Final Thesis for the Degree Master of Science
as part of
TMP Elite Master Course
Theoretical and Mathematical Physics

Static and Dynamic properties of quantum spin systems at non-zero T°

Jan Felix Kochanowski

August 12, 2023

Supervisor: Prof. Dr. Ignacio Cirac
Second Examiner: Dr. Cambyse Rouzé

Acknowledgements

I am very much indebted to my incredible supervisor Cambyse Rouzé, who always supported and pushed me and without whom this piece of work would not be possible in the slightest. I am very thankful for him believing and giving me the freedom, but also help I needed and for the great ideas and discussions we had. I could not have had a better supervisor for this thesis and I am very much looking forward to start working on my PhD under his supervision. Thank obviously goes out equally to Angela Capel and Alvaro Alhambra, my other two collaborators, who it was really fun and inspiring to work with and without who this wouldn't be possible either. I also want to thank Ignacio Cirac and all the great people at the MPQ who welcomed me there and provided a very supportive and diverse environment for me to work in. Further I want to thank Nilanjana Datta for bringing me in contact with Cambyse. Thank goes out also to my dearest friends Anna, Arthur, Ayse, Jakob, Andrea, Jonas, Merle, Lukas, and many more for their support firedship, the fun and equally the inspiring moments we had together, and their help and emotional support in tougher times. You all mean really a lot to me. But most of all, I want to thank my parents. Thank you Edeltraud and Bernd for your loving support, without which I definitely wouldn't be where I am at right now. I cannot put my dept into words. Thank you ever so much!
And finally I want to thank the 'Studienstiftung des deutschen Volkes', the 'TUM:Junge Akademie', the University of Cambridge, and my study-course TMP through which I am a member of the 'Elitenetzwerk Bayern'. These are four institutions which I am very indebted to for the incredible opportunities, support, great networks, and long lasting friendships they provided and provide me with.

Thesis advisor: Prof. Dr. Ignacio Cirac, Dr. Cambyse Rouzé *Jan Felix Kochanowski*

Static and Dynamic properties of quantum spin systems at non-zero T°

HIGH LEVEL SUMMARY

Thermalization of quantum spin systems is of large interest to the quantum information- [1]–[4] and many-body-physics-communities [5], [6]. A body of works recently approached rapid thermalization of quantum spin systems via exponential convergence of the relative entropy between the evolved state and the Gibbs equilibrium state of the system via the development of so called complete modified logarithmic Sobolev (cMLSI) inequalities [7]–[10]. The decay rate of this convergence can be characterlized by the cMLSI constant governing this inequality. Amongst others it was shown that for 1D commuting or 2-local commuting quantum spin systems a strictly positive cMLSI constant α, exists for any system size. But in general it is monotonically decreasing in system size. [11], [12]. We add to this work by showing amongst others that a 2-local, commuting quantum spin system on an arbitrary sub-exponential lattice has a positive cMLSI-constant, independent of lattice size, whenever the Lindbladian (of the evolution) is gaped. We also affirmatively show that weak clustering, i.e. exponential decay of correlations, in this setting is sufficient to guarantee a gaped Davies generator. This partially answers a long standing open question from [3] for 2-local systems. We do this via a novel concept we call strong local indistinguishability and with it show that weak clustering ($\mathbb{L}_\infty$) is equivalent to ($q\mathbb{L}_1 - \mathbb{L}_\infty$)-clustering, for certain systems. This also implies a strong form of approximate tensorization of the relative entropy for any 2-local quantum spin system and we also establish equivalence between decay of correlations ($\mathbb{L}_\infty$-clustering) and decay of mutual information for Gibbs states of commuting Hamiltonians, extending work from [7]. Applying the main result to geometrically-local, commuting quantum spin chains and b-ary trees yields existence of a systems-size independent cMLSI constant at any temperature in the former and a logarithmically decreasing one in the latter if the Davies Davies Lindbladian is gaped. This implies rapid thermalization of any uniform, commuting, 2-local system with gaped Davies Lindbladians. Amongst others this also leads to new optimal Gaussian concentration bounds, more general ensemble equivalences, and tighter entropy difference bounds.

Contents

1 Introduction **1**
1.1 Thesis outline . . . 2

2 Preliminaries **4**
2.1 Graphs . . . 4
2.2 Some general notation . . . 5
2.3 Weighted non-commutative $\mathbb{L}_{p,\sigma}$-spaces and inner products . . . 6
2.4 Uniform families of Hamiltonians . . . 7
2.5 Quantum Markov semigroups and Lindbladians . . . 8
2.6 Conditional Expectations . . . 9
2.7 The relative entropy and strong data processing . . . 11
2.8 Relative entropy decay via the complete modified logarithmic Sobolev Inequality 12
2.9 Variance Decay and Gap . . . 13
2.10 Rapid Thermalization . . . 13

3 Clustering and Strong local indistinguishability **15**
3.1 A useful relation . . . 16
3.2 Local and Strong local indistinguishability . . . 17

4 Local Generators and conditional expectations **25**
4.1 Davies Evolution and Local Davies Generators . . . 25
4.2 Schmidt Conditional Expectations . . . 27
4.3 Relating Davies and Schmidt dynamics . . . 31

5 Main results: Gap implies cMLSI **32**
5.1 Approximate tensorization for an almost classical state ω via weak clustering . . 35
5.2 Geometric argument . . . 40
5.3 Putting everything together . . . 45

6 Applications **47**
6.1 Exponential convergence to Gibbs states in the thermodynamic limit . . . 47
6.2 Tighter bounds on the entropy difference and convergence via relative entropy . . 48
6.3 Optimal 1D Gaussian concentration bound . . . 49
6.4 Ensemble equivalence under (long-range) Lipschitz observables . . . 50

7 Summary and Outlook **51**
7.1 Summary and future work . . . 51

Bibliography **54**

List of Figures **58**

Appendix **60**
A.1 Omitted proofs from the main text 60
A.2 Finite-size-coarse-graining example 63

