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2018)*, volume 111 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:21, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.TQC.2018.2, doi:10.4230/LIPICS.TQC.2018.2.

[ALG20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1319–1330. IEEE, November 2020. doi:10.1109/focs46700.2020.00125.

[Alh23] Álvaro M. Alhambra. Quantum many-body systems in thermal equilibrium. *PRX Quantum*, 4(4):040201, November 2023. doi:10.1103/prxquantum.4.040201.

[AMS22] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the sherrington-kirkpatrick gibbs measure via algorithmic stochastic localization. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 323–334. IEEE, October 2022. doi:10.1109/focs54457.2022.00038.

[Bar17] Ivan Bardet. Estimating the decoherence time using non-commutative functional inequalities, 2017. doi:10.48550/ARXIV.1710.01039.

[BCG+23] Ivan Bardet, Ángela Capel, Li Gao, Angelo Lucia, David Pérez-García, and Cambyse Rouzé. Rapid thermalization of spin chain commuting hamiltonians. *Physical Review Letters*, 130(6):060401, February 2023. doi:10.1103/physrevlett.130.060401.

[BCG+24] Ivan Bardet, Ángela Capel, Li Gao, Angelo Lucia, David Pérez-García, and Cambyse Rouzé. Entropy decay for davies semigroups of a one dimensional quantum lattice. *Communications in Mathematical Physics*, 405(2), February 2024. doi:10.1007/s00220-023-04869-5.

[BCGL23] Sergey Bravyi, Giuseppe Carleo, David Gosset, and Yinchen Liu. A rapidly mixing markov chain from any gapped quantum many-body system. *Quantum*, 7:1173, November 2023. doi:10.22331/q-2023-11-07-1173.

[BCPH22] Andreas Bluhm, Ángela Capel, and Antonio Pérez-Hernández. Exponential decay of mutual information for gibbs states of local hamiltonians. *Quantum*, 6:650, February 2022. doi:10.22331/q-2022-02-10-650.

[BCPH24] Andreas Bluhm, Ángela Capel, and Antonio Pérez-Hernández. Strong decay of correlations for gibbs states in any dimension, 2024. doi:10.48550/ARXIV.2401.10147.

[BCR21] Ivan Bardet, Ángela Capel, and Cambyse Rouzé. Approximate tensorization of the relative entropy for noncommuting conditional expectations. *Annales Henri Poincaré*, 23(1):101–140, July 2021. doi:10.1007/s00023-021-01088-3.

[BF24] Emily Beatty and Daniel Stilck França. Order p quantum wasserstein distances from couplings, 2024. doi:10.48550/ARXIV.2402.16477.

[BGJM11] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC, May 2011. doi:10.1201/b10905.

[BH24] John Bostanci and Yeongwoo Hwang. Commuting local hamiltonians beyond 2d, 2024. doi:10.48550/ARXIV.2410.10495.

[BK18] Fernando G. S. L. Brandão and Michael J. Kastoryano. Finite correlation length implies efficient preparation of quantum thermal states. *Communications in Mathematical Physics*, 365(1):1–16, May 2018. doi:10.1007/s00220-018-3150-8.

[BLMT24] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. High-temperature gibbs states are unentangled and efficiently preparable, 2024. doi:10.48550/ARXIV.2403.16850.

[BV05] S. Bravyi and M. Vyalyi. Commutative version of the local hamiltonian problem and common eigenspace problem. *Quantum Information and Computation*, 5(3):187–215, May 2005. doi:10.26421/qic5.3-2.

[CB21] Chi-Fang Chen and Fernando G. S. L. Brandão. Fast thermalization from the eigenstate thermalization hypothesis, 2021. doi:10.48550/ARXIV.2112.07646.

[Ces01] Filippo Cesi. Quasi-factorization of the entropy and logarithmic sobolev inequalities for gibbs random fields. *Probability Theory and Related Fields*, 120(4):569–584, August 2001. doi:10.1007/pl00008792.

[CHH+24] Chi-Fang Chen, Jeongwan Haah, Jonas Haferkamp, Yunchao Liu, Tony Metger, and Xinyu Tan. Incompressibility and spectral gaps of random circuits, 2024. doi:10.48550/ARXIV.2406.07478.

[CHPZ24] Chi-Fang Chen, Hsin-Yuan Huang, John Preskill, and Leo Zhou. Local minima in quantum systems. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC '24, pages 1323–1330. ACM, June 2024. doi:10.1145/3618260.3649675.

[CKBG23] Chi-Fang Chen, Michael J. Kastoryano, Fernando G. S. L. Brandão, and András Gilyén. Quantum thermal state preparation, 2023. doi:10.48550/ARXIV.2303.18224.

[CKG23] Chi-Fang Chen, Michael J. Kastoryano, and András Gilyén. An efficient and exact noncommutative quantum gibbs sampler, 2023. doi:10.48550/ARXIV.2311.09207.

[CLPG18a] Angela Capel, Angelo Lucia, and David Perez-Garcia. Superadditivity of quantum relative entropy for general states. *IEEE Transactions on Information Theory*, 64(7):4758–4765, July 2018. doi:10.1109/tit.2017.2772800.

[CLPG18b] Ángela Capel, Angelo Lucia, and David Pérez-García. Quantum conditional relative entropy and quasi-factorization of the relative entropy. *Journal of Physics A: Mathematical and Theoretical*, 51(48):484001, November 2018. doi:10.1088/1751-8121/aae4cf.

[CLV23] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of glauber dynamics: Entropy factorization via high-dimensional expansion. *SIAM Journal on Computing*, pages STOC21–104–STOC21–153, October 2023. doi:10.1137/21m1443340.

[CM15] Raffaella Carbone and Andrea Martinelli. Logarithmic sobolev inequalities in non-commutative algebras. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 18(02):1550011, June 2015. doi:10.1142/s0219025715500113.

[CM17] Eric A. Carlen and Jan Maas. Gradient flow and entropy inequalities for quantum markov semigroups with detailed balance. *Journal of Functional Analysis*, 273(5):1810–1869, September 2017. doi:10.1016/j.jfa.2017.05.003.

[CRF20] Ángela Capel, Cambyse Rouzé, and Daniel Stilck França. The modified logarithmic sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions, 2020. doi:10.48550/ARXIV.2009.11817.

[CS96] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, August 1996. doi:10.1103/physreva.54.1098.

[CS17] Anirban N. Chowdhury and Rolando D. Somma. Quantum algorithms for gibbs sampling and hitting-time estimation. *Quantum Information and Computation*, 17(1 & 2):41–64, January 2017. doi:10.26421/qic17.1-2-3.

[CW16] Richard Cleve and Chunhao Wang. Efficient quantum algorithms for simulating lindblad evolution, 2016. doi:10.48550/ARXIV.1612.09512.

[Dav76] Edward B. Davies. *Quantum theory of open systems*. Acad. Press, London [u.a.], 1976. Literaturverz. S. 161 - 167.

[Dav79] E.B. Davies. Generators of dynamical semigroups. *Journal of Functional Analysis*, 34(3):421–432, December 1979. doi:10.1016/0022-1236(79)90085-5.

[DLL24] Zhiyan Ding, Bowen Li, and Lin Lin. Efficient quantum gibbs samplers with kubo–martin–schwinger detailed balance condition, 2024. doi:10.48550/ARXIV.2404.05998.

[DLLZ24] Zhiyan Ding, Bowen Li, Lin Lin, and Ruizhe Zhang. Polynomial-time preparation of low-temperature gibbs states for 2d toric code, 2024. doi:10.48550/ARXIV.2410.01206.

[DPMTL21] Giacomo De Palma, Milad Marvian, Dario Trevisan, and Seth Lloyd. The quantum wasserstein distance of order 1. *IEEE Transactions on Information Theory*, 67(10):6627–6643, October 2021. doi:10.1109/tit.2021.3076442.

[DPP24] Giacomo De Palma and Davide Pastorello. Quantum concentration inequalities and equivalence of the thermodynamical ensembles: an optimal mass transport approach, 2024. doi:10.48550/ARXIV.2403.18617.

[DPR22] Giacomo De Palma and Cambyse Rouzé. Quantum concentration inequalities. *Annales Henri Poincaré*, 23(9):3391–3429, April 2022. doi:10.1007/s00023-022-01181-1.

[DS85] R. L. Dobrushin and S. B. Shlosman. *Completely Analytical Gibbs Fields*, pages 371–403. Birkhäuser Boston, 1985. doi:10.1007/978-1-4899-6653-7_21.

[DS87] R. L. Dobrushin and S. B. Shlosman. Completely analytical interactions: Constructive description. *Journal of Statistical Physics*, 46(5–6):983–1014, March 1987. doi:10.1007/bf01011153.

[FFS24] Hamza Fawzi, Omar Fawzi, and Samuel O. Scalet. Certified algorithms for equilibrium states of local quantum hamiltonians. *Nature Communications*, 15(1), August 2024. doi:10.1038/s41467-024-51592-3.

[FLT24] Di Fang, Jianfeng Lu, and Yu Tong. Mixing time of open quantum systems via hypocoercivity, 2024. doi:10.48550/ARXIV.2404.11503.

[GCDK24] András Gilyén, Chi-Fang Chen, Joao F. Doriguello, and Michael J. Kastoryano. Quantum generalizations of glauber and metropolis dynamics, 2024. doi:10.48550/ARXIV.2405.20322.

[GJL20] Li Gao, Marius Junge, and Nicholas LaRacuente. Fisher information and logarithmic sobolev inequality for matrix-valued functions. *Annales Henri Poincaré*, 21(11):3409–3478, September 2020. doi:10.1007/s00023-020-00947-9.

[GJLL22] Li Gao, Marius Junge, Nicholas LaRacuente, and Haojian Li. Relative entropy decay and complete positivity mixing time, 2022. doi:10.48550/ARXIV.2209.11684.

[GR22] Li Gao and Cambyse Rouzé. Complete entropic inequalities for quantum markov chains. *Archive for Rational Mechanics and Analysis*, 245(1):183–238, May 2022. doi:10.1007/s00205-022-01785-1.

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC '19. ACM, June 2019. doi:10.1145/3313276.3316366.

[HJ24] Yeongwoo Hwang and Jiaqing Jiang. Gibbs state preparation for commuting hamiltonian: Mapping to classical gibbs sampling, 2024. doi:10.48550/ARXIV.2410.04909.

[HMS20] Aram W. Harrow, Saeed Mehraban, and Mehdi Soleimanifar. Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC '20, pages 378–386. ACM, June 2020. doi:10.1145/3357713.3384322.

[HS83] Geoffrey E Hinton and Terrence J Sejnowski. Optimal perceptual inference. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, volume 448, pages 448–453. Citeseer, 1983.

[HS07] Thomas P. Hayes and Alistair Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *The Annals of Applied Probability*, 17(3), June 2007. doi:10.1214/105051607000000104.

[JI24] Jiaqing Jiang and Sandy Irani. Quantum metropolis sampling via weak measurement, 2024. doi:10.48550/ARXIV.2406.16023.

[JLR22] Marius Junge, Nicholas Laracuente, and Cambyse Rouzé. Stability of logarithmic sobolev inequalities under a noncommutative change of measure. *Journal of Statistical Physics*, 190(2), December 2022. doi:10.1007/s10955-022-03026-x.

[KACR24] Jan Kochanowski, Alvaro M. Alhambra, Angela Capel, and Cambyse Rouzé. Rapid thermalization of dissipative many-body dynamics of commuting hamiltonians, 2024. doi:10.48550/ARXIV.2404.16780.

[KB16] Michael J. Kastoryano and Fernando G. S. L. Brandão. Quantum gibbs samplers: The commuting case. *Communications in Mathematical Physics*, 344(3):915–957, May 2016. doi:10.1007/s00220-016-2641-8.

[Kit03] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, January 2003. doi:10.1016/s0003-4916(02)00018-0.

[KKB20] Tomotaka Kuwahara, Kohtaro Kato, and Fernando G.S.L. Brandão. Clustering of conditional mutual information for quantum gibbs states above a threshold temperature. *Physical Review Letters*, 124(22):220601, June 2020. doi:10.1103/physrevlett.124.220601.

[KLCT16] Anna Kómár, Olivier Landon-Cardinal, and Kristan Temme. Necessity of an energy barrier for self-correction of abelian quantum doubles. *Physical Review A*, 93(5):052337, May 2016. doi:10.1103/physreva.93.052337.

[KSM+24] Vikram Kashyap, Georgios Styliaris, Sara Mouradian, Juan Ignacio Cirac, and Rahul Trivedi. Accuracy guarantees and quantum advantage in analogue open quantum simulation with and without noise, 2024. doi:10.48550/ARXIV.2404.11081.

[KT13] Michael J. Kastoryano and Kristan Temme. Quantum logarithmic sobolev inequalities and rapid mixing. *Journal of Mathematical Physics*, 54(5), May 2013. doi:10.1063/1.4804995.

[LCP13] Olivier Landon-Cardinal and David Poulin. Local topological order inhibits thermal stability in 2d. *Physical Review Letters*, 110(9):090502, February 2013. doi:10.1103/physrevlett.110.090502.

[LPGPH23] Angelo Lucia, David Pérez-García, and Antonio Pérez-Hernández. Thermalization in kitaev's quantum double models via tensor network techniques. *Forum of Mathematics, Sigma*, 11, 2023. doi:10.1017/fms.2023.98.

[LPW08] David Levin, Yuval Peres, and Elizabeth Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, December 2008. With a chapter by James G. Propp and David B. Wilson. doi:10.1090/mbk/058.

[LW22] Xiantao Li and Chunhao Wang. Simulating markovian open quantum systems using higher-order series expansion, 2022. doi:10.48550/ARXIV.2212.02051.

[MAMW15] Markus P. Müller, Emily Adlam, Lluís Masanes, and Nathan Wiebe. Thermalization and canonical typicality in translation-invariant quantum lattice systems. *Communications in Mathematical Physics*, 340(2):499–561, September 2015. doi:10.1007/s00220-015-2473-y.

[Mar99] Fabio Martinelli. *Lectures on Glauber Dynamics for Discrete Spin Models*, pages 93–191. Springer Berlin Heidelberg, 1999. doi:10.1007/978-3-540-48115-7_2.

[MST+19] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O'Rourke, Erika Ye, Austin J. Minnich, Fernando G. S. L. Brandão, and Garnet Kin-Lic Chan. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nature Physics*, 16(2):205–210, November 2019. doi:10.1038/s41567-019-0704-4.

[MSVC15] Andras Molnar, Norbert Schuch, Frank Verstraete, and J. Ignacio Cirac. Approximating gibbs states of local hamiltonians efficiently with projected entangled pair states. *Physical Review B*, 91(4):045138, January 2015. doi:10.1103/physrevb.91.045138.

[PW09] David Poulin and Pawel Wocjan. Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer. *Physical Review Letters*, 103(22):220502, November 2009. doi:10.1103/physrevlett.103.220502.

[RD19] Cambyse Rouzé and Nilanjana Datta. Concentration of quantum states from quantum functional and transportation cost inequalities. *Journal of Mathematical Physics*, 60(1), January 2019. doi:10.1063/1.5023210.

[RFA24a] Cambyse Rouzé, Daniel Stilck França, and Álvaro M. Alhambra. Efficient thermalization and universal quantum computing with quantum gibbs samplers, 2024. doi:10.48550/ARXIV.2403.12691.

[RFA24b] Cambyse Rouzé, Daniel Stilck França, and Álvaro M. Alhambra. Optimal quantum algorithm for gibbs state preparation, 2024. doi:10.48550/ARXIV.2411.04885.

[RGE12] Arnau Riera, Christian Gogolin, and Jens Eisert. Thermalization in nature and on a quantum computer. *Physical Review Letters*, 108(8):080402, February 2012. doi:10.1103/physrevlett.108.080402.

[Rou24] Cambyse Rouzé. *Logarithmic Sobolev Inequalities for Finite Dimensional Quantum Markov Chains*, pages 263–321. Springer Nature Switzerland, 2024. doi:10.1007/978-3-031-50466-2_6.

[RWW23] Patrick Rall, Chunhao Wang, and Pawel Wocjan. Thermal state preparation via rounding promises. *Quantum*, 7:1132, October 2023. doi:10.22331/q-2023-10-10-1132.

[SL78] Herbert Spohn and Joel L. Lebowitz. Irreversible thermodynamics for quantum systems weakly coupled to thermal reservoirs. *Advances in Chemical Physics*, pages 109–142, January 1978. doi:10.1002/9780470142578.ch2.

[SM21] Oles Shtanko and Ramis Movassagh. Preparing thermal states on noiseless and noisy programmable quantum processors, 2021. doi:10.48550/ARXIV.2112.14688.

[Ste96] Andrew Steane. Multiple-particle interference and quantum error correction. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 452(1954):2551–2577, November 1996. doi:10.1098/rspa.1996.0136.

[SZ92a] Daniel W. Stroock and Boguslaw Zegarlinski. The equivalence of the logarithmic sobolev inequality and the dobrushin-shlosman mixing condition. *Communications in Mathematical Physics*, 144(2):303–323, February 1992. doi:10.1007/bf02101094.

[SZ92b] Daniel W. Stroock and Boguslaw Zegarlinski. The logarithmic sobolev inequality for discrete spin systems on a lattice. *Communications in Mathematical Physics*, 149(1):175–193, September 1992. doi:10.1007/bf02096629.

[TOV+11] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete. Quantum metropolis sampling. *Nature*, 471(7336):87–90, March 2011. doi:10.1038/nature09770.

[Ume62] Hisaharu Umegaki. Conditional expectation in an operator algebra. iv. entropy and information. *Kodai Mathematical Journal*, 14(2), January 1962. doi:10.2996/kmj/1138844604.

[vAGGdW20] Joran van Apeldoorn, András Gilyén, Sander Gribling, and Ronald de Wolf. Quantum sdp-solvers: Better upper and lower bounds. *Quantum*, 4:230, February 2020. doi:10.22331/q-2020-02-14-230.

[WT23] Pawel Wocjan and Kristan Temme. Szegedy walk unitaries for quantum maps. *Communications in Mathematical Physics*, 402(3):3201–3231, July 2023. doi:10.1007/s00220-023-04797-4.

[ZBC23] Daniel Zhang, Jan Lukas Bosse, and Toby Cubitt. Dissipative quantum gibbs sampling, 2023. doi:10.48550/ARXIV.2304.04526.

A Theoretical Framework

A.1 Basic notation

We work with finite-dimensional Hilbert spaces, denoted by $\mathcal{H}$ or $\mathcal{K}$, and their associated algebras of bounded operators, $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{K})$, respectively. Throughout this discussion, we use several standard notations: the matrix trace is denoted by $\operatorname{tr}[\,\cdot\,]$, the Hilbert-Schmidt inner product by $\langle\cdot,\cdot\rangle$, and the Hilbert-Schmidt norm by $\|\cdot\|_2$. More generally, we use $\|\cdot\|_p$ for $p\in[1,\infty]$ to denote the Schatten-p-norms, with the operator norm given by $\|\cdot\|_\infty$. Operators are represented by capital Latin or lowercase Greek letters, depending on the context. For any operator $X\in\mathcal{B}(\mathcal{H})$, we write $X^\dagger$ for its adjoint and denote the subset of self-adjoint operators by $\mathcal{B}_{\mathrm{sa}}(\mathcal{H})$. The Hilbert-Schmidt adjoint of a linear map $\Phi:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H})$ is written as Φ^*. We reserve $I_\mathcal{H}$ for the identity operator on $\mathcal{H}$ and id for the identity map on $\mathcal{B}(\mathcal{H})$, with $d_\mathcal{H}=\dim(\mathcal{H})$ denoting the dimension of $\mathcal{H}$. A quantum state (density operator) ρ is a positive semidefinite operator with unit trace, and we denote the set of states on $\mathcal{H}$ by $\mathcal{S}(\mathcal{H})$.

For any full-rank $X\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H})$, we define two fundamental linear maps: the first is $\Gamma_X:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H})$, given by $\Gamma_X(Y):=X^{\frac{1}{2}}YX^{\frac{1}{2}}$, and the second is the modular operator $\Delta_X:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H})$, defined as $\Delta_X(Y)=XYX^{-1}$. In bipartite systems AB with $\mathcal{H}_{AB}=\mathcal{H}_A\otimes\mathcal{H}_B$, we use ρ_A to denote the reduced state $\operatorname{tr}_B[\rho_{AB}]$, where tr_B represents the partial trace over system B. By convention, we may write ρ_A as an operator in $\mathcal{B}(\mathcal{H}_{AB})$, to be understood as $\rho_A\otimes I_B$. The normalized partial trace $\mathbb{E}_B:\mathcal{B}(\mathcal{H}_{AB})\to\mathcal{B}(\mathcal{H}_{AB})$ is defined as $\mathbb{E}_B(X_{AB})=\operatorname{tr}_B[X_{AB}]\otimes I_B\,/d_B$.

For a bipartite state $\rho\in\mathcal{S}(\mathbb{C}^n\otimes\mathcal{H})$ and a quantum channel $\Psi:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H})$, we write $\Psi(\rho):=(\mathrm{id}\otimes\Psi)(\rho)$, where a quantum channel is understood to be a linear completely positive trace preserving map $\Phi:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{K})$. The Umegaki relative entropy [Ume62], a fundamental measure in quantum information theory, is defined for states ρ and σ as:

$$D(\rho\|\sigma):=\begin{cases}\operatorname{tr}[\rho\log\rho-\rho\log\sigma], & \text{if } \ker(\sigma)\subseteq\ker(\rho),\\ +\infty, & \text{otherwise,}\end{cases}$$

where $\ker(\cdot)$ denotes the kernel of an operator. By Pinsker's inequality, the relative entropy relates to the trace distance as:

$$\|\rho-\sigma\|_1\leq\sqrt{2D(\rho\|\sigma)}\,. \tag{19}$$

For composite (or many-body) systems $\mathcal{H}_\mathcal{I}=\bigotimes_{i\in\mathcal{I}}\mathcal{H}_i$, we introduce the quantum Wasserstein distance of order 1 [DPMTL21]. For two quantum states $\rho,\sigma\in\mathcal{S}(\mathcal{H}_\mathcal{I})$, this distance is defined as the dual to a Lipschitz norm:

$$\|\rho-\sigma\|_{W_1}:=\max_{H\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_\mathcal{I}),\|H\|_L\leq1}\langle H,\rho-\sigma\rangle\ , \tag{20}$$

where the Lipschitz norm is defined as

$$\|H\|_L:=2\max_{i\in\mathcal{I}}\min_{\tilde{H}\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_{\mathcal{I}\setminus\{i\}})}\left\|H-I_i\otimes\tilde{H}\right\|_\infty, \tag{21}$$

for $H\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_\mathcal{I})$. The Wasserstein distance can also be regarded as a norm on the set of traceless self-adjoint operators and is an extensive quantity, scaling with the size of $\mathcal{I}$. This is reflected in the following relation with the trace distance, shown in [DPMTL21, Proposition 5]: For any traceless $X\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_\mathcal{I})$ with $\operatorname{tr}_\mathcal{J}[X]=0$, where $\mathcal{J}\subseteq\mathcal{I}$, we have

$$\|X\|_{W^1}\leq|\,\mathcal{J}\,|\|X\|_1,. \tag{22}$$

Conditional expectations on von Neumann algebras play a central role in our analysis. For a von Neumann subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$, a conditional expectation onto $\mathcal{N}$ is a completely positive unital map $E^*_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ satisfying:

1. $E^*_{\mathcal{N}}(X) = X$ for all $X \in \mathcal{N}$;
2. $E^*_{\mathcal{N}}(VXW) = VE^*_{\mathcal{N}}(X)W$ for all $V, W \in \mathcal{N}$ and $X \in \mathcal{B}(\mathcal{H})$.

For such conditional expectations, the relative entropy exhibits special properties known as the chain rule and exact entropy factorization: given conditional expectation $E^*_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ and $E^*_{\mathcal{M}} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ with $[E^*_{\mathcal{M}}, E^*_{\mathcal{N}}] = 0$, the chain rule states that for any state $\sigma = E_{\mathcal{N}}(\sigma) \in \mathcal{S}(\mathcal{H})$, and any state $\rho \in \mathcal{S}(\mathcal{H})$,

$$D(\rho\|\sigma) = D(\rho\|E_{\mathcal{N}}(\rho)) + D(E_{\mathcal{N}}(\rho)\|\sigma)\,. \tag{23}$$

The exact entropy factorisation on the other hand states that

$$D(\rho\|E_{\mathcal{N}}E_{\mathcal{M}}(\rho)) \leq D(\rho\|E_{\mathcal{M}}(\rho)) + D(\rho\|E_{\mathcal{N}}(\rho))\,. \tag{24}$$

For a full-rank state $\sigma > 0$, we define a family of inner products, $s \in [0, 1]$,

$$\langle X, Y\rangle_{s,\sigma} := \left\langle X, \sigma^s Y \sigma^{1-s}\right\rangle\,,$$

turning $\mathcal{B}(\mathcal{H})$ into a Hilbert space with norm

$$\|X\|^2_{1/s,\sigma} := \langle X, X\rangle_{s,\sigma}\,.$$

In the cases $s = 1/2$ and $s = 1$ ($s = 0$), $\langle X, Y\rangle_{s,\sigma}$ are also referred to as the KMS and GNS inner products, respectively. Note that for the KMS inner product, we will simply write $\langle \cdot, \cdot\rangle_\sigma$. Given the Hilbert space structure, we can define adjoints of linear maps $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, calling a map Φ KMS- or GNS-symmetric if it is self-adjoint with respect to the respective inner product. Note that GNS symmetry implies KMS symmetry, as it implies commutativity with the modular operator, $[\Phi, \Delta_\sigma] = 0$. These inner products play a crucial role in the analysis of quantum Markov semigroups (QMS). A QMS $(\mathcal{P}_t)_{t\geq 0}$ is a semigroup of completely positive, trace-preserving maps on $\mathcal{B}(\mathcal{H})$ in Schrödinger and of completely positive, unital maps on $\mathcal{B}(\mathcal{H})$ in Heisenberg picture. Note that one is the dual of the other w.r.t. the Hilbert-Schmidt inner-product. The unique generator of this semigroup, called a Lindbladian, is denoted by $\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ and connected to the semigroup through the linear differential equation

$$\frac{d}{dt}\rho = \mathcal{L}(\rho) \quad \text{with} \quad \rho(0) = \rho_0 \in \mathcal{S}(\mathcal{H})\,.$$

This motivates the notation $\mathcal{P}_t = e^{t\,\mathcal{L}}$, which we will use henceforth. A QMS is said to be KMS- or GNS-symmetric with respect to a full-rank state $\sigma \in \mathcal{S}(\mathcal{H})$ if the Hilbert-Schmidt dual of its generator $\mathcal{L}^* : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ exhibits the respective symmetry. For a generator of a QMS $\mathcal{L}^*$ which is further GNS-symmetric w.r.t $\sigma > 0$ the kernel $\ker(\mathcal{L}^*)$ is a von Neuman subalgebra with conditional expectation E^* given by $E^*[X] := \lim_{t\to\infty} e^{t\,\mathcal{L}^*}[X]$ for any $X \in \mathcal{B}(\mathcal{H})$.

A central question, and the primary focus of our work, concerns the speed of this convergence. Specifically, we investigate the convergence rate of a particular family of Lindbladians on quantum spin systems, measured in both the Wasserstein and trace distance.

A.2 Mixing, rapid mixing, and logarithmic Sobolev inequalities for open quantum systems

The study of convergence speeds for quantum Markov semigroups under different distance measures forms the core of this research. Here, we use the term "measure" informally as a way to compare states, rather than in its strict mathematical sense. While mixing phenomena are interesting even in simple systems, their study primarily focuses on composite systems. Therefore, we will work in the setting $\mathcal{H} = \bigotimes_{i\in\mathcal{I}} \mathcal{H}_i$. We will restrict ourselves to quantum Markov semigroups that are GNS-symmetric with respect to a full-rank state $\sigma > 0$, which guarantees both the convergence $e^{t\,\mathcal{L}} \to E$ as $t \to \infty$ and the von Neumann algebra structure of $\ker(\mathcal{L}^*)$. The mixing time in trace distance for $\varepsilon > 0$ is defined as

$$t^1_{\text{mix}}(\varepsilon) := \inf\{t \geq 0 : \forall \rho \in \mathcal{S}(\mathcal{H}), \left\|e^{t\,\mathcal{L}}(\rho) - E(\rho)\right\|_1 \leq \varepsilon\}\,. \tag{25}$$

Analogously, we define the mixing time in Wasserstein distance as

$$t^{W_1}_{\text{mix}}(\varepsilon) := \inf\{t \geq 0 : \forall \rho \in \mathcal{S}(\mathcal{H}), \left\|e^{t\,\mathcal{L}}(\rho) - E(\rho)\right\|_{W^1} \leq |\,\mathcal{I}\,|\varepsilon\}\,. \tag{26}$$

Our analysis aims to understand how these mixing times depend on both system size and ε. We classify systems as *mixing*, sometimes in literature also referred to as 'fast mixing', when mixing times scale linearly with system size, while polylogarithmic scaling characterizes *rapidly mixing* or *rapid mixing* systems and quasi-logarithmic scaling characterized *quasi-rapid mixing*. Previous research has largely focused on trace distance, commonly analyzing the generator's gap (c.f. [KB16, CKBG23, CHH+24, BCGL23])—the distance between the largest and second-largest eigenvalue of the Lindbladian. For KMS-symmetric Lindbladians, this gap is given by

$$\lambda(\mathcal{L}^*) := \inf_{X\in\mathcal{B}(\mathcal{H})} \frac{\langle X, -\,\mathcal{L}^*(X)\rangle_\sigma}{\|(\mathrm{id} - E^*)(X)\|_{2,\sigma}}\,. \tag{27}$$

By Grönwall's Lemma, we obtain $\left\|e^{t\,\mathcal{L}^*}(X) - E^*(X)\right\|_{2,\sigma} \leq e^{-t\lambda(\mathcal{L}^*)}\|X - E^*(X)\|_{2,\sigma}$. Using Hölder's inequality on its variational form, this yields an estimate for the trace distance:

$$\left\|e^{t\,\mathcal{L}}(\rho) - \sigma\right\|_1 \leq \left\|\sigma^{-1/2}\right\|_\infty \sup_{X\in\mathcal{B}_{\text{sa}}(\mathcal{H}_\mathcal{I}), \|X\|_\infty\leq 1} \left\|e^{t\,\mathcal{L}^*}(X) - E^*(X)\right\|_{2,\sigma} \leq 2\left\|\sigma^{-1/2}\right\|_\infty e^{-t\lambda(\mathcal{L}^*)}\,.$$

Since $\left\|\sigma^{-1}\right\|_\infty$ typically scales exponentially with $|\,\mathcal{I}\,|$, the mixing times must be at least linear in $|\,\mathcal{I}\,|$ to compensate, assuming the gap is independent of $|\,\mathcal{I}\,|$. Alternative approaches have focused on analyzing decay in relative entropy (c.f. [CM15, KT13, Bar17, BCG+24]), attempting to establish (weak) modified logarithmic Sobolev inequalities (wMLSI). These inequalities relate the relative entropy to the entropy production of the semigroup's generator, defined as

$$\mathrm{EP}_\mathcal{L}(\rho) := -\frac{d}{dt}\Big|_{t=0} D(e^{t\,\mathcal{L}}(\rho)\|\sigma) = -\operatorname{tr}[\mathcal{L}(\rho)(\log(\rho) - \log(\sigma))]\,. \tag{28}$$

The entropy production is always non-negative, following from monotonicity of the relative entropy under quantum channels. We say that $\mathcal{L}$ satisfies a (weak) modified logarithmic Sobolev inequality if there exist constants $c_1 > 0$ and $c_2 \geq 0$ such that for all $\rho \in \mathcal{S}(\mathcal{H})$

$$D(\rho\|E(\rho)) \leq c_1\,\mathrm{EP}_\mathcal{L}(\rho) + c_2 \tag{wMLSI}$$

Applying Grönwall's Lemma yields

$$D(e^{t\,\mathcal{L}}(\rho)\|E(\rho)) \leq e^{-t/c_1} D(\rho\|\sigma) + c_2\,,$$

and then by Pinsker's inequality (19), we obtain an estimate on the trace distance:

$$\|\rho - \sigma\|_1 \leq \sqrt{2e^{-t/c_1} D(\rho\|\sigma) + 2c_2}\,,.$$

Since $D(\rho\|\sigma) \leq \log \|\sigma^{-1}\|_\infty$, in the best case—assuming constant c_1 and $c_2 = 0$—the mixing time need only be logarithmic in system size, classifying the system as rapid mixing. The novelty of our proofs lies in accepting positive c_2 which, however, decays with system size. As with many quantum-mechanical capacity quantities, there exists a notion of complete MLSI (cMLSI), defined as the supremum of MLSI of $\mathrm{id}_R \otimes \mathcal{L}$ over arbitrary reference systems R, denoted by $\alpha_c(\mathcal{L})$. This notation yields the relation

$$\alpha_c(\mathcal{L}_A \otimes \mathrm{id}_B + \mathrm{id}_A \otimes \mathcal{L}_B) \geq \min\{\alpha_c(\mathcal{L}_A), \alpha_c(\mathcal{L}_B)\}\,,$$

which was proven in [GJL20] and is known for classical MLSI. Interestingly enough a relation between the gap of a GNS-symmetric QMS and the cMLSI was recently proven in [GR22, GJLL22, Rou24] and we want to quote this estimate here as we later will use it in our analysis:

$$\alpha_c(\mathcal{L}) > \frac{\lambda(\mathcal{L})}{2\log(10 C_{\mathrm{cb}}(E))}\,, \tag{29}$$

where $C(E) := \inf\{c \geq 0 : \rho \leq cE(\rho), \forall \rho \in \mathcal{S}(\mathcal{H})\}$, $C_{\mathrm{cb}}(E) := \sup_R C(\mathrm{id}_R \otimes E)$ is the (complete) Pimsner-Popa index, respectively.

A.3 Local commuting Hamiltonian and Davies generators on lattices

In this paper, we consider D-dimensional hypercubes $\Lambda := \Lambda_L := [\![-L, L]\!]^D$ of side length $2L+1$ consisting of $N = |\Lambda|$ sites, where each site hosts a qudit system $\mathcal{H}_i = \mathbb{C}^d$. The Hilbert space of the complete system is thus $\mathcal{H}_\Lambda = \bigotimes_{k\in\Lambda} \mathcal{H}_k$, and for any subsystem $A \subseteq \Lambda$, we have $\mathcal{H}_A = \bigotimes_{k\in A} \mathcal{H}_i$. In isolation, the system's dynamics are fully characterized by its Hamiltonian $H_\Lambda \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H})$. When in thermal equilibrium with a heat bath at fixed inverse temperature $\beta > 0$, the system eventually assumes its thermal state, also known as the Gibbs state:

$$\sigma := \sigma^\Lambda := \frac{e^{-\beta H_\Lambda}}{\mathrm{tr}[e^{-\beta H_\Lambda}]}\,.$$

Our analysis focuses on a special class of Hamiltonians that are (κ, r)-local and commuting, with bounded interaction strength and connectivity. Specifically, there exists a representation

$$H_\Lambda = \sum_{A\subseteq\Lambda} h_A \otimes I_{\overline{A}}$$

where $\overline{A} := \Lambda \backslash A$ denotes the complement in Λ and $h_A \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}_A)$, satisfying the following conditions:

1. Commutativity: All terms $h_A \otimes I_{\overline{A}}$, $A \subseteq \Lambda$ pairwise commute;
2. (κ, r)-locality: $h_A = 0$ if either $|A| > \kappa$ or $\mathrm{diam}(A) > r$;
3. Bounded interaction strength: $\max_{A\subseteq\Lambda} \|h_A\|_\infty =: J < \infty$;
4. Bounded connectivity: $\max_{k\in\Lambda} |\{A \subseteq \Lambda : h_A \neq 0, k \in A\}| =: g < \infty$.