1 Introduction

Dynamics of open quantum systems are of large interest to the quantum optics, condensed matter, chemical physics, and mathematical physics communities [13]. One very important class of such open quantum evolutions are Markovian ones, which are of large interest to the quantum information community as well as all other just mentioned ones. Such Markovian dynamics describe amongst others thermalization of quantum systems [2], and hence are also important in quantum algorithms that sample thermal or ground states of certain Hamiltonians or are supposed to simulate physical thermalization of quantum lattice systems. Quantum Markov Semigroups (QMS) which describe thermalization are also called quantum Gibbs samplers, since the thermal equilibrium state is also called the *Gibbs state* of a system. These are a mathematical tool, which can be studied systematically in high generality. They form the quantum analogue of the seminal Monte Carlo algorithm and are in contrast to the often considered heuristic approaches of Gibbs state preparation such as the variational eigensolver or adiabatic algorithms. One important example, and the main one we are considering in this work, is the *Davies evolution* [13], [14]. It describes the physical process of thermalization[1] of a quantum system weakly coupled to its environment. Hence it is often considered in the mathematical physics community when describing thermalization processes. It is, hence, also often used as a sub-routine for certain types of quantum algorithms. [2], [3], [7], [11]. Quantum lattice-spin systems are also some of the most promising candidate architectures for quantum computing and quantum information storage. For these reasons, understanding their thermalization is of great importance, especially in the current regime of noisy intermediate scale quantum systems, where the number of qubits and their coherence times are too small/short to enable the use of quantum error correction. Hence thermalization of quantum lattice systems is also of interest to the computer science community. There are many heuristic and numerical results about thermalization of certain physical systems, but an overarching theory of thermalization of quantum systems does not exist in the mathematics literature, however, there is a rapidly growing body of work. This piece of work is a step towards establishing such a theory.

Quantum Markovian evolutions, just like their classical counterparts, are known to mix under certain conditions, that is that the evolution approaches certain fixed points. When considering dynamics which describe thermalization, this fixed point is the thermal equilibrium state of the spin system, called the Gibbs state of the system. This work is concerned with the quest of finding quantitative bounds on the mixing times of Davies evolutions $t_{\text{mix}}(\epsilon)$. This is how quickly these evolutions evolve any initial state ϵ-close to the stationary state. We will be doing this via entropic and functional inequalities that these evolutions follow. Systems for which the mixing time scales logarithmically with system size are said to be *rapidly mixing*. In turn rapid mixing is known to imply many static and dynmaical properties for these systems, among which stability of these evolutions under local perturbations [15], existence of area laws for the mutual information of fixed points of such evolutions [16], efficient preparation and simulability of the fixed points [2], and even rigorous connections to the elusive eigenstate thermalization hypoethsis (ETH) [17], [18]. Our main result directly applies to these.

Quantum Gibbs samplers find also applications as sub-routines in larger algorithms, like ones

[1] though this is sometimes questioned since its derivation requires a few non-trivial assumptions see e.g. [2], [13]

solving semidefinite programs [19], [20], quantum machine learning [21], fault tolerant algorithms and analogue quantum simulators for Gibbs states. [18].
Most notably, we know that fast thermalization of certain lattice-systems occurs under certain clustering conditions [7], [11], [22] In recent years, an extensive amount of research has focused on using concepts from quantum information theory and developing entropic tools and bounds for the study of thermalization of quantum lattice-spin systems. One of the most notable is the socalled *complete modified logarithmic Sobolev inequality*, an entropic inequality used to establish exponential decay of relative entropy under a Markovian evolution towards its stationary states. [7], [11], [12], [23] This exponential decay directly implies rapid thermalization if the decay rate is at most poly-logarithmically decreasing in system size. Using these tools it was established that under certain conditions, 1D quantum spin chains with Hamiltonians which arise from commuting potentials [12] and hypercubic systems with Hamiltonians arising from nearest neighbour-interacting commuting potentials at high temperatures [11] are rapidly mixing.
Main results in words:(informal)
In this work we will improve amongst others on both of these works and establish new ones for more general lattice systems. In the one dimensional case we will show that the entropy decay rate of the system towards its equilibrium is system size independent (see Theorem 5.2) and in the hypercubic setting we will show that a much weaker condition, namely that of a gapped generator, is actually a sufficient condition for exponential entropic decay and thus rapid thermalization. We will also establish that nearest neighbour commuting systems on b-ary trees satisfy rapid thermalization with a logarithmically decreasing cMLSI constant under the condition of gapped generators. For a graphical representation of some of these implications see Figure 5.1.

1.1 Thesis outline

In the following Chapter 2 will set the notation, present necessary mathematical definitions and prerequisites, and some preliminary results required in the rest of this thesis. Most notably we formally introduce the complete Modified Logarithmic Sobolev Inequality (cMLSI) and show its connection to rapid thermalization. The proof of the main result may conceptually be split into two parts, one static and one dynamic. We establish some preliminary results of potentially independent interest concerning the static part in Chapter 3. There we first define the concept of strong local indistinguishability, which, as the name suggests gives a strictly stronger notion of local indistinguishability and which will be central in the proof of an approximate tensorization- and a 'weak-implies-strong clustering' result in Theorem 5.6 later on. We also show in Theorem 3.4 that for Gibbs states of geometrically local, commuting, and bounded Hamiltonians, weak clustering, i.e. exponential decay of spacial correlations, is actually equivalent to exponential decay of mutual information and implies strong local indistinguishability. This is a in general stronger form of clustering. For 1D systems we show this without the commutativity assumption also in Theorem 3.5.
Next in Chapter 4 we first discuss the local Davies generators associated to a uniform family of Hamiltonians and then construct the Schmidt conditional expectations, establishing some further notation for this work, and derive some of their properties.
In Chapter 5 we first present the main results of this work in Theorem 5.1 and some immediate Corollaries in Theorem 5.2, Theorem 5.3 and discuss them. The rest of that section is then devoted to proving it. One of the main steps of the proof is establishing a very strong clustering condition for nearest-neighbour, commuting systems on 2-colorable graphs, from existence of a strictly positive gap of the generator. This is done in Theorem 5.4 and although this is a static property it is in this

chapter because it requires the notion of Schmidt conditional expectations. This is a result very much of independent interest. Its immediate consequences, such as Theorem 5.5 and Theorem 5.6, are discussed thereafter. Finally through a novel geometric divide-and-conquer strategy, averaging over suitable partitions, we establish the main result in Section 5.2.