Note that the *growth constant* g is a function of κ, r and D, and that κ, r are not independent. Despite that, we will treat them in this work as separate parameters.
When considering a family of Hamiltonians on increasing lattice sizes, such as those arising from an interaction on $\mathbb{Z}^D$, all the above constants are implicitly assumed to be independent of system size. In what follows, we will suppress the identity and simply write h_A, where the subindex indicates the support of the Hamiltonian term. For any set $R \subseteq \Lambda$, we define its closure dictated by the interaction structure of the Hamiltonian as

$$R\partial := \{k \in \Lambda : \exists A \subseteq \Lambda, h_A \neq 0, A \cap R \neq \emptyset, k \in A\}$$

and its boundary as $\partial R := R\partial \backslash R$. For our analysis, we introduce local Hamiltonians and local Gibbs states. The local Hamiltonian of $R \subseteq \Lambda$ is defined as

$$H_R := \sum_{A \subseteq R} h_A$$

and is strictly supported in R. Its corresponding local Gibbs state is $\sigma^R := \frac{e^{-\beta H_R}}{\operatorname{tr}[e^{-\beta H_R}]}$, which should not be confused with the marginal Gibbs state on R given by $\sigma_R := \operatorname{tr}_{\overline{R}}[\sigma]$. Using the locality and connectivity constraints, one readily obtains

$$\|H_R\|_\infty \leq gJ|R|\,, \tag{30}$$

which will prove useful in subsequent derivations.

We previously mentioned that, when a system is brought into contact with a heat bath, it is assumed to thermalize and eventually conform to the Gibbs state statistics. Assuming a specific interaction form between the bath and the system, i.e., a combined Hamiltonian of the form

$$H = H_\Lambda + H^{\mathrm{HB}} + \sum_{k \in \Lambda, \alpha} S_{\alpha,k} \otimes B_{\alpha,k}\,, \tag{31}$$

where H^{HB} is the bath Hamiltonian and $\{S_{\alpha,k} \otimes B_{\alpha,k}\}_{k \in \Lambda, \alpha}$ are self-adjoint single-site couplings one can derive an effective system dynamic through a weak coupling limit [SL78]. Note that we will not only assume that the $\{S_{\alpha,k}\}_\alpha$ are self-adjoint but for every k form a Kraus decomposition of the partial trace on the site k, i.e. $\sum_\alpha S_{\alpha,k} \cdot S_{\alpha,k} = \operatorname{tr}_k[\,\cdot\,]$. An example would be the Pauli operators on $\mathbb{C}^2$. This choice ensures that in the trivial case $H_\Lambda = 0$ one recovers the depolarising semigroup. The effective dynamics obtained by the weak-coupling limits are governed by the so-called Davies Lindbladien

$$\mathcal{L}_\Lambda^{D,H}(\rho) := -i[H_\Lambda, \rho] + \mathcal{L}_\Lambda^D(\rho) \; = -i[H_\Lambda, \rho] + \sum_{k \in \Lambda} \mathcal{L}_k^D(\rho)\,, \tag{32}$$

for some local generators

$$\mathcal{L}_k^D(\rho) = \sum_{\omega,\alpha} \chi_{\alpha,k}^{\beta,\omega} \left(S_{\alpha,k}^\omega \rho S_{\alpha,k}^{\omega,\dagger} - \frac{1}{2} \left\{ \rho, S_{\alpha,k}^{\omega,\dagger} S_{\alpha,k}^\omega \right\} \right). \tag{33}$$

The sum in (33) ranges over the index α of the local basis $\{S_{\alpha,k}\}$ as well as the Bohr frequencies ω of the Hamiltonian $H_{k\partial}$, i.e. all pairwise differences of its eigenvalues that arise as a result of the decomposition

$$e^{-itH_\Lambda}\, S_{\alpha,k} e^{itH_\Lambda} = e^{-itH_{k\partial}}\, S_{\alpha,k} e^{itH_{k\partial}} = \sum_\omega e^{it\omega} S_{\alpha,k}^\omega\,, \tag{34}$$

for arbitrary $t \in \mathbb{R}$. Note that the validity of the first equality relies on the commuting, local nature of the Hamiltonian, leading to the cancellation of all Hamiltonian terms that do not intersect k, thereby localizing the Bohr frequencies (i.e., their dependence only on $H_{k\partial}$). For better readability, we will not make the dependence on k explicit, however. Lastly the $\chi^{\beta,\omega}_{\alpha,k}$ in (33) are the Fourier coefficients of the two-point correlation functions of the environment, which we assume to be uniformly bounded above and below: $0 < \chi^{\beta}_{\min} \leq \chi^{\beta,\omega}_{\alpha,k} \leq \chi^{\beta}_{\max}$. By assumptions made in the weak coupling limit, those coefficients satisfy the KMS condition, i.e. $\chi^{\beta,-\omega}_{\alpha,k} = e^{-\beta\omega}\chi^{\beta,\omega}_{\alpha,k}$ making the $\mathcal{L}^D_k$, GNS-symmetric w.r.t. to the Gibbs state of the same temperature. Similarly to the local Hamiltonians, we can define the Lindbladian on a set $R \subseteq \Lambda$ by $\mathcal{L}^D_R$ by restricting the sum in (32) to R. Unlike the local Hamiltonians, however, the local Lindbladiens act non-trivially on $R\partial$ instead of R. One can define semigroups from these local Lindbladiens that, due to their GNS-symmetry also converge to projections respectively conditional expectation in the Heisenberg picture, we will denote by

$$E^*_A := \lim_{t\to\infty} e^{t\,\mathcal{L}^{D,*}_A} \quad \text{or equivalently} \quad E_A := \lim_{t\to\infty} e^{t\,\mathcal{L}^D_A}\,.$$

Note that these maps satisfy the following properties [BCR21, BCG+24, KB16]:

1. $E_\emptyset = \mathrm{id}$ and $E_\Lambda = \mathrm{tr}[\,\cdot\,]\,\sigma^\Lambda$;
2. for each $A \subset A\partial \subseteq B \subseteq \Lambda$, σ^B is a fixed point of E_A;
3. for each $A \subseteq B \subseteq \Lambda$, $E_A E_B = E_B E_A = E_B$;
4. for any two regions $A, B \subseteq \Lambda$ such that $A\partial \cap B\partial = \emptyset$, $E_{A\sqcup B} = E_A E_B = E_B E_A$,

that will become useful later. Note that here and in the following we will use $\sqcup$ to denote a disjoint union of sets. By using the representation

$$E_A = \lim_{t\to\infty} \mathcal{R}^n_{\sigma,A}\,, \tag{35}$$

in terms of the Petz recovery

$$\mathcal{R}_{A,\sigma}(X) := \sigma^{1/2}(\sigma_{\overline{A}}^{-1/2}\,\mathrm{tr}_A[X]\sigma_{\overline{A}}^{-1/2}) \otimes I_A\ \sigma^{1/2}\,,$$

proven in [BCR21, Theorem 1] and the locality of the Hamiltonian one readily verifies

$$E_A\,\mathbb{E}_A = E_A \quad \text{and} \quad \mathbb{E}_{A\partial}\,E_A = \mathbb{E}_{A\partial}\ . \tag{36}$$

Although $\mathcal{L}^D_A$ and E_A for $A \subseteq \Lambda$, as well as σ^A and σ_A, depend on the temperature, we will only make this dependence explicit where necessary to avoid a cluttering of indices.

A.4 Decay of correlations in Gibbs states

The theory of thermalization in classical spin systems [Mar99] suggests a strong connection between equilibrium and non-equilibrium properties of quantum spin systems. In classical systems, a well-established principle links the convergence rate of local stochastic dynamics to the decay of correlations in the Gibbs equilibrium measure. Specifically, the Glauber dynamics, which models the thermalization of a discrete spin system, rapidly approaches the equilibrium Gibbs measure if and only if correlations between spatially separated regions exhibit exponential decay with distance. Here, we consider a different notion of the decay of correlations and relate it to more standard classical notions.

A.4.1 Correlation measures and the MCMI

The notion we want to introduce is that of the matrix-valued quantum conditional mutual information or MCMI for short. It has previously appeared as a tool to prove the decay of correlations in Gibbs states. More precisely the decay of the (conditional) mutual information [KKB20]. This result, unfortunately, contains a mistake, meaning we cannot recycle the decay of MCMI from there and have to rely on the toolbox of effective Hamiltonians, introduced later in Appendix A.4.3.

Definition A.1 (Uniform exponential decay of matrix-valued quantum conditional mutual information) Given any partition $\Lambda = A \sqcup B \sqcup C \sqcup D$ of the lattice, we define the *matrix-valued quantum conditional mutual information* (MCMI) of a state σ as

$$\mathbf{H}_\sigma(A:C|D) := \log \sigma_{ACD} + \log \sigma_D - \log \sigma_{AD} - \log \sigma_{CD} \,. \tag{37}$$

Then the Gibbs state $\sigma = \sigma^\Lambda$ is said to satisfy *uniform exponential decay of its matrix-valued quantum conditional mutual information* with constants K, ξ if there exist constants $K, \xi > 0$ independent of the regions A, B, C, D and such that

$$H_\sigma(A:C|D) := \|\mathbf{H}_\sigma(A:C|D)\|_\infty \leq K|\Lambda| e^{-\frac{\mathrm{dist}(A,C)}{\xi}} \,. \tag{38}$$

Note that under slight abuse of notation, we will be referring to both $\mathbf{H}$ and H as the MCMI. To illustrate this decay and its geometric aspects, let us examine three two-dimensional situations in Figure 4 which, despite their simplicity, capture all use cases later. The leftmost and the rightmost can be regarded as the extremal cases while the one in the middle in some sense interpolates the two. What all of them have in common and what is the crucial property is the separation of the regions A and C. This can be achieved either by averaging as in the case of B (operationally a partial trace) or through conditioning as in the case of D.

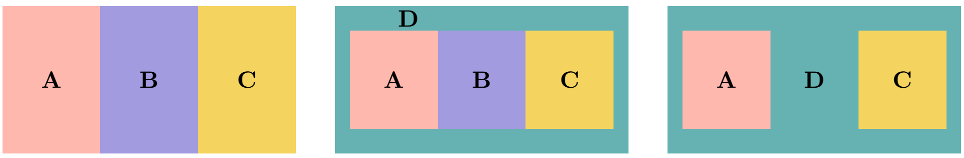


Figure 4: A lattice Λ is partitioned into four distinct regions, such that A and C are separated by either B or D or both. B is a system that is averaged over (through a partial trace) while D we condition on.

The system size-dependent prefactor we impose in our examples can be replaced by one that only depends on the minimum sizes of A and C (see Appendix D). For our purposes, the linear dependence is, however, sufficient. As the work in [KKB20] suggests the MCMI is closely tied to other correlation measures. More precisely: Given MCMI decay, one in particular has the decay of the conditional mutual information

$$\begin{aligned} I_\sigma(A:C|D) &:= -S(\sigma_{ACD}) - S(\sigma_D) + S(\sigma_{CD}) + S(\sigma_{AD}) \\ &= \mathrm{tr}[\sigma_{ACD}\, \mathbf{H}_\sigma(A:C|D)], \end{aligned}$$

the mutual information

$$\begin{aligned} I_\sigma(A:C) &:= -S(\sigma_{AC}) + S(\sigma_A) + S(\sigma_C) \\ &= \mathrm{tr}[\sigma_{AC}\, \mathbf{H}_\sigma(A:C|\emptyset)], \end{aligned}$$

and in consequence also the decay of the covariance (see [BCPH22]). Note that $S(\sigma) := -\operatorname{tr}[\sigma \log \sigma]$ here denotes the von Neuman entropy. Besides its connection to quantum information theoretic quantities, the MCMI also is connected to a notion of decay of correlation from classical statistical physics, namely complete analyticity. This comes as no big surprise since complete analyticity in classical systems implies a logarithmic Sobolev inequality and thereby also rapid mixing [Ces01, Theorem 4.1].

A.4.2 MCMI decay from complete analyticity in classical systems

Let us more closely investigate the connection to the *complete analyticity* also called the Dobrushin-Shlosmann condition of classical spin systems [DS85, DS87, Ces01]. We will show that indeed complete analyticity implies uniform decay of the MCMI. Although we believe that also the reverse implication holds, we are still missing the proof as of the day of writing.
Note that in this section we will follow the notation of [Mar99] mostly (re)defining the necessary objects ad-hoc. We will further suppress all temperature dependence to improve readability. Let us consider the single spin system $S = \{-1, 1\}$ with counting measure leading to a full state space $\Omega = \{\omega : \mathbb{Z}^D \to S\}$. We will write $\omega_V = \omega|_V$ for the restriction of $\omega \in \Omega$ to V. The interaction which completely characterises the model is given as $J : \{V \Subset \mathbb{Z}^D\} \to \mathbb{R}$ and we write J_V instead of $J(V)$. Here and in the following $V \Subset \mathbb{Z}^D$ refers to finite subsets. The corresponding Hamiltonian on $V \Subset \mathbb{Z}^D$ is given by

$$H_V(\omega) = -\sum_{A:\, A \cap V \neq \emptyset} J_A \prod_{x \in A} \omega(x)\,,$$

and the Gibbs measure for the boundary condition $\tau \in \Omega$ as

$$\mu_V^\tau(\omega) = \begin{cases} (Z_V^\tau)^{-1} \exp[-H_V(\omega)] & \text{if } \omega_{\overline{V}} = \tau_{\overline{V}}\,, \\ 0 & \text{else.} \end{cases}$$

Here and in the rest of this section a line over a set denotes the complement w.r.t. $\mathbb{Z}^D$, e.g. $\overline{V} = \mathbb{Z}^D \setminus V$ in the above definition. Interestingly for $V \subset \Lambda \Subset \mathbb{Z}^D$, we find that μ_V^τ alternatively can be written as

$$\mu_V^\tau(\omega) = \begin{cases} (\tilde{Z}_V^\tau)^{-1} \exp[-H_\Lambda(\omega)] & \text{if } \omega_{\overline{V}} = \tau_{\overline{V}}\,, \\ 0 & \text{else.} \end{cases} \tag{39}$$

i.e. the Hamiltonian on V can be replaced by the Hamiltonian on Λ through a suitable change of normalisation $Z_V^\tau \to \tilde{Z}_V^\tau$. This is due to the conditioning with τ which as a consequence ensures the existence of a $c > 0$ only dependent on $\tau_{\overline{V}}$, s.t. for all σ_V, $H_V(\sigma_V \tau_{\overline{V}}) + c = H_\Lambda(\sigma_V \tau_{\overline{V}})$. The marginal of a Gibbs distribution on $\Delta \subset V$ is defined as $\mu_{V,\Delta}^\tau : \Omega \to \mathbb{R}$ with

$$\mu_{V,\Delta}^\tau(\omega) = \sum_{\nu : \nu|_\Delta = \omega|_\Delta} \mu_V^\tau(\nu)\,. \tag{40}$$

By the definition of [Ces01], a system is *completely analytic* if there exists $K > 0$, $\xi > 0$ such that for all $V \Subset \mathbb{Z}^D$, $x \in \partial V := \{x \in \overline{V}\ :\ \operatorname{dist}(x, V) \leq r\}$, $\Delta \subset V$ and $\tau, \tau' \in \Omega$ with $\tau(y) = \tau'(y)$ for all $y \neq x$

$$\left\| \frac{\mu_{V,\Delta}^\tau}{\mu_{V,\Delta}^{\tau'}} - 1 \right\|_\infty \leq K e^{-\frac{\operatorname{dist}(x,\Delta)}{\xi}}\,. \tag{41}$$

Now let us try to make the connection to Definition A.1. Let us first fix the equivalent of σ^Λ of $\Lambda \Subset \mathbb{Z}^D$. Since quantumly we are considering only Hamiltonians on finite lattice Λ let us fix a boundary condition $\nu \in \Omega$ and set classical analogue of our Gibbs measure σ_Λ on Λ to be $\mu^\nu = \mu^\nu_\Lambda : \Omega \to \mathbb{R}$. Marginals of this distribution are defined through (40) and for $\Sigma \subset \Lambda$ denoted by μ^ν_Σ. We further adapt the shorthand notation $\mu^\nu_{\Sigma|\Lambda} := \frac{\mu^\nu_{\Sigma\Lambda}}{\mu^\nu_\Lambda}$. Incorporating the implicit boundary condition explicitly, Definition A.1 states: There exist $\xi > 0$ and $K > 0$ such that for all partitions $\Lambda = A \sqcup B \sqcup C \sqcup D$

$$\left\|\log \frac{\mu^\nu_{A|CD}}{\mu^\nu_{A|D}}\right\|_\infty \leq K|\Lambda| e^{-\frac{\operatorname{dist}(A,C)}{\xi}} . \tag{42}$$