In Chapter 6 we derive some applications of the main result, establish exponential convergence to the Gibbs state in the thermodynamic limit, optimal Gaussian concentration bounds, improved entropy difference bounds in the settings considered here. Finally in Chapter 7 we will discuss some conjectures stemming from this work and future related directions of research.

Hence the main results of this work and discussions thereof are found in Chapter 3, Chapter 5, and Chapter 6 whereas Chapter 2 and Chapter 4 introduce necessary prerequisites.

2 Preliminaries

2.1 Graphs

A graph is a tuple $\Lambda = (V, E_V)$ of a set of vertices V and a set of edges $E_V \subset V \times V$ connecting vertices. A complete subgraph $\Gamma \subset \Lambda$ is a tuple (G, E_G), where $G \subset V$ and E_G contains/connects all edges in E_V, which contain the vertices in G. Abusing notation slightly we will call complete subgraphs *subsets*, write $\Gamma \subset \Lambda$.
For simplicity of notation we associate the graph with its vertex set, hence we may write $x \in \Lambda$, $A \subset \Lambda$, or $x \in \Gamma$ for an $x \in V$, when the edge set E_V of $\Lambda = (V, E_V)$ is clear from context.
We define the size of a graph Λ, or of a subset $\Gamma \subset \Lambda$, denotes as $|\Lambda|, |\Gamma|$, respectively, as the number of the vertices it contains. When emphasizing that Γ is a finite subset of Λ, i.e. $|\Gamma| < \infty$, we write $\Gamma \subset\subset \Lambda$.
We write $CD \subset \Lambda$ for the complete subgraph containing all of the vertices of C and D, in this sense $CD = C \cup D$. Note that this does not require C, D to be disjoint. We call a subset of vertices $\Gamma \subset \Lambda$ connected, if for any two vertices $x, y \in \Gamma$ there exists a sequence of pairwise overlapping edges in E_G, s.t. the first overlaps with x and the last with y.
The graph *distance* d (on Λ) between two vertices $x \neq y \in \Lambda$ is defined as the minimal length of a connected subset of edges which overlap both with x and y. We also set $d(x, x) = 0 \forall x \in \Lambda$. The length of a subset of edges is given by the number of edges it contains. The distance between two subsets $A, B \subset \Lambda$ is defined as the minimal graph distance between pairs of points in A and B, respectively. It is denoted, with slight abuse of notation, with the same symbol d. We define the diameter of a set $A \subset \Lambda$ as $\mathrm{diam}(A) := \sup_{x,y \in A} d(x, y)$.
The graph has *growth constant* $\nu > 0$ defined as the smallest real number s.t., for any $m \in \mathbb{N}$, the number of connected subsets of size m containing some edge, for any edge, is bounded by ν^m:

$$n_m := \sup_{e \in E_V} |\{F \subset E_V \text{ connected } | e \in F, |F| = m\}| \leq \nu^m.$$

Note that any *regular graph*, i.e. one where every vertex has the same number of neighbours as every other, has finite growth constant. For example, the growth constant of a D-dimensional hypercubic lattice ($\mathbb{Z}^d$) is bounded by $2De$, where e is Euler's number [6], [24].
We say a graph is *2-colorable* if there exists a labeling of the graph with labels 0 and 1, i.e. a map which assigns each vertex one label, such that adjacent vertices, i.e. ones which are connected by some edge, have different labels.

Definition 1. *For an infinite graph Λ we define $N(l) := \sup_{x \in \Lambda} |B_l(x)|$, where $B_l(x) := \{v \in \Lambda | d(x, v) \leq l|\}$ is the ball of radius l around vertex x. We call a graph sub-exponential if there exists a $\delta \in (0, 1)$ s.t. $N(l) \leq \exp(l^\delta)$ holds eventually, i.e. if $\ln N(l) = O(l^\delta)_{j \to \infty}$.*
Equally we call it exponential if no such δ exists, i.e. if $\ln N(l) = O(l)_{l \to \infty}$.

First note that all graphs with finite growth constant are in either of these two classes since we can crudely bound $|B_l(x)| \leq |\{F \subset \Lambda | F \text{connected} x \in F, |F| = l\}| \leq \nu^l$ and hence $N(l) \leq \nu^l$.
Note that hypercubic lattices are sub-exponential under this definition, whereas $b-$ary trees are

exponential. We will often consider geometrically$-r-$local interactions, with $r > 1$ some integer, on such graphs. For some fixed r we define the boundary of a subset $A \subset \Lambda$, denoted with ∂A, to be all vertices in $\Lambda \setminus A$ that are within graph distance $r - 1$ from vertices in A

$$\begin{aligned} \partial A &:= \{x \in \Lambda \setminus A | d(x, A) < r\}, \\ A\partial &:= A \cup \partial A. \end{aligned}$$

It will be clear from context what r and hence the set-boundary ∂ is. Hence for nearest neighbour interactions ($r = 2$), ∂A coincides with usual set boundary. A first important class of graphs considered in this work is the of hypercubic lattices of dimension $D \in \mathbb{N}$, with $\Lambda = \mathbb{Z}^D$, and the graph distance equal to the Hamming distance. Hence hypercubic lattices are subexponential two colorable graphs. In the case $D = 1$, the 1-dimensional infinite chain this is $\Lambda \equiv \mathbb{Z} = (\mathbb{Z}, \{x, x+1\}_{x\in\mathbb{Z}})$. Another example is the complete infinite $b-$ary tree, for some integer $b \geq 1$. These are loop-free, exponential, and two-colorable graphs. where each vertex has exactly b neighbours. Each tree has one vertex, called the *root*, from which the tree extends, and who's b neighbours are called its children or leaves. Every other vertex has exactly $b - 1$ children or leaves.