First note that it suffices to investigate the case when C only contains a single site, as we can write

$$\left\|\log \frac{\mu^\nu_{A|CD}}{\mu^\nu_{A|D}}\right\|_\infty \leq \sum_{i=1}^{|C|} \left\|\log \frac{\mu^\nu_{A|\{x_i\}D_i}}{\mu^\nu_{A|D_i}}\right\|_\infty \tag{43}$$

where $D_1 = D$, $D_i = D_{i-1} \sqcup C_{i-1}$ and $C_i = \{x_i\}$ with $\{x_i\}_{i=1}^{|C|}$ an enumeration of C. In the following, we will drop ν for better readability. By the argument above we are now in the setting of $|C| = 1$. Let us assume $\tau \in \Omega$ to be the optimiser such that $\left\|\log \frac{\mu_{A|CD}}{\mu_{A|D}}\right\|_\infty = \log \frac{\mu_{A|D}(\tau)}{\mu_{A|CD}(\tau)}$. In the case that the optimum is achieved for the reversed fraction, the convexity of $x \mapsto \frac{1}{x}$ gives an inequality instead of equality in the second line of (44) while also the roles of τ and τ' change, however, leaving the rest of the argument invariant. Let us denote by τ' the element of Ω which agrees with τ on $\overline{C}$ but has a flipped spin at site C. By $\log(x+1) \leq x$, the fact that $\mu_{C|D}(\tau) = 1 - \mu_{C|D}(\tau')$ with both $\mu_{C|D}(\tau)$ and $\mu_{C|D}(\tau')$ in $[0,1]$, we get

$$\begin{aligned}
\log \frac{\mu_{A|D}(\tau)}{\mu_{A|CD}(\tau)} &\leq \frac{\mu_{A|D}(\tau)}{\mu_{A|CD}(\tau)} - 1 = \frac{\mu_{A|CD}(\tau)\mu_{C|D}(\tau) + \mu_{A|CD}(\tau')\mu_{C|D}(\tau')}{\mu_{A|CD}(\tau)} - 1 \\
&= \mu_{C|D}(\tau')\left(\frac{\mu_{A|CD}(\tau')}{\mu_{A|CD}(\tau)} - 1\right) \leq \left|\frac{\mu_{A|CD}(\tau')}{\mu_{A|CD}(\tau)} - 1\right| = \left|\frac{\mu^{\tau'}_{AB,A}(\tau)}{\mu^{\tau}_{AB,A}(\tau)} - 1\right| \\
&\leq \left\|\frac{\mu^{\tau'}_{AB,A}}{\mu^{\tau}_{AB,A}} - 1\right\|_\infty .
\end{aligned} \tag{44}$$

In the last equality, we used the fact that

$$\mu^\tau_{AB,A}(\omega) = \mu_{A|CD}(\omega_A \tau_{\overline{A}}) , \tag{45}$$

which one readily obtains from (39). As the system satisfies complete analyticity, this last norm is uniformly bounded by $Ke^{-\frac{\operatorname{dist}(A,C)}{\xi}}$ which for the general setting of (43) allows us to conclude

$$\left\|\log \frac{\mu_{A|CD}}{\mu_{A|D}}\right\|_\infty \leq K|C| e^{-\frac{\operatorname{dist}(A,C)}{\xi}} .$$

A.4.3 The effective Hamiltonians as tool to proof correlation decay

Let us conclude the section about the theoretical framework by introducing a tool which has proved useful when showing the decay of some correlation measures for explicit models. This tool is called effective Hamiltonians and has appeared in [KKB20] and was further developed in [BCPH24]. We will restrict our discussion to the case where $\Lambda \subseteq \mathbb{Z}^D$ is a finite hypercube and address the Hamiltonian directly, whereas the results in [BCPH24] apply to interactions on general graphs. Next, we introduce an essential norm originally developed for interactions; here, we define it specifically for Hamiltonians with a fixed local representation (cf. (A.3)): For H_Λ with representation as in (A.3) and $\mu > 0$ the interaction norm is defined as

$$|||H_\Lambda|||_\mu := \sup_{x\in\Lambda} \sum_{X\subseteq\Lambda: x\in X} \|h_X\|_\infty e^{\mu\,\mathrm{diam}(X)}\,. \tag{46}$$

Slightly adapting the definition from [BCPH24, Definition 3.1] we say that H_Λ admits a *strong local effective Hamiltonian at inverse temperature* $\beta > 0$ if for every $A \subseteq \Lambda$ there exists a Hamiltonian with a local decomposition $\tilde{H}^A = \sum_{X\subseteq\Lambda} \tilde{h}^A_X$ such that

1. $\tilde{h}^A_X$ is supported in $X \cap A$ for every $X \subseteq \Lambda$;
2. For $A' \subseteq \Lambda$, $\tilde{h}^A_X = \tilde{h}^{A'}_X$ for all $X \subseteq \Lambda$ for which $X \cap A' = X \cap A$;
3. One has $\log d_A^{-1}\, I_A \otimes \mathrm{tr}_A[e^{-\beta H_\Lambda}] = \sum_{X\subseteq\Lambda} \tilde{h}^A_X = \tilde{H}^A$.

The existence of an effective Hamiltonian alone does not suffice to ensure the decay of the MCMI, as the above definition lacks information on the system's locality. In general, this decay will require an additional locality assumption (see Appendix D.1). However, when the system Hamiltonian is also marginal commuting, [BCPH24, Theorem 4] directly guarantees the existence of an effective Hamiltonian with explicit decay properties. This result allows us to establish uniform decay of the MCMI at high temperatures in Appendix D.2. To clarify, we define what it means for H_Λ to have commuting marginals: We say that $H_\Lambda = \sum_{A\subseteq\Lambda} h_A$ is *marginal commuting* if there exists a commuting algebra $\mathcal{A}$ generated by all the local terms h_A closed under the application of $\mathbb{E}_A$ for any subset $A \subseteq \Lambda$, i.e., $\mathbb{E}_A[\mathcal{A}] \subset \mathcal{A}$. This definition is exactly [BCPH24, Definition 3.5] and is named here to highlight that, under this assumption, arbitrary marginals of the Gibbs state commute.

B Proof techniques - extended

B.1 A general weak entropy factorization

This section is dedicated to the proof and a short discussion of a weak entropy factorization. More precisely a splitting of the conditional relative entropy defined for $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\overline{A}}$ with $A \subseteq \Lambda$ as

$$D_A(\rho\|\sigma) := D(\rho\|\sigma) - D(\rho_{\overline{A}}\|\sigma_{\overline{A}}) \tag{47}$$

for $\rho, \sigma \in \mathcal{S}(\mathcal{H}_\Lambda)$. We can understand this as the relative entropy in the A subsystem conditioned on the rest of the lattice. It made its first appearance in [CLPG18a] where it also was connected to the Davies channel on A, more precisely, the following important inequality was shown:

$$D_A(\rho\|\sigma) \le D(\rho\|E_A(\rho))\,. \tag{48}$$

Naturally $D_\Lambda(\rho\|\sigma) = D(\rho\|E_\Lambda(\sigma)) = D(\rho\|\sigma)$.

The following lemma details the weak entropy factorization which is one of the key ingredients in our main results.

Lemma B.1 (Weak entropy factorization) *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$ and $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ with $\sigma > 0$, then*

$$D_{ABC}(\rho\|\sigma) \le D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma) + \|\mathbf{H}_\sigma(A : C|D)\|_\infty\,. \tag{49}$$

Proof. We have that

$$\begin{aligned}
&D_{ABC}(\rho\|\sigma) - D_{AB}(\rho\|\sigma) - D_{BC}(\rho\|\sigma)\\
&\qquad = -D(\rho_D\|\sigma_D) - D(\rho_{ABCD}\|\sigma_{ABCD}) + D(\rho_{CD}\|\sigma_{CD}) + D(\rho_{AD}\|\sigma_{AD})\\
&\qquad \leq -D(\rho_D\|\sigma_D) - D(\rho_{ACD}\|\sigma_{ACD}) + D(\rho_{CD}\|\sigma_{CD}) + D(\rho_{AD}\|\sigma_{AD}) && \text{(DPI)}\\
&\qquad \leq \operatorname{tr}[\rho_{ACD}(\log\sigma_{ACD} + \log\sigma_D - \log\sigma_{AD} - \log\sigma_{CD})] && \text{(SSA)}\\
&\qquad \leq \|\mathbf{H}_\sigma(A:C|D)\|_\infty\,. && \text{(Hölder)}
\end{aligned}$$

□

The correction's additive nature is why we only manage to obtain the weak approximate tensorization for the Davies (Appendix B.3) and as a consequence a weak modified logarithmic Sobolev inequality and a weak transport cost inequality (Appendix B.4). Lifting the above inequality to one which has a multiplicative correction would strengthen all results in this paper to their strong counterparts and further allow one to eliminate the requirement for a polynomial gap in Appendix C.2 and in the strengthening of the W^1 mixing in Appendix C.2. At the date of writing, we only managed to prove the inequality in the fully classical case where we obtain

$$(1 - \|\exp(\mathbf{H}_\sigma(A:C|D)) - I\|_\infty) D_{ABC}(\rho\|\sigma) \leq D_{AB}(\rho\|\sigma) + D_{BC}(\rho\|\sigma)\,.$$

Note that in the special case where D is empty, i.e. $\dim \mathcal{H}_D = 1$, the authors in [CLPG18b] show a multiplicative correction. Their proof yields a slightly altered factor, classically equivalent to the above bound.

B.2 A coarse-graining of the hypercubic lattice

In this section, we present the construction of the coarse-graining w.r.t. which we apply the weak entropy factorization (c.f. Appendix B.1) to prove the weak approximate tensorization later (c.f. Appendix B.3). This coarse-graining will be at the heart of its proof and our divide-and-conquer scheme involving both exact and weak approximate tensorizations at every step.

We do this by recalling the construction of a coarse-graining of the 2-dimensional hypercube due to [BK18] and extending it to any dimension $D \in \mathbb{N}$. We denote by $\Lambda_L := [\![-L, L]\!]^D \subset \mathbb{Z}^D$ the D-dimensional hypercube of side length $2L+1$. Our decomposition relies on the fixation of three parameters:

1. $c \in \mathbb{N}$ with $c \geq r$ the *overlap length*, ensuring a sufficiently fast decay when using a weak approximate tensorization to consecutively cut out cells reducing the extensive dimension in every step.
2. $k \in \mathbb{N}$ with $k \geq r$ the *buffer length*, separating cells of the same dimensionality such that the corresponding conditional expectations commute and hence allow us to use tensorization of the relative entropy.
3. $\ell \in \mathbb{N}$ with $L \geq \ell$ and ℓ odd the 'extensive' side length of the cells we decompose into. As we will consecutively reduce the 'extensive' dimensions until we reach zero subtracting ℓ in every step by $2(k+c)$, we further require $\ell > 2D(k+c)$.

Remark 3. In the first step of the decomposition, we aim to divide Λ_L into hypercubes of side length ℓ. This might not be possible due to the choice of ℓ we made prior as divisibility of $2L+1$ by ℓ may not be guaranteed. We can, however, always embed Λ_L into some Λ_{L+r} with $r \in [\![1, l-1]\!]$

such that $2(L+r)+1$ is now divisible by ℓ, when ℓ is odd. Since we are interested in how certain quantities scale in ℓ we may always choose ℓ odd, however, we will overlook this subtlety.

The original Hamiltonian is likewise embedded by padding it with zeros, leading to a Gibbs state of the form $\sigma^{\Lambda_L} \otimes \pi_{\text{rest}} \in \mathcal{S}(\mathcal{H}_{\Lambda_{L+r}})$ where π_{rest} is the maximally mixed state on $\Lambda_{L+r} \backslash \Lambda_L$. The (κ, r)-locality and interaction strength of both Hamiltonians, original and padded agree, due to the specific embedding. Furthermore, the condition of the global and local gaps and the decay of MCMI hold for the embedded if and only if they hold for the original Hamiltonian. This means we employ our proof for the embedded version and immediately get a result for the original one.

Consider the example of the MLSI, where we denote by $\mathcal{L}$ the Davies Lindbladian in Λ_L and by $\mathcal{L}_{\text{rest}}$ that of $\Lambda_{L+r} \setminus \Lambda_L$. The separation between them is due to the embedding which in addition gives $[\mathcal{L}, \mathcal{L}_{\text{rest}}] = 0$. Assuming the embedded system has MLSI, we have $D(e^{t(\mathcal{L}+\mathcal{L}_{\text{rest}})}(\rho') \| \sigma \otimes \pi_{\text{rest}}) \leq e^{-\frac{t}{c_1}} D(\rho' \| \sigma \otimes \pi_{\text{rest}})$. By setting $\rho' = \rho \otimes \pi_{\text{rest}}$ and utilizing commutativity of the Lindbladians, we obtain $D(e^{t\mathcal{L}}(\rho) \otimes \pi_{\text{rest}} \| \sigma \otimes \pi_{\text{rest}}) \leq e^{-\frac{t}{c_1}} D(\rho \otimes \pi_{\text{rest}} \| \sigma \otimes \pi_{\text{rest}})$. Through the tensorization of relative entropy, we can conclude that c_1 also serves as an MLSI constant for the original system. This approach extends to other contexts (e.g. the decay of Wasserstein distance). We've established that all conditions imposed on the original system remain unchanged for the embedding, meaning the only change in the obtained bounds is in the system size. Notably, $|\Lambda_{L+r}| \leq 2^D |\Lambda_L|$ since $\ell \leq L$, meaning the correction only depends on the lattice dimension and does not affect the scaling with system size. Throughout the paper, Λ_L is treated as general. However, in the proofs, we always assume without loss of generality that ℓ divides $(2L+1)$. When providing explicit constants for the bounds, we include the correction factor 2^D.

In the process of constructing the coarse-graining, we will define various cells of dimension a, where a ranges from 0 to $D-1$. These cells should be understood as embedded within the corresponding D-dimensional hypercube for which $e_1, \ldots, e_D$ denote the canonical orthonormal basis. We begin by partitioning the D-dimensional hypercube $C^D := \Lambda_L$ into smaller D-hypercubes, each with side length ℓ. We denote these partition elements as $\{C^\partial_{D,i}\}_{i \in \mathcal{I}_D}$ with $\mathcal{I}_D = [\![1, ((2L+1)/\ell)^D]\!]$. For each cell $C^\partial_{D,i}$, we define two subsets:

1. The interior of $C^\partial_{D,i}$, excluding a boundary layer of buffer length k:

$$C_{D,i} := \{v \in C^\partial_{D,i} \ : \ \text{dist}(v, \overline{C^\partial_{D,i}}) > k\} \, ;$$

2. A further restricted interior, excluding a boundary layer of width $k+c$, i.e excluding buffer and overlap length:

$$\mathring{C}_{D,i} := \{v \in C^\partial_{D,i} \ : \ \text{dist}(v, \overline{C^\partial_{D,i}}) > k + c\} \, .$$

In the above $R \subseteq \Lambda_L$ is given by $\overline{R} = \Lambda_L \backslash R$. We define the aggregate sets:

$$C_D := \bigsqcup_{i \in \mathcal{I}_D} C_{D,i} \qquad \mathring{C}_D := \bigsqcup_{i \in \mathcal{I}_D} \mathring{C}_{D,i},$$

where $\bigsqcup$ denotes the disjoint union. Finally, we set $C^{D-1} := \Lambda_L \setminus \mathring{C}_D$. For a visual representation of this construction in two dimensions ($D = 2$), refer to Figure 5.

We continue the deconstruction by defining cells of lower extensive dimensions. For each dimension $a \in [\![1, D-1]\!]$, we iteratively construct a-cells from C^a as follows:

1. First, we identify the set of vertices that "join" each pair of neighbouring $\mathring{C}_{a+1,i}$:

$$C^\partial_a := \left\{ v \in C^a \ : \ \exists j \in \mathbb{Z}, b \in [\![1, D]\!] : v + j e_b \in \overline{C^a} \right\} \, .$$

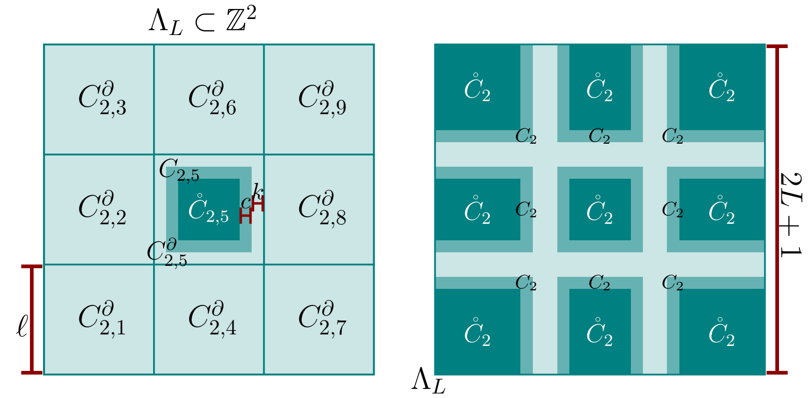


Figure 5: Splitting of $\Lambda_L = C^D$ for $D = 2$ as described in the text. On the left side, Λ_L has been split into 9 2-cells. For the cell in the middle $C_{2,5}$ and $\mathring{C}_{2,5}$ are explicitly represented. On the right side, C_2 is the union of the midblue and darkblue regions, and $\mathring{C}_2$ is the union of the dark blue regions.

2. By construction, C^∂_a is a disjoint union of fattened a-cells, which we denote as $C^\partial_{a,i}$ with index set $\mathcal{I}_a$.

3. We again define the two subsets:

$$C_{a,i} := \{v \in C^\partial_{a,i} \ : \ \mathrm{dist}(v, \overline{C^\partial_{a,i}} \cap C^a) > k\}, \qquad \mathring{C}_{a,i} := \{v \in C^\partial_{a,i} \ : \ \mathrm{dist}(v, \overline{C^\partial_{a,i}} \cap C^a) > k + c\}$$

separated from the skeleton that remains after removal of $\bigsqcup_{b=a+1}^{D} \mathring{C}_b$ by buffer and overlap plus buffer length respectively. We set the aggregated sets to be

$$C_a := \bigsqcup_{i \in \mathcal{I}_a} C_{a,i}, \qquad \mathring{C}_a := \bigsqcup_{i \in \mathcal{I}_a} \mathring{C}_{a,i}.$$

4. Finally, we define the set for the next lower dimension: $C^{a-1} := C^a \setminus \mathring{C}_a$.

This process is iterated for decreasing values of a, creating a hierarchical structure of cells of different extensive dimensions. For a visual representation of this construction for $a = 1$ in two dimensions ($D = 2$), refer to Figure 6.

Lastly for $a = 0$ we set $C^0 := C^1 \backslash \mathring{C}_1$ and $\mathring{C}_{0,i} := C_{0,i} := C^\partial_{0,i}$ so that $C^0 = \bigsqcup_{i \in \mathcal{I}_0} C_{0,i}$. This is shown on the left side of Figure 7. Let us summarise the properties of the above construction in the following lemma.

Lemma B.2 *The decomposition of Λ_L described in the previous paragraphs satisfies the following properties:*

1. *For each $a \in [\![0, D]\!]$, $\mathring{C}_a := \bigsqcup_{i \in \mathcal{I}_a} \mathring{C}_{a,i}, C_a := \bigsqcup_{i \in \mathcal{I}_a} C_{a,i}$, and $C^\partial_a := \bigsqcup_{i \in \mathcal{I}_a} C^\partial_{a,i}$ are unions of disjoint sets, with size bounded as $|\mathring{C}_{a,i}| \leq |C_{a,i}| \leq |C^\partial_{a,i}| \leq \ell^D$.*[1]

[1] This is a non-tight bound and it holds that $|C_{a,i}| \leq [2(D-a)(k+c)]^{D-a}[\ell - 2(D-a)(k+c)]^a \leq \ell^D$.

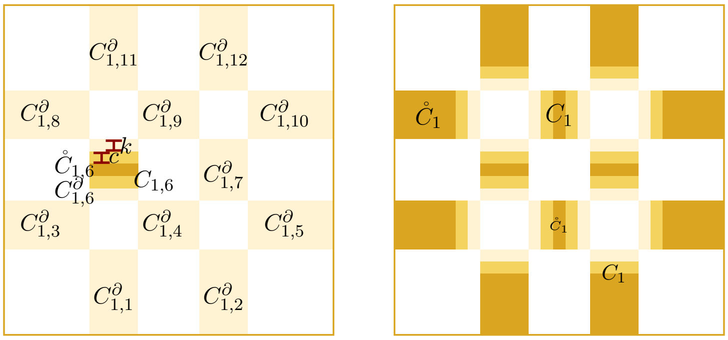


Figure 6: The figure shows the construction of C_1^∂, contained in $C^1 = \Lambda_L \backslash \mathring{C}_1$. On the left side, C^1 is the whole region, and C_1^∂ has been split into several fattened 1-cells respectively; we represent explicitly, $C_{1,6}$, $\mathring{C}_{1,6}$. On the right side, $\mathring{C}_1$ is the union of the dark green regions, $C_1 \backslash \mathring{C}_1$ is that of the medium green ones, and $C_1^\partial \backslash C_1$ is the union of the lighter green, dashed regions without the lighgreen squares in the centre, however.

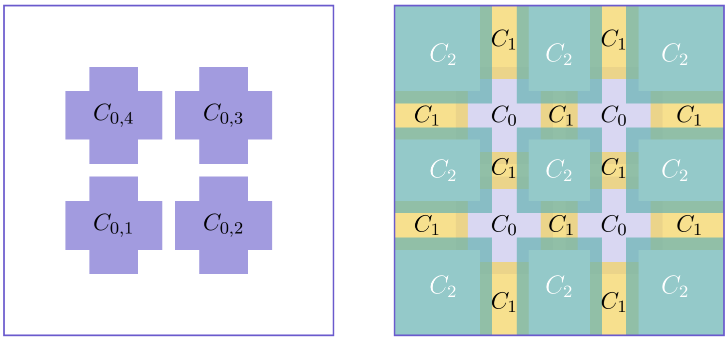


Figure 7: On the left side we show C_0, which is defined as $C^1 \backslash \mathring{C}_1$. On the right side, we show the coarse-graining in terms of the combined C_0, C_1 and C_2. We omit the corresponding C_x^∂ and $\mathring{C}_x$, for $x = 0, 1, 2$, for simplicity.

2. $C^0 = C_0 = \mathring{C}_0 = \bigsqcup_{i \in \mathcal{I}_0} C^\partial_{0,i}$ *is a disjoint union of "fat"* 0*-cells, in the shape of* D*-dimensional "crosses", with each* $C^\partial_{0,i}$ *included in a hypercube of sidelength* $2D(k+c)$ *and distance*

$$\operatorname{dist}(C^\partial_{0,i}, C^\partial_{0,j}) \geq \ell - 2D(k+c) > 0$$

from each other.

3. *The hierarchy* $\{C_a\}_{a=0}^D$ *induces a suitably overlapping coarse-graining of* Λ_L*, i.e.* $\bigcup_{a=0}^D C_a = C^D = \Lambda_L$*, and each site* $x \in \Lambda_L$ *is included in at most* $D+1$ *sets* $\{C_{a,i}\}_{a,i}$.

4. $c \leq \operatorname{dist}(C^a \backslash C_a, \mathring{C}_a) = \operatorname{dist}(C^a \backslash C_a, C^a \backslash C^{a-1})\ \forall a \in [\![1, D]\!]$,

5. $2k \leq \operatorname{dist}(C_{a,i}, C_{a,j}) \quad \forall i, j \in \mathcal{I}_a, \forall a \in [\![0, D]\!]$.

For a visualization, find an example of the case $D = 2$ on the right side of Figure 7.

Proof. 1. We proceed by induction on a. For $a = D$, each of the sets $C^{\partial}_{D,i}$ are hypercubes with sidelength ℓ, hence $C_{D,i}, \mathring{C}_{D,i}$ are hypercubes of side length at most $\ell - k$ and $\ell - (k + c)$, respectively. Disjointness is clear. For $a < D$, it holds that

$$\begin{aligned} C^{\partial}_a &:= \Big\{ v \in C^a \ : \ \exists j \in \mathbb{Z}, b \in [\![1, D]\!] : \ v + je_b \in \overline{C^a} \Big\} \\ &= \bigcup_{b=1}^{D} \Big(C^a \cap \Big\{ \mathring{C}_{a+1} + \mathbb{Z} e_b \Big\} \Big) \\ &= \bigcup_{a'=1}^{D} \Big(C^a \cap \Big\{ \mathring{C}_{a+1} + \mathbb{Z}_{2(D-a)(k+c)} e_b \Big\} \Big) . \end{aligned}$$

More precisely: For each $\mathring{C}_{a+1,i}$ and each canonical direction b, we construct a set $C^{\partial}_{a,j(i,b)}$ by translating $\mathring{C}_{a+1,i}$ by $2(D - a)(k + c)$ along e_b. We then define $C^{\partial}_{a,j(i,b)}$ as this translated set minus its intersection with $\mathring{C}_{a+1,i}$. We eliminate duplicates from the collection $\{C^{\partial}_{a,j(i,b)}\}_{(i,b) \in \mathcal{I}_{a+1} \times [\![1,D]\!]}$ to obtain the index set $\mathcal{I}_a$, which we use to obtain C^{∂}_a. The sets $C^{\partial}_{a,i}$ are mutually disjoint by construction. Each $C^{\partial}_{a,i}$ has $D - a$ sides of length $2(D - a)(k + c)$, while the remaining sides retain the length $\ell - 2(D - a)(k + c)$ inherited from the sets composing $\mathring{C}_{a+1}$. The disjointness of $C_{a,i}$ and $\mathring{C}_{a,i}$ as well as the bounds on their sizes immediately follow.

2. The disjointedness statement in 2. follows equally from the proof in 1. The statement $C^0 = \mathring{C}_0$ follows by their definitions.

3. The proof proceeds by finite induction with at most D steps, starting from $a = D$. We consider the following cases: Base case: If $x \in C_D$, we are done. Inductive step: If $x \notin C_D$, then $x \in C^{D-1} \supseteq \Lambda_L \setminus C_D$. For each subsequent step a, we have three possibilities:

   (a) If $a = 0$, then $x \notin \bigcup_{b=1}^{D} C_b$. Consequently, $x \in C^0 = C^{\partial}_0 = C_0 = \mathring{C}_0$, and we are done.

   (b) If $x \in C_a$, we are done.

   (c) If neither (a) nor (b) holds, then $x \notin \bigcup_{b=a}^{D} C_b$. Hence, $x \in \Lambda_L \setminus \bigcup_{b=a}^{D} C_b \subset C^{a-1}$, and the induction continues to the next step.

   The induction terminates after at most D steps because there are only $D + 1$ sets C_a. And since each is a disjoint union of $C_{a,i}$ every $x \in \Lambda_L$ is contained in at most $D + 1$ of the $C_{a,i}$.

4. Recall that $k, c \geq r$, the range of the interactions. The claim follows via

$$\begin{aligned} \operatorname{dist}(C^a \backslash C_a, \mathring{C}_a) &= \operatorname{dist}(C^a \backslash C_{a,i}, \mathring{C}_{a,i}) = \operatorname{dist}(\overline{C_{a,i}} \cap C^a, \mathring{C}_{a,i} \cap C^a) \\ &= \operatorname{dist}(\{x \in C^a : \operatorname{dist}(x, \overline{C^{\partial}_{a,i}}) \leq k\}, \{x \in C^a : \operatorname{dist}(x, \overline{C^{\partial}_{a,i}}) > k + c\}) = c \end{aligned}$$

   and it is easy to check that $C^a \backslash C^{a-1} = C^a \backslash (C^a \backslash \mathring{C}_a) = \mathring{C}_a$ for $a \in [\![1, D]\!]$.

5. This follows directly from the disjointness of $C_{a,i}$ and $C_{a,j}$, see 1. and their definition.

□

B.3 A weak approximate tensorization for Davies channels

By combining the general weak entropy factorization from Appendix B.1 with the coarse-graining in Appendix B.2 this section contains a weak approximate tensorization (wAT) for the Davies channels with corrections that are the MCMIs of the splitting. Under the assumption that the Gibbs state exhibits exponential decay of its matrix-valued quantum conditional mutual information this correction term decays exponentially with the overlap length l (see Lemma B.2).

Theorem B.3 (Weak approximate tensorization) *Given the decomposition of the lattices Λ with constants $(k \geq r, c, l)$ described in the previous section (c.f. Appendix B.2) and then in the context of Appendix A.3, the family $\{E_{C_{a,i}} : a \in [\![0, D]\!], i \in \mathcal{I}_a\}$ of Davies conditional expectations with respect to σ satisfies the following inequality: for all $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$,*

$$D(\rho\|\sigma) \leq \sum_{a=0}^{D} \sum_{i_a \in \mathcal{I}_a} D(\rho\|E_{C_{a,i_a}}(\rho)) + \sum_{a=1}^{D} \zeta_a(\sigma)\,, \tag{50}$$

with

$$\zeta_a(\sigma) := \|\mathbf{H}_\sigma(X_a : Z_a|W_a)\|_\infty\,, \tag{51}$$

and $W_a \sqcup X_a \sqcup Y_a \sqcup Z_a =: \overline{C^a} \sqcup \mathring{C}_a \sqcup (C_a \backslash \mathring{C}_a) \sqcup (C^a \backslash C_a)$ with $d(X_a, Z_a) = c \geq r$, for $a \in [\![1, D]\!]$. Assuming further that the Gibbs state satisfies exponential decay of MCMI with constants K and ξ, then we can estimate (50) with

$$D(\rho\|\sigma) \leq \sum_{a=0}^{D} \sum_{i_a \in \mathcal{I}_a} D(\rho\|E_{C_{a,i_a}}(\rho)) + DK|\Lambda_L|e^{-c/\xi}\,. \tag{52}$$

Proof. We proceed by induction on a, starting at $a = D$ and decreasing to zero. Using the notation and result from Lemma B.1, and conditioning on $W_D = \emptyset = \overline{C^D} = \overline{\Lambda}$, we obtain:

$$\begin{aligned}
D(\rho\|\sigma) = D_{C^D}(\rho\|\sigma) = D_{X_D Y_D Z_D}(\rho\|\sigma) &\leq D_{X_D Y_D}(\rho\|\sigma) + D_{Y_D Z_D}(\rho\|\sigma) + \|\mathbf{H}_\sigma(X_D : Z_D|\emptyset)\|_\infty \\
&= D_{C_D}(\rho\|\sigma) + D_{C^{D-1}}(\rho\|\sigma) + \zeta_D(\sigma) \\
&\overset{(1)}{\leq} D(\rho\|E_{C_D}(\sigma)) + D_{C^{D-1}}(\rho\|\sigma) + \zeta_D(\sigma) \\
&\overset{(2)}{\leq} \sum_{i \in \mathcal{I}_D} D(\rho\|E_{C_{D,i}}(\sigma)) + D_{C^{D-1}}(\rho\|\sigma) + \zeta_D(\sigma).
\end{aligned}$$

In step (1), we apply (48). Then, we use the fact that $C_D = \bigsqcup_{i \in \mathcal{I}_D} C_{D,i}$ with $\mathrm{dist}(C_{i,D}, C_{j,D}) > k \geq r$, as per Lemma B.2. This allows us to apply 3. of the properties of the Davies channels from Appendix A.3, i.e split $E_D = \prod_{i \in \mathcal{I}_D} E_{C_{D,i}}$ into commuting constituents. Consequently, in step (2), we split the relative entropy according to (24). The induction proceeds with $D_{C^{D-1}}(\rho\|\sigma)$.

The induction step follows a similar line of reasoning: For $a \neq 0$, we again employ the weak entropy factorisation from Lemma B.1:

$$\begin{aligned}
D_{C^a}(\rho\|\sigma) = D_{X_a Y_a Z_a}(\rho\|\sigma) &\leq D_{X_a Y_a}(\rho\|\sigma) + D_{Y_a Z_a}(\rho\|\sigma) + \|H_\sigma(X_a : Z_a|W_a)\|_\infty \\
&= D_{C_a}(\rho\|\sigma) + D_{C^{a-1}}(\rho\|\sigma) + \|\mathbf{H}_\sigma(X_a : Z_a|W_a)\|_\infty \\
&\leq \sum_{i \in \mathcal{I}_a} D(\rho\|E_{C_{a,i}}(\rho)) + D_{C^{a-1}}(\rho\|\sigma) + \zeta_a(\sigma)\,.
\end{aligned}$$

For $a = 0$, we only need to estimate $D_{C^0}(\rho\|\sigma) = D_{C_0}(\rho\|\sigma) \leq \sum_{i \in \mathcal{I}_0} D(\rho\|E_{C_{0,i}}(\rho))$. This inequality holds because we first can employ (48) and then use that $C_0 = \bigsqcup_{i \in \mathcal{I}_0} C_{0,i}$, where the $C_{0,i}$ have

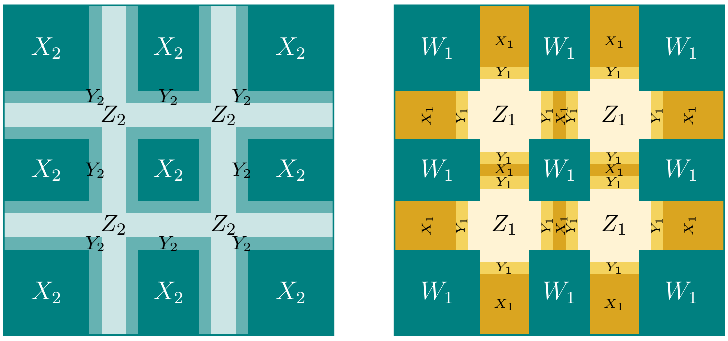


Figure 8: The decomposition of the lattice used in the weak approximate tensorization for the case $D = 2$ in the steps $a = 2$ on the left and $a = 1$ on the right.

mutual distances greater than r. Thus, 3. for the Davies channels from Appendix A.3 applies, allowing us to decompose E_{C_0} into a mutually commuting composition of the $E_{C_{0,i}}$ and then apply (24). The final inequality holds because again $C_a = \bigsqcup_{i\in\mathcal{I}_a} C_{a,i}$ with $\operatorname{dist}(C_{a,i}, C_{a,j}) > r$ for $i \neq j$ by Lemma B.2, allowing us to apply property 3. for Davies channels from Appendix A.3 and, subsequently, (24). We thereby complete the induction. In Figure 8 the decomposition of the lattice for the weak approximate tensorization is demonstrated for the case $D = 2$ and $a = D$ and $a = D - 1$ respectively.

In the case of exponential decay of matrix-valued quantum conditional mutual information, we have:

$$\zeta_a(\sigma) \leq K|\Lambda|e^{-c/\xi}\,,$$

independent of a, which immediately gives (52). □

B.4 A weak transport cost inequality for Davies channels

Tailored to our weak approximate tensorization from the previous section Appendix B.3 but valid also for arbitrary coarse-grainings of the lattice with Davies channels, we derive in this section a weak transport cost (wTC) inequality which forms a crucial ingredient in the proof of the Wasserstein mixing result. Let us begin with an auxiliary lemma.

Lemma B.4 *Let $\mathcal{L}$ a generator of a GNS-symmetric quantum Markov semigroup with invariant state $\sigma > 0$ that satisfies (wMLSI) with constants c_1 and c_2 and let $\rho \in \mathcal{S}(\mathcal{H})$. Assume that $D(\rho\|\sigma) \geq c_2$. Then for $t \in (0, t_{c_2}(\rho))$ where $t_{c_2}(\rho) := \sup\{t \geq 0 : D(e^{t\,\mathcal{L}}(\rho)\|\sigma) \geq c_2\}$*

$$\int_0^t \sqrt{\mathrm{EP}_{\mathcal{L}}(\rho_s)}ds \leq 2\sqrt{c_1 D(\rho\|\sigma)} \quad \textit{with} \quad \rho_s := e^{s\,\mathcal{L}}(\rho)\,. \tag{53}$$

Proof. Let us define $F : (0, t_{c_2(\rho)}) \to \mathbb{R}$ as

$$F(t) := \int_0^t \sqrt{\mathrm{EP}_{\mathcal{L}}(\rho_s)}ds + 2\sqrt{c_1(D(\rho_t\|\sigma) - c_2)}\,. \tag{54}$$

This function is non-increasing in t, as

$$\frac{d}{dt}F(t) = \sqrt{\mathrm{EP}_{\mathcal{L}}(\rho_t)} - \frac{\sqrt{c_1}\,\mathrm{EP}_{\mathcal{L}}(\rho_t)}{\sqrt{D(\rho_t\|\sigma) - c_2}} \leq 0 \tag{55}$$

due to wMLSI. The claim then immediately follows. □

With this lemma in place, we can now move our attention to the proof of the wTC from the wAT.