2.2 Some general notation

A quantum spin system on a finite graph $\Gamma = (V, E_V)$ is described by the Hilbert space

$$\mathcal{H}_\Gamma := \bigotimes_{x\in V} \mathcal{H}_x,$$

where each local Hilbert space $\mathcal{H}_x$ has dimension $d < \infty$, i.e. describes a qudit system. Hence the global dimension of the system is $\dim(\mathcal{H}_\Gamma) = d^{|\Gamma|}$. We will only be considering finite dimensional Hilbert spaces in this work. We denote the von Neumann-algebra (vN algebra) of bounded linear operators, also called *observables*, over $\mathcal{H}_\Gamma$ by $\mathcal{B}(\mathcal{H}_\Gamma) = \mathscr{A}_\Gamma$ and the set of density operators with $\mathcal{D}(\mathcal{H}_\Gamma) := \{\rho \in \mathcal{B}(\mathcal{H}_\Gamma) | \operatorname{Tr}[\rho] = 1, \rho \geq 0\}$. Note that this von Neumann-algebra is *-homeomorphic to $\mathscr{A}_{\Gamma *}$, the predual of $\mathscr{A}_\Gamma = \mathcal{B}(\mathcal{H}_\Gamma)$ w.r.t. the canonical trace on the finite dimensional Hilbert space $\mathcal{H}_\Gamma$.[1] Since we can associate each (normalized) state (positive, linear functional) with its density operator representation, i.e. for $\omega \in \mathscr{A}_{\Gamma *}$ there exists a $\rho \in \mathcal{D}(\mathcal{H}_\Gamma)$, s.t. $\omega(X) = \operatorname{Tr}[\rho X]$, and the other way around. We denote the trace-class operators on a Hilbert space $\mathcal{H}$ with $B_1(\mathcal{H})$. The norm on $\mathscr{B}(\mathcal{H})$ is the usual operator norm, denoted by $\|A\| \equiv \|A\|_\infty$ for $A \in \mathcal{B}(\mathcal{H})$. The norm on $\mathscr{D}(\mathcal{H})$ is the usual trace-norm, denoted by $\|\rho\|_1 := \operatorname{Tr}[|\rho|]$ for $\rho \in \mathcal{D}(\mathcal{H})$.
We denote the the identity operator on $\mathcal{H}$ as $\mathbb{1} \equiv \mathbb{1}_\mathcal{H} \in \mathcal{B}(\mathcal{H})$ and the identity map $\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ as $\mathrm{id} \equiv \mathrm{id}_{\mathcal{B}(\mathcal{H})}$. Given a linear map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ we denote its pre-dual w.r.t the Hilbert-Schmidt inner product as Φ_*. We call such a map Φ a *quantum channel in the Heisenberg picture* if it is completely positive and unital [2]. We will refer to such maps simply as unital CP maps. Their pre-duals $\Phi_* : B_1(\mathcal{H}) \to B_1(\mathcal{H})$, i.e. *quantum channels in the Schrödinger picture*, are completely positive trace preserving maps (CPTP) [3]. We will call such maps CPTP or quantum channels.
We denote the spectrum of an operator $A \in \mathscr{A}$ with $\sigma(A) \equiv \operatorname{spec}(A)$. The support-projection of a

[1] I.e. the inner product here is the map $\langle X, Y\rangle = \operatorname{Tr}[X^* Y]$, which is also called the *Hilbert Schmidt inner product*.
[2] A map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is completely positive, if $(\mathrm{id}_n \otimes \Phi) : \mathcal{B}(\mathbb{C}^n \otimes \mathcal{H}) \to \mathcal{B}(\mathbb{C}^n \otimes \mathcal{H})$ is a positive map for all $n \in \mathbb{N}$. It is *unital*, if it is identity preserving $\Phi(\mathbb{1}) = \mathbb{1}$.
[3] As the name suggests, a map Φ_* is *trace preserving* if $\operatorname{Tr}[\Phi_*(\rho)] = \operatorname{Tr}[\rho]$ for all $\rho \in \mathcal{D}(\mathcal{H})$.

self-adjoint operator $A \in \mathcal{B}(\mathcal{H})$ is defined as the smallest projection $P \in \mathcal{B}(\mathcal{H})$, s.t. $PA = AP = A$. The trace on the full Hilbert space $\mathcal{H}_\Gamma$ is denoted as $\mathrm{Tr}[\cdot]$ and the partial trace tracing out the Hilbert space corresponding to a region $A \subset \Gamma$ is denoted as $\mathrm{tr}_A[\cdot] : B_1(\mathcal{H}_\Gamma) \to B_1(\mathcal{H}_{\Gamma\setminus A})$.
We will employ the following *"big-O"*-notation $O(g(x))_{x\to\infty}$ when meaning that $f(x) = O(g(x))$ for $x \to \infty$, i.e. to indicate in which limit the scaling $O(g(x))$ holds for a function $f(x)$. We use the same for the $big - \Omega$ notation.