Theorem B.5 (Weak transport cost inequality) *In the context of Appendix A.3, assume that there exists $A = \{A_i\}_{i=1}^{n_A}$ with $A_i \subseteq \Lambda$ an overlapping coarse-graining of Λ, i.e. $\bigcup_{i=1}^{n_A} A_i = \Lambda$ with corresponding Davies projectors E_{A_i} such that one has*

$$D(\rho\|\sigma) \leq \sum_{i=1}^{n_A} D(\rho\|E_{A_i}(\rho)) + c_2 \tag{56}$$

for all $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$, then the following holds

$$\|\rho - \sigma\|_{W^1} \leq \max_i 2\sqrt{2}|A_i\partial|\sqrt{n_A D(\rho\|\sigma)} + |\Lambda|\sqrt{2c_2} \tag{57}$$

for all $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$.

Proof. In the following let us denote $\mathcal{L}_{A_i}^H := E_{A_i} - \mathrm{id}$. By non-negativity of the relative entropy, it holds that

$$\begin{aligned} D(\rho\|E_{A_i}(\rho)) &\leq D(\rho\|E_{A_i}(\rho)) + D(E_{A_i}(\rho)\|\rho) = \mathrm{tr}[(\mathrm{id} - E_{A_i})(\rho)(\log(\rho) - \log E_{A_i}(\rho))] \\ &= \mathrm{tr}[(\mathrm{id} - E_{A_i})(\rho)(\log(\rho) - \log\sigma)] \\ &= \mathrm{EP}_{\mathcal{L}_{A_i}^H}(\rho)\,, \end{aligned} \tag{58}$$

where in the last equality we used that $\log E_{A_i}(\rho) - \log\sigma$ is a fixpoint of E_A^* which immediately gives $\mathrm{tr}[(\mathrm{id} - E_{A_i})(\rho)(\log E_{A_i}(\rho) - \log\sigma)] = 0$. Using (58), the assumption can be rewritten as

$$D(\rho\|\sigma) \leq \mathrm{EP}_{\mathcal{L}_\Lambda^H}(\rho) + c_2 \tag{59}$$

where we defined the primitive GNS-symmetric generator $\mathcal{L}_\Lambda^H := \sum_{i=1}^{n_A} \mathcal{L}_{A_i}^H$. As E_{A_i} only acts non-trivially on $A_i\partial$, we find that $\mathrm{tr}_{A_i\partial} \circ \mathcal{L}_{A_i}^H = 0$ which is a direct consequence of (36). From this fact one immediately obtains

$$\left\|\mathcal{L}_{A_i}^H(\rho)\right\|_{W^1} \leq |A_i\partial|\sqrt{2\,\mathrm{EP}_{\mathcal{L}_{A_i}^H}(\rho)}\,, \tag{60}$$

as by (22) one has

$$\left\|\mathcal{L}_{A_i}^H(\rho)\right\|_{W^1} \leq |A_i\partial|\left\|\mathcal{L}_{A_i}^H(\rho)\right\|_1.$$

Now using Pinskers' inequality we get $\left\|\mathcal{L}_{A_i}^H(\rho)\right\|_1 \leq \sqrt{2D(\rho\|E_{A_i}(\rho))}$ which combined with (58) gives (60). We can now shift our attention to the main result. First assume that $D(\rho\|\sigma) \leq c_2$, then the inequality holds trivially by $\|\cdot\|_{W^1} \leq |\Lambda|\|\cdot\|_1$ (60) and Pinsker's inequality (19). If this is not the case we set $t_{c_2} := \sup\{t \geq 0 : D(e^{t\mathcal{L}_\Lambda^H}(\rho)\|\sigma) \geq c_2\}$ and get

$$\|\rho - \sigma\|_{W^1} \leq \left\|\rho - \rho_{t_{c_2}}\right\|_{W^1} + \left\|\rho_{t_{c_2}} - \sigma\right\|_{W^1} \leq \left\|\rho - \rho_{t_{c_2}}\right\|_{W^1} + |\Lambda|\sqrt{2c_2}$$

where we defined $\rho_t := e^{t\mathcal{L}_\Lambda^H}(\rho)$ for $t \geq 0$. The second inequality follows again by the above reasoning. For the first term, we write

$$\left\|\rho - \rho_{t_{c_2}}\right\|_{W^1} \leq \int_0^{t_{c_2}} \left\|\mathcal{L}_\Lambda^H(\rho_s)\right\|_{W^1} ds \leq \int_0^{t_{c_2}} \sum_{i=1}^{n_A} \left\|\mathcal{L}_{A_i}^H(\rho_s)\right\|_{W^1} ds\,.$$

By (60) we get

$$\int_0^{t_{c_2}} \sum_{i=1}^{n_A} \|\mathcal{L}^H_{A_i}(\rho_s)\|_{W^1} ds \le \max_i \sqrt{2}|A_i\partial| \int_0^{t_{c_2}} \sum_{i=1}^{n_A} \sqrt{\mathrm{EP}_{\mathcal{L}^H_{A_i}}(\rho_s)} ds$$

and lastly through concavity of $x \mapsto \sqrt{x}$,

$$\|\rho - \rho_{t_{c_2}}\|_{W^1} \le \max_i \sqrt{2}|A_i\partial|\sqrt{n_A} \int_0^{t_{c_2}} \sqrt{\mathrm{EP}_{\mathcal{L}^H_\Lambda}(\rho_s)} ds\,.$$

Using (59) the claim follows by Lemma B.4. □

B.5 A MLSI alike inequality for local Davies semigroups at every temperature

This section is dedicated to collecting some results about Davies conditional expectations and entropy productions that will all be used in an inequality that almost resembles an MLSI for the local Davies Lindbladiens. More precisely the following inequality

$$D(\rho\|E_A(\rho)) \le \mathcal{O}(e^{\mathcal{O}(|A\partial|)})\, \mathrm{EP}_{\mathcal{L}_{A\partial}}(\rho)\,, \tag{61}$$

which will be shown and used in the next section, i.e. Appendix B.6. As our proof will rely on going through the infinite temperature Davies (i.e. $\lim_{\beta\to 0} \mathcal{L}^\beta_A = \mathcal{L}^0_A$) we will keep the temperature explicit in the following. The strategy will be to relate the LHS and the entropy production on the RHS of (61) to their infinite temperature counterparts through Lemma B.6 and then to proof (61) only for the infinite temperature Davies, where the semigroup is almost the depolarizing one.

Lemma B.6 *In the context of Appendix A.3, let $A \subseteq \Lambda$ and define $E^\beta_A = \lim_{t\to\infty} e^{t\,\mathcal{L}^{D,\beta}_A}$ and $E^0_A = \lim_{t\to\infty} e^{t\,\mathcal{L}^{D,0}_A}$. Then the following inequalities hold:*

$$D(\rho\|E^\beta_A(\rho)) \le e^{2gJ\beta|A\partial|} D(\rho\|E^0_A(\rho))\,. \tag{62}$$

and

$$\mathrm{EP}_{\mathcal{L}^{D,0}_A}(\rho) \le e^{2gJ(\beta|A\partial|+1)}\, \mathrm{EP}_{\mathcal{L}^{D,\beta}_A}(\rho) \tag{63}$$

Proof. Let $\sigma = E^\beta_A(d^{-1}_\Lambda I_\Lambda)$ and $\sigma' = E^0_A(d^{-1}_\Lambda I_\Lambda) = d^{-1}_\Lambda I_\Lambda$. Then $\mathcal{L}^{D,\beta}_A$ is GNS-symmetric with respect to σ, and $\mathcal{L}^{D,0}_A$ is GNS-symmetric with respect to σ'. Applying [JLR22, Proposition 4.2], we obtain:

$$D(\rho\|E^\beta_A(\rho)) \le \|E^\beta_A(I_\Lambda)\|_\infty D(\rho\|E^0_A(\rho))\,.$$

Observe that $E^\beta_A(e^{-\beta H_{A\partial}}) = e^{-\beta H_{A\partial}}$, hence $\|E^\beta_A(I_\Lambda)\|_\infty \le \|e^{\beta H_{A\partial}}\|_\infty \|e^{-\beta H_{A\partial}}\|_\infty \le e^{2\beta gJ|A\partial|}$. For the entropy productions, [JLR22, Proposition 4.3] yields:

$$\mathrm{EP}_{\mathcal{L}^{D,0}_A}(\rho) \le \max_\omega e^{|\omega|/2} \|(E^\beta_A(I_\Lambda))^{-1}\|_\infty\, \mathrm{EP}_{\mathcal{L}^{D,\beta}_A}(\rho) \tag{64}$$

The Bohr frequencies are localized and can be bounded above by $|\omega| \le \max_{k\in A} 2\|H_{k\partial}\|_\infty \le 2gJ$ as per (34). Furthermore, $\|(E^\beta_A(I_\Lambda))^{-1}\|_\infty \le \|e^{\beta H_{A\partial}}\|_\infty \|e^{-\beta H_{A\partial}}\|_\infty \le e^{2\beta gJ|A\partial|}$, completing the proof. □

To conclude let us present (61) for the infinite temperature Davies.

Lemma B.7 *In the context of Appendix A.3, let $A \subseteq \Lambda$ and $E_A^0 = \lim_{t\to\infty} e^{t\,\mathcal{L}_A^{D,0}}$. Then:*

$$\chi^0_{\min} D(\rho \| E_A^0(\rho)) \leq \mathrm{EP}_{\mathcal{L}_{A\partial}^{D,0}}(\rho)\,. \tag{65}$$

Proof. By the GNS-symmetry of $\mathcal{L}^0_{A\partial}$, we can employ [CM17, Theorem 5.10] to express $\mathrm{EP}_{\mathcal{L}_{A\partial}^{D,0}}(\rho)$ as a sum of non-negative inner products with positive coefficients, including $\chi^0_{\omega,k}$. Using the bound $\chi^0_{\omega,k} \geq \chi^0_{\min}$, we obtain:

$$\chi^0_{\min}\,\mathrm{EP}_{\tilde{\mathcal{L}}_{A\partial}^{D,0}} \leq \mathrm{EP}_{\mathcal{L}_{A\partial}^{D,0}}\,,$$

where

$$\begin{aligned}
\tilde{\mathcal{L}}_{A\partial}^{D,0} &= \sum_{k\in A\partial}\sum_{\omega,\alpha}\left(S^\omega_{\alpha,k}\rho S^{\omega,\dagger}_{\alpha,k} - \frac{1}{2}\left\{\rho, S^{\omega,\dagger}_{\alpha,k}S^\omega_{\alpha,k}\right\}\right)\\
&= \sum_{t\geq 0}\sum_{k\in A\partial}(\sum_\alpha \Delta_{e^{itH}}\left(S_{\alpha,k}\Delta_{e^{-itH}}(\rho)S_{\alpha,k}\right) - \rho)\,.
\end{aligned}$$

The last simplification follows from [KB16, Theorem 31]. Define $\rho_t := \Delta_{e^{-itH}}(\rho)$. Using the gauge freedom of entropy production with respect to the choice of fixed point, and noting that I_Λ / d_Λ is a fixed point, we have:

$$\begin{aligned}
\mathrm{EP}_{\tilde{\mathcal{L}}_{A\partial}^{D,0}}(\rho) &= -\operatorname{tr}\left[\tilde{\mathcal{L}}_{A\partial}^{D,0}(\rho)(\log(\rho) - \log I/d)\right] = \operatorname{tr}\left[\tilde{\mathcal{L}}_{A\partial}^{D,0}(\rho)\log(\rho)\right]\\
&= \sum_{t\geq 0}\sum_{k\in A\partial,}\operatorname{tr}\left[(\rho_t - \sum_\alpha S_{\alpha,k}\rho_t S_{\alpha,k})\log(\rho_t)\right] = \sum_{t\geq 0}\mathrm{EP}_{\mathcal{L}_{A\partial}^{\mathrm{depol}}}(\rho_t) \geq \min_t \mathrm{EP}_{\mathcal{L}_{A\partial}^{\mathrm{depol}}}(\rho_t)\,.
\end{aligned}$$

Where $\mathcal{L}^{\mathrm{depol}}_{A\partial}(\,\cdot\,) = \sum_{k\in A\partial}(\sum_\alpha S_{\alpha,k}(\,\cdot\,)S_{\alpha,k} - \mathrm{id}) = \sum_{k\in A\partial}(\mathrm{tr}_k[\,\cdot\,] - \mathrm{id})$. Since the depolarizing channel has a cMLSI of 1 [CLPG18b], and using the chain rule of relative entropy (23), we have:

$$\begin{aligned}
\mathrm{EP}_{\mathcal{L}_{A\partial}^{\mathrm{depol}}}(\rho_t) &\geq D(\rho_t\|\,\mathbb{E}_{A\partial}(\rho_t)) = D(\rho_t\|d_\Lambda^{-1}\,I_\Lambda) - D(\mathbb{E}_{A\partial}(\rho_t)\|d_\Lambda^{-1}\,I_\Lambda)\\
&= D(\rho_t\|d_\Lambda^{-1}\,I_\Lambda) - D(\mathbb{E}_{A\partial}(E^0_A(\rho_t))\|\,\mathbb{E}_{A\partial}(d_\Lambda^{-1}\,I_\Lambda))\\
&\overset{\mathrm{DPI}}{\geq} D(\rho_t\|d_\Lambda^{-1}\,I_\Lambda) - D(E^0_A(\rho_t)\|d_\Lambda^{-1}\,I_\Lambda)\\
&= D(\rho_t\|E^0_A(\rho_t))
\end{aligned}$$

We used that $\mathbb{E}_{A\partial}\circ E_A = \mathbb{E}_{A\partial}$ (which follows from $\mathbb{E}_{A\partial}\circ\mathcal{A}_{A,\sigma} = \mathbb{E}_{A\partial}$ and (35)), the data-processing inequality (DPI), and the chain rule. Finally, observe that $[E^0_A, \Delta_{e^{-itH}}] = 0$ by the GNS-symmetry of $\mathcal{L}^\beta_A$ which is preserved in the limit, therefore:

$$D(\rho_t\|E^0_A(\rho_t)) = D(\Delta_{e^{-itH}}(\rho)\|\Delta_{e^{-itH}}(E^0_A(\rho))) = D(\rho\|E^0_A(\rho))\,.$$

□

B.6 Weak MLSI for the global Davies semigroup

In this section, we are going to show two weak MLSIs for the global Davies semigroups relying in one case solely on the MCMI decay (c.f. Theorem B.8), while in the other case (c.f. Theorem B.10), we replace the result from Appendix B.5 with an assumption on the local gap of the generator to improve the scaling with system size. Hence the second result relies on the existence of MCMI and a uniformly polynomial local gap. More precisely we require the following *uniform local polynomial gap*: There exists $\mu \in \mathbb{N}_0$ and $C > 0$ such that for all $A \subseteq \Lambda$

$$\lambda(\mathcal{L}_A) \geq C|A|^{-\mu} = \Omega(|A|^{-\mu})\,. \tag{66}$$

The first result will be then used in Appendix C to derive the quasi-rapid Wasserstein mixing under every temperature while the second one improves the scaling quasi-rapid Wasserstein mixing and further gives rapid mixing in trace distance under the assumption of a polynomial local gap. Let us begin with the first result only requiring MCMI decay and no additional assumptions on the local or global gap.

Theorem B.8 (weak MLSI at every β) *In the setting of Appendix A.3 assume that the Gibbs state at inverse temperature β satisfies MCMI decay with constants K, ξ, then the semigroup $\{e^{t\mathcal{L}_\Lambda^D}\}_{t\geq 0}$ at that temperature satisfies the following weak MLSI for $\epsilon \geq KD2^D(2L+1)^D \exp\Big(\frac{2r}{\xi} - \frac{L}{D\xi}\Big) = \mathcal{O}(L^D e^{-\mathcal{O}(L)})$.*

$$D(\rho_t\|\sigma) \leq e^{-\alpha(\epsilon)t} D(\rho\|\sigma) + \epsilon \tag{67}$$

with $\frac{1}{\alpha(\epsilon)} = (D+1)(\chi^0_{\min})^{-1} e^{2gJ} e^{4\beta gJ(2D(2r+\xi\log\frac{KD2^DN}{\epsilon})+1)^D} = \mathcal{O}(\exp\{\mathcal{O}((\log\frac{N}{\epsilon})^D)\})$, where $\rho_t = e^{t\mathcal{L}_\Lambda^D}(\rho)$ for arbitrary $\rho \in \mathcal{S}(\mathcal{H}_\Lambda)$.

Such a lower bound on the admissible ϵ is generic and to be expected. Likewise we can also interpret this as giving a minimal lattice size N for which this MLSI eventually holds for a given fixed ϵ.