2.3 Weighted non-commutative $\mathbb{L}_{p,\sigma}$-spaces and inner products

We will make frequent use of so called (weighted) non-commutative $\mathbb{L}_p$ spaces in this work. For a general overview and construction of such spaces on von Neumann algebras, see e.g. [25]. Since we are only considering finite dimensional Hilbert spaces, the von Neumann algebra $\mathcal{B}(\mathcal{H})$ with the canonical Hilbert space trace Tr on $\mathcal{H}$ is finite, i.e. of type I.[4] Hence all the non-commutative $\mathbb{L}_p \equiv \mathbb{L}_p(\mathcal{B}(\mathcal{H}), \mathrm{Tr})$ spaces are just the $p-$Schatten spaces $B_p(\mathcal{H}) := \{X \in \mathcal{B}(\mathcal{H}) | \infty > \|X\|_p := \mathrm{Tr}[|X|^p]^{\frac{1}{p}})\}$ with norms

$$\|X\|_p := \mathrm{Tr}[|X|^p]^{\frac{1}{p}} \qquad 1 \le p < \infty, \tag{2.1}$$
$$\|X\|_\infty = \|X\|, \tag{2.2}$$

which are all equivalent to each other. This is due to the finite dimension of the Hilbert space. Hence the L_p-norms $\|\cdot\|_p$ are all equivalent to each other. This is not true in general.
Given a full-rank state $\sigma \in \mathcal{D}(\mathcal{H})$, we define the *weighted non-commutative* $\mathbb{L}_{p,\sigma}$ spaces as the subsets of $\mathcal{B}(\mathcal{H})$ s.t. its elements are bounded by the following associated norms

$$\|X\|_{p,\sigma} := \mathrm{Tr}[|\sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}}|^p]^{\frac{1}{p}} \qquad 1 \le p < \infty, \tag{2.3}$$
$$\|X\|_{\infty,\sigma} := \|X\|_\infty \equiv \|X\|, \tag{2.4}$$

respectively. Note that these norms turn these spaces into Banach spaces for $p \in [1, \infty]$ and satisfy the usual Hölder-type inequality, Hölder duality, and monotonicity in p for fixed σ, see e.g. [3]. Equally one can show that $\mathbb{L}_{2,\sigma}$ is, as expected, a Hilbert space with respect to the *KMS-inner product*

$$\langle X, Y\rangle_\sigma^{\mathrm{KMS}} := \mathrm{Tr}[\sqrt{\sigma} X^* \sqrt{\sigma} Y]. \tag{2.5}$$

There exists a natural embedding $\Gamma_\sigma : \mathbb{L}_{1,\sigma} \to \mathbb{L}_1$ via

$$\Gamma_\sigma(X) := \sqrt{\sigma} X \sqrt{\sigma}. \tag{2.6}$$

Hence the weighted p, σ-norm can also be expressed as $\|X\|_{p,\sigma} = \|\Gamma_\sigma^{\frac{1}{p}}(X)\|_p$. For completeness we define the *modular operator* of σ here as

$$\Delta_\sigma(X) := \sigma X \sigma^{-1}, \tag{2.7}$$

and the *modular group* of σ as $\{\Delta_\sigma^{is}\}_{s\in\mathbb{R}}$. For completeness, we define the *GNS-inner product* on $\mathcal{B}(\mathcal{H})$, for a finite dimensional Hilbert space $\mathcal{H}$ as

$$\langle X, Y\rangle_\sigma^{\mathrm{GNS}} := \mathrm{Tr}[\sigma X^* Y]. \tag{2.8}$$

[4] A von Neumann algebra is said to be finite if a tracial state exists. In the case here, it is $d^{-1}\mathbb{1} \in \mathcal{D}(\mathcal{H})$, whenever $d = \dim(\mathcal{H}) < \infty$.

These two weighted inner products on $\mathcal{B}(\mathcal{H})$ are the most relevant ones and we give the variance and covariance of our many-body observables with respect to these as

$$\mathrm{Cov}_\sigma(X,Y) := \langle X - \mathrm{Tr}[\sigma X]\mathbb{1}, Y - \mathrm{Tr}[\sigma Y]\mathbb{1}\rangle_\sigma^{\mathrm{KMS}} = |\,\mathrm{Tr}[\sqrt{\sigma}X\sqrt{\sigma}Y] - \mathrm{Tr}[\sigma X]\,\mathrm{Tr}[\sigma Y]|, \tag{2.9}$$

$$\mathrm{Cov}_\sigma^{(0)}(X,Y) := \langle X - \mathrm{Tr}[\sigma X]\mathbb{1}, Y - \mathrm{Tr}[\sigma Y]\mathbb{1}\rangle_\sigma^{\mathrm{GNS}} = |\,\mathrm{Tr}[\sigma XY] - \mathrm{Tr}[\sigma X]\,\mathrm{Tr}[\sigma Y]|, \tag{2.10}$$

$$\mathrm{Var}_\sigma(X) := \mathrm{Cov}_\sigma(X,X), \tag{2.11}$$

$$\mathrm{Var}_\sigma^{(0)}(X) := \mathrm{Cov}_\sigma^{(0)}(X,X). \tag{2.12}$$

A distance on $\mathcal{D}(\mathcal{H}_\Lambda)$, which will be important for some applications is the *quantum Wasserstein distance of order 1* [26] between two finite dimensional quantum states $\rho,\sigma \in \mathcal{D}(\mathcal{H}_\Lambda)$. It is defined as

$$W_1(\rho,\sigma) \equiv \|\rho-\sigma\|_{W_1} := \tag{2.13}$$
$$\frac{1}{2}\min\left\{\sum_{i\in\Lambda}\|X^{(i)}\|_1 \,\middle|\, \mathrm{Tr}[X^{(i)}]=0, X^{(i)*}=X^{(i)}, \mathrm{tr}_i X^{(i)}=0\ \forall i\in\Lambda, \rho-\sigma=\sum_{i\in\Lambda}X^{(i)}\right\}.$$

Its dual norm w.r.t the Hilbert-Schmidt inner product distance is the *Lipschitz distance* [26], i.e. for any self-adjoint observable $A \in \mathcal{B}(\mathcal{H}_\Lambda)$

$$\|A\|_L := \max\{\mathrm{Tr}[AX]\,|\,\mathrm{Tr}[X]=0, X=X^*, \|X\|_{W_1}\le 1\} = 2\max_{i\in\Lambda}\min_{A^{(i)}\in\mathcal{A}_{\Lambda\setminus\{i\}}}\|A-\mathbb{1}_i\otimes A^{(i)}\|,$$

where $\mathbb{1}_i \in \mathcal{B}(\mathcal{H}_i)$ is the identity on system i and $A^{(i)}$ does not act on system i. Thus by definition it holds that $|\,\mathrm{Tr}[AX]| \le \|X\|_{W_1}\|A\|_L$ for suitable X, A. For a thorough overview and some properties see [26].