Proof. Using Theorem B.3 w.r.t. a coarse-graining with parameters $(k = 2r - \frac{r}{D}, c, l)$ and both Lemma B.6 and Lemma B.7 yields

$$\begin{aligned}
D(\rho\|\sigma) &\leq \sum_{a,i} D(\rho\|E^D_{C_{a,i}}(\rho)) + D2^D K N e^{-\frac{c}{\xi}} \\
&\leq \sum_{a,i} e^{2\beta gJ|C_{a,i}\partial|} e^{2\beta gJ|C_{a,i}\partial\partial|} e^{2gJ} (\chi^0_{\min})^{-1}\, \mathrm{EP}_{\mathcal{L}^D_{C_{a,i}\partial}}(\rho) + D2^D K N e^{-\frac{c}{\xi}} \\
&\leq (D+1)(\chi^0_{\min})^{-1} e^{4\beta gJ(\ell+2r)^D + 2gJ}\, \mathrm{EP}_{\mathcal{L}^D_\Lambda}(\rho) + D2^D K N e^{-\frac{c}{\xi}},
\end{aligned}$$

where we used that $|C_{a,i}\partial| \leq |C_{a,i}\partial\partial| \leq (\ell+2r)^D$ by construction of this coarse-graining (c.f. Lemma B.2). Picking $c = \max\{\xi\log\frac{KD2^DN}{\epsilon}, r\} = \mathcal{O}(\log\frac{N}{\epsilon})$ and $\ell = 2D(k+c)+1 = 2D(2r+c)+1-2r = \mathcal{O}(\log\frac{N}{\epsilon})$ gives a valid coarse-graining iff $c, \ell = 2D(k+c)+1 \leq 2L+1$, since $r \leq k \leq 2r \leq 2L+1$. This is equivalent to $c \leq \frac{L}{D} - k$, which is implies by $c \leq \frac{L}{D} - 2r$ and thus $\epsilon \geq KD2^DN\exp\Big(\frac{2r}{\xi} - \frac{L}{D\xi}\Big) \geq KD2^D(2L+1)^D\exp\Big(\frac{2r}{\xi} - \frac{L}{D\xi}\Big)$. We get

$$D(\rho\|\sigma) \leq \frac{(D+1)e^{2gJ}}{\chi^0_{\min}} e^{4\beta gJ(2D(2r+\xi\log\frac{KD2^DN}{\epsilon})+1)^D}\, \mathrm{EP}_{\mathcal{L}_\Lambda}(\rho) + \epsilon \tag{68}$$

$$= \mathcal{O}\left(\exp\left(\mathcal{O}\left(\left(\log\frac{N}{\epsilon}\right)^D\right)\right)\right) \mathrm{EP}_{\mathcal{L}_\Lambda}(\rho) + \epsilon. \tag{69}$$

Applying Grönwalls lemma now yields the desired statement. □

For the second result, let us first convert the uniform polynomial local gap into a uniform polynomial cMLSI of the local generators $\mathcal{L}_A$:

Lemma B.9 *In the context of Appendix A.3 a uniform polynomial lower bound on the local gap with constants C and μ implies a uniform lower bound on the cMLSI constant of the form*

$$\alpha_c(\mathcal{L}_A) \geq \frac{C}{2\log 10 + 2(2\beta gJ + 3\log d)|A\partial|}|A|^{-\mu} \geq \Omega(|A|^{-\mu-1})\,. \tag{70}$$

Proof. The bound is a consequence of (29) and a suitable bound on the Pimsner Popa index in the setting of local Davies generators. Let us denote with $E_A = \lim_{t\to\infty} e^{t\,\mathcal{L}_A}$. In [GR22] the authors showed that

$$C_{\mathrm{cb}}(E_A) \leq \left\|\tau^{-1}\right\|_\infty \sum_{i=1}^n d_{\mathcal{K}_i}^2\,, \tag{71}$$

where

$$E_A(\rho) = \bigoplus_{i=1}^n \mathrm{tr}_{\mathcal{K}_i}[P_i\rho P_i]\otimes\tau_i \quad \text{and} \quad \tau = \bigoplus_{i=1}^n I_{\mathcal{H}_i}\otimes\tau_i\,. \tag{72}$$

is the decomposition of the Davies channel. As E_A only acts non-trivially on $A\partial$, we can conclude that $\sum_{i=1}^n d_{\mathcal{K}_i} \leq d_{A\partial} = d^{|A\partial|}$. By superadditivity of $x\mapsto x^2$, we hence get $\sum_{i=1}^n d_{\mathcal{K}_i}^2 \leq \big(\sum_{i=1}^n d_{\mathcal{K}_i}\big)^2 \leq d^{2|A\partial|}$. To estimate $\left\|\tau^{-1}\right\|$ note that

$$E_A(d_{A\partial}^{-1}\,I_\Lambda) = \bigoplus_{i=1}^n \frac{d_{\mathcal{K}_i}}{d_{A\partial}}\,I_{\mathcal{H}_i}\otimes\tau_i \leq \bigoplus_{i=1}^n I_{\mathcal{H}_i}\otimes\tau_i = \tau\,,$$

hence $\left\|\tau^{-1}\right\|_\infty \leq d^{|A\partial|}\left\|(E_A(I_\Lambda))^{-1}\right\|_\infty \leq d^{|A\partial|}e^{2\beta gJ|A\partial|}$. The last inequality follows by the same argument as in the proof of Lemma B.6. □

Theorem B.10 (weak MLSI under polynomial gap) *In the setting of Appendix A.3 assume that the local Davies generators satisfy the polynomial local gap assumption for some $C>0, \mu\in\mathbb{N}_0$ and further that the Gibbs state satisfies MCMI decay with constants K,ξ both for a fixed temperature β, then the semigroup $\{e^{t\,\mathcal{L}^D_{\Lambda_L}}\}_{t\geq 0}$ at that temperature satisfies the following weak MLSI for $\epsilon \geq KD2^D(2L+1)^D\exp\Big(\frac{r}{\xi}-\frac{L}{D\xi}\Big) = \mathcal{O}(L^De^{-\mathcal{O}(L)})$.*

$$D(e^{t\,\mathcal{L}^D_{\Lambda_L}}(\rho)\|\sigma) \leq e^{-\alpha(\epsilon)t}D(\rho\|\sigma)+\epsilon \tag{73}$$

with $\frac{1}{\alpha(\epsilon)} \leq \frac{D+1}{C}(5+4\beta gJ+8\log d)\left(2D\left(r+\xi\log\frac{KD^2DN}{\epsilon}\right)\right)^{D(1+\mu)} = \mathcal{O}\left(\left(\log\frac{N}{\epsilon}\right)^{D(1+\mu)}\right)$, where $\rho_t = e^{t\,\mathcal{L}^D_{\Lambda_L}}(\rho)$ for arbitrary $\rho\in\mathcal{S}(\mathcal{H}_\Lambda)$.

Remark 4. Note that instead of a uniform polynomial local gap, one can also require very high temperature, i.e. $\beta\sim\frac{1}{N}$ leading to a constant correction in Lemma B.6, giving a result analogous to the one above. This means at such very high temperatures, one gets the strengthened Theorem C.2 only from the uniform decay of MCMI.

Proof of Theorem B.10. We fix a valid coarse-graining with constants (k,c,l) to be chosen later.

Then by Theorem B.3 and the MCMI decay assumed in the statement of the theorem we have

$$\begin{aligned} D(\rho\|\sigma) &\leq \sum_{a,i} D(\rho\|E_{C_{a,i}}(\rho)) + KD2^D Ne^{-\frac{c}{\xi}} \\ &\leq \sum_{a,i} \frac{1}{\alpha(\mathcal{L}_{C_{a,i}})} \operatorname{EP}_{\mathcal{L}_{C_{a,i}}}(\rho) + KD2^D Ne^{-\frac{c}{\xi}} \\ &\leq \frac{D+1}{C}[2\log 10 + (4\beta gJ + 6\log d)\ell^D]\ell^{D\mu} \operatorname{EP}_{\mathcal{L}_\Lambda}(\rho) + KD2^D Ne^{-\frac{c}{\xi}}. \end{aligned}$$

Now choosing $k = r, c = \xi \log \frac{KD2^D N}{\epsilon} = \mathcal{O}\left(\log \frac{N}{\epsilon}\right)$, and $\ell = 2D(r+c)+1 = \mathcal{O}\left(\log \frac{N}{\epsilon}\right)$ gives a valid coarse-graining iff $c, \ell = 2D(k+c)+1 \leq 2L+1$. This is equivalent to $c \leq \frac{L}{D} - r$ and thus $\epsilon \geq KD2^D N \exp\left(-\frac{c}{\xi}\right) \geq KD2^D(2L+1)^D \exp\left(\frac{r}{\xi} - \frac{L}{D\xi}\right)$. We get

$$D(\rho\|\sigma) \leq \alpha(\epsilon) \operatorname{EP}_{\mathcal{L}_\Lambda}(\rho) + \epsilon,$$

where

$$\begin{aligned} \frac{1}{\alpha(\epsilon)} &= \frac{D+1}{C}[2\log 10 + (4\beta gJ + 6\log d)\ell^D]\ell^{D\mu} \\ &\leq \frac{D+1}{C}\ell^{D(1+\mu)}(5 + 4\beta gJ + 6\log d) \\ &= \frac{D+1}{C}(5 + 4\beta gJ + 6\log d)\left(2D(r + \xi\log\frac{KD2^D N}{\epsilon}) + 1\right)^{D(1+\mu)} \\ &= \mathcal{O}\left(\left(\log\frac{N}{\epsilon}\right)^{D(1+\mu)}\right), \end{aligned}$$

where we used that $\ell \geq 1$ and $\max\{2\log 10, 6\log d\} < 8\log d$ with $d \geq 2$. Integration now gives the statement of the theorem. □

C Main results - extended

C.1 W_1-mixing from MCMI-decay

Here we state and prove the main result of this work concerning the quasi-rapid W_1-decay.

Theorem C.1 (Quasi rapid Wasserstein mixing) *Let $\mathcal{L}_\Lambda^D$ be a Davies Lindbladian at inverse temperature $\beta > 0$ corresponding to a $(\kappa, r)-$local, J-bounded, commuting Hamiltonian H_Λ. Denote the growth constant of H_Λ on $\Lambda = [\![-L, L]\!]^D$ with g. Then if the Gibbs state of H_Λ with the same temperature (the invariant state of $\mathcal{L}_\Lambda^D$) satisfies uniform exponential decay of the MCMI with constants $K, \xi > 0$ the semi group generated by $\mathcal{L}_\Lambda^D$ is quasi-rapidly mixing in normalized W_1-distance* *with mixing time*

$$\begin{aligned}
t^{W_1}_{mix}(\varepsilon) &= \frac{(D+1)e^{2gJ}}{\chi^0_{\min}} \\
&\times \exp\left(4\beta gJ\left[2D\left(2r+\xi\log\left\{\frac{64KD(D+1)2^D}{\varepsilon^2}\left(2D\left(r+\xi\log\frac{8D2^DKN}{\varepsilon^2}\right)+1\right)^{2D}\right\}\right)+1\right]^D\right) \\
&\times \log\left\{\frac{64(2\beta gJ+\log d)(D+1)}{\varepsilon^2}\left(2D\left(r+\xi\log\left\{\frac{8NKD2^D}{\varepsilon^2}\right\}\right)+1\right)^{2D}\right\} \\
&= \mathcal{O}(\exp\left(\text{poly}\left(\log\left(\frac{1}{\varepsilon^2}\log\frac{N}{\varepsilon^2}\right)\right)\right)) \\
&= \text{quasi-poly}\left(\frac{1}{\varepsilon^2}\text{poly}\log\frac{N}{\varepsilon^2}\right) = \text{quasi-poly}(\varepsilon^{-1})_{\varepsilon\to 0}\text{quasi-log}(N)_{N\to\infty}
\end{aligned}$$

whenever $\varepsilon \geq 8N\sqrt{(D+1)KD2^D}\exp\left(\frac{r}{\xi}-\frac{N^{\frac{1}{D}}-1}{4D\xi}\right) = \mathcal{O}\left(N\exp\left(-\mathcal{O}(\sqrt[D]{N})\right)\right)$ *is fixed.*

Remark 5. Note that the requirement of $\epsilon \geq \mathcal{O}\left(N\exp\left(-\mathcal{O}(\sqrt[D]{N})\right)\right)$ is not of relevance in implying quasi-rapid mixing, since the property of quasi-rapid mixing is only determined by the asymptotic scaling of the mixing time $t^{W_1}_{\text{mix}}(\varepsilon)$ in the system size, for any fixed ε. The asymptotics are, however, not affected by this requirement since they hold eventually for any fixed $\varepsilon > 0$ as $\lim_{N\to\infty}\mathcal{O}(N\exp\left(-\mathcal{O}(\sqrt[D]{N})\right)) = 0$. It is also a crude bound on the actual tight one scaling as $\varepsilon \geq \mathcal{O}\left((\log N)^D\exp\left(-\mathcal{O}(\sqrt[D]{N})\right)\right)$, as can be seen in the proof.

The proof will follow from the combined application of Theorem B.5 and Theorem B.8, taking care of the explicit constants along the way.

Proof of Theorem C.1. We set $N := |\Lambda|$ in the following. First, we choose a valid (k,c,ℓ) coarse-graining w.r.t which the weak AT Theorem B.3 implies the following weak TC due to Theorem B.5.

$$\begin{aligned}
\|\rho_t-\sigma\|_{W_1} &\leq \max_{a,i} 2\sqrt{2}|C_{a,i}|\sqrt{N(D+1)D(\rho_t\|\sigma)} + N\sqrt{2D2^DKN}e^{-\frac{c}{2\xi}} \\
&\leq 2\sqrt{2}\ell^D\sqrt{D+1}\sqrt{ND(\rho_t\|\sigma)} + N\sqrt{N}\sqrt{2D2^DK}e^{-\frac{c}{2\xi}}.
\end{aligned}$$

Now choosing $k=r, c=\xi\log\frac{2D2^DKN}{\delta^2}$ and $\ell = 2D(r+c)+1 = 2D\left(r+\xi\log\frac{2D2^DKN}{\delta^2}\right)+1 = \mathcal{O}\left(\log\frac{N}{\delta^2}\right)$, yields

$$\|\rho_t-\sigma\|_{W_1} \leq 2\sqrt{2(D+1)}\sqrt{Nl^{2D}D(\rho_t\|\sigma)} + N\delta.$$

Inserting the following weak MLSI from Theorem B.8, which is implied by the uniform MCMI decay:

$$D(\rho_t\|\sigma) \leq e^{-\alpha(N\ell^{-2D}\epsilon)t}D(\rho\|\sigma) + \frac{N}{\ell^{2D}}\epsilon \leq Ne^{-\alpha(N\ell^{-2D}\epsilon)t}(2\beta gJ+\log d) + \frac{N}{\ell^{2D}}\epsilon,$$

where

$$\frac{1}{\alpha(N\ell^{-2D}\epsilon)} = \frac{D+1}{\chi^0_{\min}} e^{2gJ} e^{4\beta gJ\left[2D\left(2r+\xi\log\frac{KD2^D\ell^{2D}}{\epsilon}\right)+1\right]^D}$$
$$= \frac{D+1}{\chi^0_{\min}} e^{2gJ} e^{4\beta gJ\left[2D\left(2r+\xi\log\frac{KD2^D\left(2D\left(r+\xi\log\frac{2D2^DKN}{\delta^2}\right)+1\right)^{2D}}{\epsilon}\right)+1\right]^D}.$$

Combining this with the weak MLSI above yields

$$\|\rho_t-\sigma\|_{W_1} \le 2\sqrt{2}\sqrt{D+1}N\sqrt{e^{-\alpha(N\ell^{-2D}\epsilon)t}\ell^{2D}(2\beta gJ+\log d)+\epsilon}+N\delta\,.$$

So for times t larger than

$$\frac{1}{\alpha(N\ell^{-2D}\epsilon)}\log\frac{(2\beta gJ+\log d)\ell^{2D}}{\epsilon}$$
$$= \frac{D+1}{\chi^0_{\min}} e^{2gJ} e^{4\beta gJ\left[2D\left(2r+\xi\log\left\{\frac{KD2^D}{\epsilon}\left(2D\left(r+\xi\log\frac{2D2^DKN}{\delta^2}\right)+1\right)^{2D}\right\}\right)+1\right]^D}$$
$$\times\log\left\{\frac{(2\beta gJ+\log d)}{\epsilon}\left(2D\left(r+\xi\log\frac{2D2^DKN}{\delta^2}\right)+1\right)^{2D}\right\}$$
$$= \mathcal{O}\left(\exp\left(\mathrm{poly}_D\left(\log\frac{1}{\epsilon}\log\frac{N}{\delta^2}\right)\right)\right),$$

it holds that

$$\|\rho_t-\sigma\|_{W_1} \le 2\sqrt{2(D+1)}N\sqrt{2\epsilon}+N\delta \le N\varepsilon,$$

where we set $\epsilon := \frac{\delta^2}{16(D+1)}$ and $\delta = \frac{\varepsilon}{2}$ completes the proof with the mixing time claimed in the statement of the theorem. The bound on the minimal value of comes from the one in Theorem B.8, rewritten in terms of $N = (2L+1)^D$, yielding

$$\delta \ge \sqrt{2D2^DKN}\exp\left(\frac{r}{2\xi}-\frac{\sqrt[D]{N}-1}{4D\xi}\right)$$
$$\epsilon \ge KD2^D\ell^{2D}\exp\left(\frac{2r}{\xi}-\frac{\sqrt[D]{N}-1}{2D\xi}\right).$$

Crudely bounding $\ell^D \le N$ and $\varepsilon = 8\sqrt{D+1\epsilon} = 2\delta$ yields the result, i.e. it is easy to check that the in the theorem claimed bound on ε satisfies both the inequalities above. □

C.2 Rapid mixing from MCMI-decay and polynomial local gap

Theorem C.2 (Rapid mixing and hyper rapid W_1 mixing under polynomial local gap) *Let $\mathcal{L}^D_\Lambda$ be a Davies Lindbladian corresponding to a (κ, r)−local, J-bounded, commuting Hamiltonia H_Λ, to inverse temperature $\beta > 0$. Denote the growth constant of H_Λ on $\Lambda = [\![-L, L]\!]^D$ with g. Then if the Gibbs state of H_Λ to inverse temperature β (the invariant state of $\mathcal{L}^D_\Lambda$) satisfies uniform exponential decay of the MCMI with constants $K, \xi > 0$ and the gap of the local Davies generators are at most* *polynomially decaying in local region size with degree $\mu \in \mathbb{N}_0$, the semigroup generated by $\mathcal{L}_\Lambda$ is rapidly mixing in trace-distance with mixing time*

$$\begin{aligned} t^1_{mix}(\varepsilon) &= \frac{D+1}{C}(5+4\beta gJ+8\log d)\left(2D\left(r+\xi\log\frac{2KD^2DN}{\varepsilon^2}\right)\right)^{D(1+\mu)}\log\frac{4(2\beta gJ+\log d)N}{\varepsilon^2} \\ &= \mathcal{O}\left(\left(\log\frac{N}{\varepsilon^2}\right)^{1+D(1+\mu)}\right) \end{aligned}$$

whenever $\varepsilon \geq \sqrt{2KD2^DN}\exp\Big(\frac{r}{2\xi}-\frac{\sqrt[D]{N}-1}{4D\xi}\Big) = \mathcal{O}(\sqrt{N}\exp\Big(-\mathcal{O}(\sqrt[D]{N})\Big))$. *And it is hyper rapidly mixing in normalized W_1 distance with mixing time*

$$\begin{aligned} t^{W_1}_{mix}(\varepsilon) &= \frac{D+1}{C}(5+4\beta gJ+8\log d) \\ &\times\left(2D\left(r+\xi\log\left(\frac{64KD2^D}{\varepsilon^2}\left(2D(r+\xi\log\frac{8D2^DKN}{\varepsilon^2})+1\right)^{2D}\right)\right)+1\right)^{D(1+\mu)} \\ &\times\log\left(\frac{64(2\beta gJ+\log d)}{\epsilon^2}\left(2D(r+\xi\log\frac{8D2^DKN}{\varepsilon^2})+1\right)^{2D}\right) \\ &= \mathcal{O}\left(\left(\log\left(\frac{1}{\varepsilon^2}\log\frac{N}{\varepsilon^2}\right)\right)^{1+D(1+\mu)}\right) \end{aligned}$$

whenever $\varepsilon \geq 8N\sqrt{(D+1)KD2^D}\exp\left(\frac{r}{2\xi}-\frac{N^{\frac{1}{D}}-1}{4D\xi}\right) = \mathcal{O}\left(N\exp\Big(-\mathcal{O}(\sqrt[D]{N})\Big)\right)$ *is fixed.*

Note here equally the remark on the minimal value of $\varepsilon(N)$ as 5.

Proof. The rapid mixing part follows directly from Theorem B.10, which is implied by the MCMI decay and uniform local gap and an application of Pinsker's inequality. In detail we get

$$\begin{aligned} \|\rho_t-\sigma\|_1 &\leq \sqrt{2D(\rho_t\|\sigma)} \leq \sqrt{2e^{-\alpha(\epsilon)t}D(\rho\|\sigma)+\epsilon} \\ &\leq \sqrt{2e^{-\alpha(\epsilon)t}N(2\beta gJ+\log d))+\epsilon}. \end{aligned}$$

Hence for times t larger than

$$\begin{aligned} &\frac{1}{\alpha(\epsilon)}\log\frac{2(2\beta gJ+\log d)N}{\epsilon} \\ &= \frac{D+1}{C}(5+4\beta gJ+8\log d)\left(2D\left(r+\xi\log\frac{KD^2DN}{\epsilon}\right)\right)^{D(1+\mu)}\log\frac{2(2\beta gJ+\log d)N}{\epsilon} \\ &= \mathcal{O}\left(\log\left(\frac{N}{\epsilon}\right)^{1+D(1+\mu)}\right), \end{aligned}$$

it holds that

$$\|\rho_t-\sigma\|_1 \leq \sqrt{2\epsilon}.$$

Since this time also upper bounds the mixing time, replacing ϵ by $\frac{\varepsilon^2}{2}$ yields the first claim. For the W_1 mixing proceed just as in the proof of Theorem C.1, but with Theorem B.10 instead of Theorem B.8. So fix a valid coarse-graining with constants (k, c, ℓ) and consider the weak TC from Theorem B.5 w.r.t this coarse-graining.

$$\begin{aligned}\|\rho_t-\sigma\|_{W_1} &\leq \max_{a,i} 2\sqrt{2}|C_{a,i}|\sqrt{ND(D+1)D(\rho_t\|\sigma)}+N\sqrt{2KD2^DN}e^{-\frac{c}{2\xi}}\\ &\leq 2\sqrt{2(D+1)}\ell^D\sqrt{N\left(e^{-\alpha(N\ell^{-2D}\epsilon)t}N(2\beta gJ+\log d)+\frac{N}{\ell^{2D}}\epsilon\right)}+N\sqrt{N}\sqrt{2D2^DK}e^{-\frac{c}{\xi}}.\end{aligned}$$

Now setting $k=r, c=\xi\log\frac{2D2^DKN}{\delta^2}=\mathcal{O}\left(\frac{N}{\delta^2}\right)$ and $l=2D(r+c)+1=\mathcal{O}\left(\frac{N}{\delta^2}\right)$ yields a valid coarse-graining and it is easily checked that for times t larger than

$$\frac{1}{\alpha(N\ell^{-2D}\epsilon)}\log\left(\frac{2\beta gJ+\log d}{\epsilon}\ell^{2D}\right)=\mathcal{O}\left(\left(\log\left(\frac{1}{\epsilon}\right)\right)^{1+D(1+\mu)}\right),$$

it holds that

$$\|\rho_t-\sigma\|_{W_1}\leq 2\sqrt{2(D+1)}N\sqrt{2\epsilon}+N\delta.$$

Setting $\epsilon=\frac{\delta^2}{16(D+1)}$ renaming $\delta=\frac{\varepsilon}{2}$ finishes the proof with the in the theorem claimed mixing time. In the first case, the requirement on the minimal size of ε comes from the one in Theorem B.10 rewritten in terms of $N=(2L+1)^D$ and applied to $\varepsilon=\sqrt{2\epsilon}$. In the second case, this is analogous to the proof of Theorem C.1. □

We want to emphasise again that alternatively to the assumption of the existence of a uniform polynomial gap one can also ask for a very high temperature to again reduce the requirements only to the decay of the MCMI (see Remark 4)

C.3 Quasi-Optimal Gibbs state preparation from MCMI-decay

To convert the prior established mixing time bounds of the Davies semi-group into efficiency results of preparing their fixed points, we require (Davies) Lindblad simulation theorems. These give explicit constructions of quantum circuits, with circuit complexity bounds, which approximate $e^{t\mathcal{L}}$ in diamond norm. Such suitable circuits are for example constructed in [CW16, RWW23, LW22, CKBG23], with the latter two being the more efficient ones.

Specifically in [CKBG23, Theorem III.2] the authors construct a Lindblad simulation algorithm for which it was shown that the complexity of implementing $e^{t\mathcal{L}_\Lambda^D}$ in terms of two-qubits gates, ancilla qubits, and block-encodings of the dissipative part and Hamiltonian part (see [CKBG23]) scales linear up to poly logarithmic corrections in $\frac{Nt_{\text{mix}}}{\epsilon}$. Combining this Lindblad simulation algorithm we get the following main result.

Theorem C.3 (Quasi-optimal sampling from Gibbs states that satisfy MCMI decay) *Let σ be a Gibbs states of $(\kappa, r)-$local, commuting, J-bounded Hamiltonian on $\Lambda\subset\mathbb{Z}^D$ which satisfies uniform MCMI decay. Then there exists a quantum algorithm (circuit) that outputs an ϵ-close in normalized W_1-distance, state using*

1. $\mathcal{O}\left(\text{poly}\log\left(N\,\text{quasi-poly}\left(\frac{1}{\epsilon^2}\,\text{poly}\log\frac{N}{\epsilon^2}\right)\right)\right)=\mathcal{O}\left(\text{poly}\log N,\text{poly}\log\frac{1}{\epsilon^2}\right)$ *ancilla qubits,*
2. $\mathcal{O}\left(N\,\text{quasi-poly}\left(\frac{1}{\epsilon^2}\,\text{poly}\log\frac{N}{\epsilon^2}\right)\right)=\mathcal{O}\left(N\text{quasi-log}(N),\text{quasi-poly}(\epsilon^{-2})\right)$ *two-qubit gates, block encodings of the Hamiltonian H_Λ and the dissipative part of $\mathcal{L}_\Lambda^D$.*

Remark 6. Such algorithms are usually referred to as optimal if the complexity and/or runtime scales as $\mathcal{O}(N)$ up to polylogarithmic corrections. Since our scaling is $\mathcal{O}(N)$ up to quasi-logarithmic corrections we call it *quasi-optimal*, since it still scales better than $\mathcal{O}(N^2)$. In [CKBG23] they also explicitly construct these block encodings for the dissipative part of $\mathcal{L}_\Lambda^D$ in terms of simpler gates and the jump operators $S_{\alpha,k}^\omega$. Note though that any efficient Lindbaldian simulation algorithm, such as the one from [LW22], may be used to construct an efficient algorithm to sample from such Gibbs states. Our rapid mixing results imply optimal efficient preparation, analogous to the above, yielding circuit complexities and runtimes of $\mathcal{O}(N)$ up to poly-logarithmic corrections.

Proof. The above bounds of the number of ancilla quits and necessary gates follow directly from implementing a quantum circuit $\mathcal{C}_\epsilon$ that ϵ approximates $e^{t\,\mathcal{L}_\Lambda^D}$ in diamond norm and Theorem C.1. The former is done in, e.g. [CKBG23, Theorem III.2] and gives the in the theorem mentioned bound when substituting their t for $\mathcal{O}(N)t_{\text{mix}}^{W_1}(\varepsilon)$, since they consider normalized Lindbaldians, see (1.10) in [CKBG23]. This suffices since any quantum circuit that is ϵ close in diamond norm to $e^{t\,\mathcal{L}}$ also is ϵ-close in stabilized $1 \to \tilde{W}_1$ norm, where $\tilde{W}_1$ denotes the normalized W_1 distance. I.e. let ρ be the output of this circuit $\mathcal{C}_\epsilon$, then

$$\begin{aligned}\frac{1}{N}\|\rho-\sigma\|_{W_1} &\le \frac{1}{N}\|\rho - e^{t_{\text{mix}}^{W_1}(\varepsilon)\,\mathcal{L}_\Lambda^D}(\rho)\|_{W_1} + \frac{1}{N}\|e^{t_{\text{mix}}^{W_1}(\varepsilon)\,\mathcal{L}_\Lambda^D}(\rho)-\sigma\|_{W_1}\\ &\le \|\rho - e^{t\,\mathcal{L}_\Lambda^D}(\rho)\|_1 + \varepsilon \le \|\mathcal{C}_\epsilon - e^{t_{\text{mix}}^{W_1}(\varepsilon)\,\mathcal{L}_\Lambda^D}\|_\diamond \|\rho\|_1 + \varepsilon \le \epsilon + \varepsilon.\end{aligned}$$

We conclude by setting $\epsilon = \varepsilon$ and rescaling. □

D Examples

In this section, we study systems that exhibit MCMI decay, which consequently leads to quasi-rapid Wasserstein mixing, and, under a uniform local polynomial gap condition further to hyper-rapid Wasserstein mixing and rapid mixing. We begin by deriving MCMI decay from the existence of a strong effective Hamiltonian in Appendix D.1, that exhibits a uniform bound in interaction norm. Next, in Appendix D.2, we assume the system Hamiltonian is marginal commuting, which, based on a result from [BCPH24], directly implies the existence of a strong effective Hamiltonian at high temperatures with a specified decay rate. By applying the result from Appendix D.1, we thus establish MCMI decay at high temperatures. We know that this commuting marginal assumption is fulfilled for every Hamiltonian composed of commuting Pauli strings (e.g. the Toric code in any dimension)[2]. Lastly, note that in all the following results, we will not restate that MCMI decay implies quasi-rapid Wasserstein mixing, which, under a uniform local polynomial gap, can be strengthened to hyper-rapid Wasserstein mixing and supplemented by rapid mixing in trace distance. If wanted the constants of the MCMI decay of all results below could be inserted into the result in Appendix C.1 and Appendix C.2 to obtain explicit decay rates.

D.1 MCMI-decay from strong effective Hamiltonian

The following lemma adapts the proof of [BCPH24, Theorem 4.1] to show the uniform decay of the MCMI (c.f. (38)) given the existence of an effective Hamiltonian with a uniform bound on its interaction norm.

[2] This result is shown by Sebastian Stengele in some currently unpublished notes.

Lemma D.1 (MCMI-decay from locality of effective Hamiltonian) *In the setting of Appendix A.3 assume the existence of a strong effective Hamiltonian for H_Λ, with a uniform bound on the interaction norm, i.e.*

$$\left\|\left|\tilde{H}^A\right|\right\|_\mu := \sup_{x\in\Lambda} \sum_{X\subseteq\Lambda : x\in X} \|\tilde{h}_X^A\|_\infty e^{\mu\,\mathrm{diam}(X)} \le \Delta \tag{74}$$

for $\mu, \Delta \ge 0$ independent of $A \subseteq \Lambda$. Then for any partition $\Lambda = A \sqcup B \sqcup C \sqcup D$

$$\|\boldsymbol{H}(A:C|D)_\sigma\|_\infty \le 4\min\{|A|,|C|\}\Delta e^{-\mu\,\mathrm{dist}(A,C)}\,. \tag{75}$$

Proof. Without loss of generality, we can assume that $|A| \le |C|$. Inserting the definition of the MCMI and using 3. from the definition of a strong effective Hamiltonian in Appendix A.4.3 we get

$$\begin{aligned}
\|\mathbf{H}(A:C|D)_\sigma\|_\infty &= \Big\| \sum_{X\subseteq\Lambda} \tilde{h}_X^{ACD} + \tilde{h}_X^D - \tilde{h}_X^{AD} - \tilde{h}_X^{CD} \Big\|_\infty \\
&= \Big\| \sum_{\substack{X\subseteq\Lambda:\\ X\cap A\neq\emptyset, X\cap C\neq\emptyset}} \tilde{h}_X^{ACD} + \tilde{h}_X^D - \tilde{h}_X^{AD} - \tilde{h}_X^{CD} \Big\|_\infty \\
&\le \sum_{\substack{X\subseteq\Lambda:\\ X\cap A\neq\emptyset, X\cap C\neq\emptyset}} \left\|\tilde{h}_X^{ACD}\right\|_\infty + \left\|\tilde{h}_X^{D}\right\|_\infty + \left\|\tilde{h}_X^{AD}\right\|_\infty + \left\|\tilde{h}_X^{CD}\right\|_\infty
\end{aligned}$$

The first equality stems from the fact that

$$\tilde{h}_X^{ACD} + \tilde{h}_X^D - \tilde{h}_X^{AD} - \tilde{h}_X^{CD} = 0$$

if $X \cap C = \emptyset$ or if $X \cap A = \emptyset$. Indeed, if $X \cap C = \emptyset$, then $X \cap ACD = X \cap AD$ and $X \cap CD = X \cap D$ which respectively give $\tilde{h}_X^{ACD} = \tilde{h}_X^{AD}$ and $\tilde{h}_X^{CD} = \tilde{h}_X^{D}$ by 2. of Appendix A.4.3. The argument is analogous for $X \cap A = \emptyset$ by symmetry in A and C. Using (74) we obtain

$$\begin{aligned}
\|\mathbf{H}(A:C|D)_\sigma\|_\infty &\le \sum_{\substack{X\subseteq\Lambda:\\ X\cap A\neq\emptyset, X\cap C\neq\emptyset}} e^{-\mu\,\mathrm{diam}(X)} e^{\mu\,\mathrm{diam}(X)} \left(\left\|\tilde{h}_X^{ACD}\right\|_\infty + \left\|\tilde{h}_X^{D}\right\|_\infty + \left\|\tilde{h}_X^{AD}\right\|_\infty + \left\|\tilde{h}_X^{CD}\right\|_\infty\right) \\
&\le \sum_{x\in A} e^{-\mu\,\mathrm{dist}(x,C)} \sum_{\substack{X\subseteq\Lambda:\\ x\in X, X\cap C\neq\emptyset}} e^{\mu\,\mathrm{diam}(X)} \left(\left\|\tilde{h}_X^{ACD}\right\|_\infty + \left\|\tilde{h}_X^{D}\right\|_\infty + \left\|\tilde{h}_X^{AD}\right\|_\infty + \left\|\tilde{h}_X^{CD}\right\|_\infty\right) \\
&\le 4\Delta|A|e^{-\mu\,\mathrm{dist}(A,C)}\,,
\end{aligned}$$

proving the claim. □

D.2 MCMI-decay from commuting marginals at high temperature

Assuming that the system Hamiltonian H_Λ is marginal commuting now immediately gives an explicit uniform decay of the MCMI going through the result of [BCPH24, Theorem 3.8]. This is detailed in the theorem below.

Theorem D.2 (MCMI decay for marginal commuting systems at high temperature) *In the setting of Appendix A.3, assume that H_Λ is marginal commuting (c.f Appendix A.4.3 for a definition). Then, for $\beta < \frac{1}{\kappa g(1+\kappa g)e^2 gJ}$ and for every partition $\Lambda = A \sqcup B \sqcup C \sqcup D$, we have*

$$\|\boldsymbol{H}_\sigma(A:C|D)\|_\infty \le 4\min\{|A|,|C|\}e^{-\mu\,\mathrm{dist}(A,C)}$$

with $\mu = \frac{1}{r}\log\left(\frac{1}{g\kappa(1+g\kappa)e^2 gJ\beta}\right)$.

Proof. Assume that $\beta < \frac{1}{\kappa g(1+\kappa g)e^2 gJ}$ and set $\mu_\epsilon = \frac{1+\epsilon}{r} \log\left(\frac{1}{g\kappa(1+g\kappa)e^2 gJ\beta}\right) > 0$ for $\epsilon > 0$. Then

$$\left\|\left\| H_\Lambda \right\|\right\|_{\mu_\epsilon} \leq gJe^{\mu_\epsilon r} < \frac{1}{g\kappa(1+g\kappa)e^2\beta}\,.$$

Reordering the above gives

$$\beta < \frac{1}{g\kappa(1+g\kappa)e^2 \left\|\left\| H_\Lambda \right\|\right\|_{\mu_\epsilon}} \leq \frac{1}{\mathfrak{d}(1+\mathfrak{d})e^2 \left\|\left\| H_\Lambda \right\|\right\|_{\mu_\epsilon}}\,. \tag{76}$$

In the last inequality, we used $\mathfrak{d} \leq g\kappa$, where $\mathfrak{d}$ is the degree of the interaction graph (see [BCPH24]). Combining this temperature constraint with the commutativity of the generated algebra and its preservation under partial traces, all conditions of [BCPH24, Theorem 3.8] are satisfied. We can therefore conclude the existence of a strong effective Hamiltonian with decay

$$\left\|\left\| \tilde{H}^A \right\|\right\|_{\mu_\epsilon} \leq 1$$

for every $A \subseteq \Lambda$. Now, by Lemma D.1, we immediately conclude the claim with decay rate μ_ϵ. Taking $\epsilon \to 0$ then concludes the claim with the stated decay rate. □