2.4 Uniform families of Hamiltonians

A *Hamiltonian* $H : \mathcal{B}(\mathcal{H}_\Lambda)$ is a self-adjoint operator which governs the dynamics of closed quantum systems and describes the state of a quantum system in thermal equilibrium at the same time. In this work we consider many-body Hamiltonians of the form,

$$H_\Lambda := \sum_{X\subset\subset\Lambda}\Phi_X,$$

where for each $X\subset\Lambda$, Φ_X is a self-adjoint operator acting only non-trivially on the sub-region X. The map $X\mapsto\Phi_X$ for a finite $X\subset\Lambda$ is called the *potential* of the system. The potential is called *commuting* (on Λ) if for each $X,Y\subset\Lambda$ Φ_X and Φ_Y commute. It is said to have *bounded interaction strength* $J := \max_{X\subset\Lambda}\{\|\Phi_X\|\}$ and *interaction range* $r := \max\{\mathrm{diam}(X)|X\subset\Lambda, \Phi_X\neq 0\}$. We will call potentials with interaction range r *geometrically-r-local*. The family of Hamiltonians $\{H_\Gamma\}_{\Gamma\subset\subset\Lambda}$, s.t.

$$H_\Gamma = \sum_{X\subset\Gamma}\Phi_X, \tag{2.14}$$

is called a *uniform* J-bounded, geometrically-r-local, commuting family if the potential satisfies these properties for all $\Gamma\subset\subset\Lambda$ independent of $|\Gamma|$. In this work, we will only consider such

uniform families, unless explicitly stated otherwise.[5]
The associated Gibbs state of the local Hamiltonian on $A \subset \Gamma$ at inverse temperature β is denoted by

$$\sigma^A := \frac{e^{-\beta H_A}}{\mathrm{Tr}[e^{-\beta H_A}]}, \tag{2.15}$$

while the reduced states onto some subregion $A \subset \Gamma$ is denoted by

$$\sigma_A := \mathrm{tr}_{\Gamma \backslash A}\, \sigma^\Gamma, \tag{2.16}$$

where $\sigma \equiv \sigma^\Gamma = \sigma_\Gamma$. We will employ the convenient notation $E_{X,Y} := e^{-H_{XY}} e^{H_X + H_Y}$ for Araki's expansionals for two disjoint subsets $X, Y \subset \Lambda$ from [7].

2.5 Quantum Markov semigroups and Lindbladians

A *quantum Markov semigroup* (QMS) is a strongly continuous one-parameter semigroup of unital CP maps $\{\Phi_t\}_{t\geq 0} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$. This is a family s.t. $\Phi_0 = \mathrm{id}_{\mathcal{B}(\mathcal{H})}, \Phi_{s+t} = \Phi_s \circ \Phi_t \; \forall s, t \geq 0$, and $\lim_{t\downarrow 0} \|(\Phi_t - \mathrm{id})(X)\| = 0 \forall X \in \mathcal{B}(\mathcal{H})$. By Hille-Yosida theorem there exists a densely defined generator, called the *Lindbladian*

$$\mathcal{L}(X) := \lim_{t\to 0} \frac{1}{t}(\Phi_t - \mathrm{id})(X),$$

such that the semigroup is given as $\Phi_t = e^{t\mathcal{L}} \; \forall t \geq 0$. In our case of a finite dimensional Hilbert space, the Lindbaldian is defined on all of $\mathcal{B}(\mathcal{H})$ and its pre-dual on all of $\mathcal{D}(\mathcal{H})$.
A QMS with generator $\mathcal{L}$ gives the unique solution to the master equation $\frac{d}{dt}\rho(t) = \mathcal{L}(\rho)$. We call a QMS and its generator *faithful* if the QMS admits a full rank invariant state $\sigma \in \mathcal{D}(\mathcal{H})$ and *primitive* if this state is unique. A state is invariant if $\Phi_{t*}(\sigma) = \sigma$ for all $t \geq 0$, which is equivalent to $\mathcal{L}_*(\sigma) = 0$.

We call a QMS and its generator *reversible* or *KMS-symmetric* w.r.t. a state σ if the QMS is symmetric w.r.t the KMS-inner product and similarly *GNS-symmetric* w.r.t. a state σ, if it is symmetric w.r.t. the GNS inner product. In the latter case, we also say that the QMS satisfies the *detailed balance condition*. If $\mathcal{L}$ is its generator, then this is equivalent to

$$\mathrm{Tr}[\sigma X^* \mathcal{L}(Y)] = \mathrm{Tr}[\sigma \mathcal{L}(X)^* Y] \qquad \forall X, Y \in \mathcal{B}(\mathcal{H}).$$

This is because one can think of the GNS symmetry of a QMS to be a quantum generalization of the detailed balance property of a classical Markovian process w.r.t its invariant distribution. Note that if a QMS is GNS-symmetric w.r.t a state σ, then this state is necessarily a stationary one. Given a graph Λ and a finite subset $\Gamma \subset\subset \Lambda$, we will be considering a family of Lindbladians $\mathcal{L}_\Lambda = \{\mathcal{L}_\Gamma\}_{\Gamma\subset\subset\Lambda}$, s.t.

$$\mathcal{L}_\Gamma = \sum_{X\subset\Gamma} L_X, \tag{2.17}$$

where $\{\mathcal{L}_X\}_{X\subset\Lambda}$ is a set family of local Lindbladians, s.t. $\tilde{J} := \sup_{X\subset\Lambda} \|\mathcal{L}_{X*}\|_{1\to 1,\mathrm{cb}} < \infty$ and $\mathcal{L}_{X*} = 0$ whenever $\mathrm{diam}(X) > r$.[6] Hence we call these a *uniform geometrically−r−local,*

[5] Hence these constants J, r do not depend on the regions A on which the local Hamiltonians are defined.

[6] Here $\|\Phi_*\|_{1\to 1,\mathrm{cb}}$ is the completely bounded $1 \to 1$ norm, i.e. $\|\Phi_*\|_{1\to 1,\mathrm{cb}} := \sup_{n\in\mathbb{N}} \sup_{\rho\in\mathcal{D}(\mathbb{C}^n\otimes\mathcal{H})} \|(\mathrm{id}_n \otimes \Phi_*)(\rho)\|_1$.

$\tilde{J}-$*bounded* family of (bulk) Lindbladians, very much in analogy to the Hamiltonian case. When we call a family of Lindbadians *uniform*, we imply that there are some $\tilde{J}, r < \infty$ such that it is geometrically$-r-$local and $\tilde{J}-$bounded in the sense above. Such a family is called *locally primitive* if there exists a full rank state for each finite $\Gamma \subset\subset \Lambda$, $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$, s.t. σ^Γ is the unique full rank stationary state of $\mathcal{L}_\Gamma$. The family is called *locally reversible* if each $\mathcal{L}_\Gamma$ is KMS-symmetric w.r.t σ^Γ and it is called *frustration free*, if for any two finite subsets $A \subset B \subset \Lambda$, the stationary states of $\mathcal{L}_B$ are also stationary under $\mathcal{L}_A$, i.e. the $\ker(\mathcal{L}_B)\subset\ker(\mathcal{L}_A)$. Note, that if a QMS is GNS symmetric,i.e. satisfies detailed balance, then it is also KMS symmetric, i.e. reversible. Hence, for a region $\Gamma \subset\subset \Lambda$, write the projection onto the fixed point subalgebra of $\mathcal{L}_\Gamma$ as

$$E_\Gamma(X) := \lim_{t\to\infty} e^{t\mathcal{L}_\Gamma}(X). \tag{2.18}$$

It turns out that for a primitive, frustration free uniform family, these projections are conditional expectations w.r.t the family of stationary states. See Section 2.6 for more details. For a slightly more general, but in this work unnecessary, notion of uniform families of Lindbladians, see e.g. [11].

In this work we will be working with the *Davies generators* $\mathcal{L}^D_\Lambda = \{\mathcal{L}^D_\Gamma\}_{\Gamma\subset\subset\Lambda}$, which is a physically motivated suitable uniform family of Lindbladians associated to a uniform family of Hamiltonians. They are introduced in Section 4.1. In the setting we are considering, they are a uniform geometrically$-r-$bounded, locally primitive, locally reversible, locally GNS-symmetric,[7] frustration free family of Lindbladians which describe thermalization of a spin system.

We call a uniform family of Lindbladians, which are locally reversible, locally GNS symmetric, and frustration free w.r.t to a set of Gibbs states $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$ a *quantum Gibbs sampler* of the system (H_Λ, β), whenever H_Λ is a uniform family of Hamiltonians and σ^Γ are the Gibbs states of H_Γ to inverse temperature β. The Davies generators, see Section 4.1, but also the Heat-bath generators w.r.t to the Davies or Schmidt conditional expectation [3], [11] are examples of quantum Gibbs samplers.

2.6 Conditional Expectations

A very important tool we are working with are (quantum) conditional expectations. Given a von Neumann subalgebra $\mathcal{N} \subset \mathcal{B}(\mathcal{H})$, a *conditional expectation* onto $\mathcal{N}$ is a completely positive unital map $E_\mathcal{N} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$, s.t.

$$\begin{aligned} &E_\mathcal{N}(X) = X \;\forall\, X \in \mathcal{N} \\ &E_\mathcal{N}(aXb) = aE_\mathcal{N}(X)b \;\forall\, a, b \in \mathcal{N}, X \in \mathcal{B}(\mathcal{H}). \end{aligned}$$

By complete positivity and unitality, it follows that the preadjoint of any conditional expectation w.r.t the Hilbert-Schmidt inner product $E_{\mathcal{N}*} : \mathcal{N}_* \to \mathcal{B}(\mathcal{H})_*$ is a completely positive trace preserving map, i.e. a quantum channel. Any *conditional expectation* onto $\mathcal{N}$ s.t. there exists a full rank state $\sigma \in \mathcal{D}(\mathcal{H})$ which satisfies

$$E_{\mathcal{N}*}(\sigma) = \sigma \iff \mathrm{Tr}[\sigma E_\mathcal{N}(X)] = \mathrm{Tr}[\sigma X] \;\forall\, X \in \mathcal{B}(\mathcal{H}),$$

is said to be *with respect to the state* σ. [11], [12] Let E be a conditional expectation w.r.t. a full rank state σ onto $\mathcal{N}$, then from the definition it follows that it is self-adjoint w.r.t. the $\sigma-$KMS

[7]The KMS and GNS symmetry is w.r.t to the global Gibbs state.

inner product, i.e.

$$\sigma^{\frac{1}{2}}E(X)\sigma^{\frac{1}{2}} = E_*(\sigma^{\frac{1}{2}}X\sigma^{\frac{1}{2}})$$

holds for any $X \in \mathcal{B}(\mathcal{H})$.
Furthermore, it can be shown that it commutes with the modular group of σ, i.e.

$$\Delta_\sigma^{is} \circ E = E \circ \Delta_\sigma^{is} \; \forall s \in \mathbb{R}.$$

Moreover, given a von Neumann *-subalgebra $\mathcal{N} \subset \mathcal{B}(\mathcal{H})$ and a faithfull state $\sigma \in \mathcal{D}(\mathcal{H})$, it turns out that the existence of a conditional expectation w.r.t. σ onto $\mathcal{N}$ is equivalent to the invariance of $\mathcal{N}$ under the modular automorphism group $\{\Delta_\sigma^{is}\}_{s\in\mathbb{R}}$. Furthermore, in the case that the vN *-subalgebra $\mathcal{N}$ is invariant under the modular automorphism group of said faithfull state σ this conditional expectation is uniquely determined by σ [11], [27]. It turns out that any conditional expectation w.r.t. some full rank state σ between finite dimensional matrix algebras, as all the ones in this work, can be given in an explicit form, see e.g. [12]. Any finite dimensional von Neumann subalgebra $\mathcal{N} \subset \mathcal{B}(\mathcal{H})$ can be decomposed as

$$\mathcal{N} = \bigoplus_{i=1}^{n} \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{C}\mathbb{1}_{\mathcal{K}_i}, \text{ where } \mathcal{H} = \bigoplus_{i=1}^{n} \mathcal{H}_i \otimes \mathcal{K}_i.$$

Now there exist density operators $\{\tau_i \in \mathcal{D}(\mathcal{K}_i)\}_{i=1}^n$ and projections $\{P_i \in \mathcal{B}(\mathcal{H})\}_{i=1}^n$, respectively, onto $\{\mathcal{H}_i \otimes \mathcal{K}_i\}$ s.t.

$$E_\mathcal{N}(X) = \bigoplus_{i=1}^{n} \operatorname{tr}_{\mathcal{K}_i}[P_i X P_i(\mathbb{1}_{\mathcal{K}_i} \otimes \tau_i)] \otimes \mathbb{1}_{\mathcal{K}_i} \iff E_{\mathcal{N}*}(\rho) = \bigoplus_{i=1}^{n} \operatorname{tr}_{\mathcal{K}_i}[P_i \rho P_i] \otimes \tau_i,$$

for $X \in \mathcal{B}(\mathcal{H})$ and $\rho \in \mathcal{B}(\mathcal{H})$. [12]

Since conditional expectations are, by definition, projections on closed-*-subalgebras (which are convex), the following chain rule holds for states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, whenever $E_{\mathcal{N}*}(\sigma) = \sigma$ [23][Lemma 3.4]

$$D(\rho\|\sigma) = D(\rho\|E_{\mathcal{N}*}(\rho)) + D(E_{\mathcal{N}*}(\rho)\|\sigma). \tag{2.19}$$

Example 1 (Local Lindbaldian Projectors)**.** An important case of conditional expectations are so called *local Lindbladian projectors*.
Let Γ be some finite graph. Let $\mathcal{L}_\Gamma = \{\mathcal{L}_A\}_{A\subset\subset\Lambda}$ be a uniform, frustration free family of local primitive Lindbladians with stationary states $\{\sigma^A\}_{A\subset\Gamma}$. The local Lindbladian projector associated with the family $\mathcal{L}_\Gamma$ on $A \subset \Gamma$ is given by

$$E_A(X) := \lim_{t\to\infty} e^{t\mathcal{L}_A}(X) \tag{2.20}$$

for $X \in \mathcal{B}(\mathcal{H}_\Gamma)$. If $\mathcal{L}_\Gamma$ is locally primitive, i.e. each $\mathcal{L}_A$ is a primitive Lindbladian, then E_A acts only non-trivially on A^c. If $\mathcal{L}_\Gamma$ is frustration free, then E_A is a conditional expectation w.r.t. the stationary state σ^Γ onto the subalgebra $\mathbb{1}_{A\partial} \otimes \mathcal{B}(\mathcal{H}_{(A\partial)^c})$.

Proof. Complete positivity and unitality follows from that of $e^{t\mathcal{L}_\Gamma}$ for any $t \geq 0$. Taking the limit $t \to \infty$ does not change these properties. Clearly we have $e^{t\mathcal{L}_{\Gamma*}}(\sigma^\Gamma) = \sigma^\Gamma$, and by frustration freeness it follows that $e^{t\mathcal{L}_{A*}}(\sigma^\Gamma) = \sigma^\Gamma$. Hence

$$\operatorname{Tr}[\sigma E_A(X)] = \lim_{t\to\infty} \operatorname{Tr}[\sigma e^{t\mathcal{L}_A}(X)] = \lim_{t\to\infty} \operatorname{Tr}[e^{t\mathcal{L}_{A*}}(\sigma)X] = \operatorname{Tr}[\sigma X].$$

Furthermore, by locality we have that E_A acts only non trivially on $\mathcal{B}(\mathcal{H}_{A\partial})$ since $\mathcal{L}_A$ does. □

Note that in this case it holds that the expectation value of any observable w.r.t. the invariant state on the full system is also given by

$$\mathrm{Tr}[\sigma X] = \mathrm{Tr}[\sigma E_\Lambda(X)] = \mathrm{Tr}[\sigma E_A(X)] \ \forall A \subset \Lambda.$$

On the other hand, given a family of local conditional expectation $E_A : \mathcal{B}(\mathcal{H}) \to \mathbb{1}_{A\partial} \otimes \mathcal{B}(\mathcal{H})_{(A\partial)^c}$ w.r.t. the same state $\sigma \in \mathcal{D}(\mathcal{H})$,

$$\overline{\mathcal{L}}_A := \sum_{X \cap A \neq \emptyset} (E_X - \mathrm{id}), \tag{2.21}$$

is a family of locally primitive, frustration free Lindbladians with invariant state σ.

Since the *Davies-Lindbladian* are a family of locally primitive...

2.7 The relative entropy and strong data processing

The *Umegaki relative entropy* [28] between two finite dimensional quantum states given by their density operators $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is defined as

$$D(\rho\|\sigma) := \begin{cases} \mathrm{Tr}[\rho(\log\rho - \log\sigma)] & \text{if } \mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma) \\ \infty & \text{else} \end{cases},$$

where the logarithm here is the natural logarithm to base e.
It is a very important information theoretic quantity which can be interpreted as a statistical distinguishability quantity between states, e.g. as Stein exponent in asymptotic asymmetric quantum hypothesis testing [29]. The well known quantum Pinsker inequality (2.22), gives an upper bound on the trace-distance, which is related to the one-shot symmetric distinguishability, in terms of the relative entropy:

$$\|\rho - \sigma\|_1^2 \leq 2D(\rho\|\sigma) \tag{2.22}$$

Hence, the relative entropy is positive semi-definite, however, unlike a proper mathematical distance, it is in general not symmetric in its two arguments, nor does it satisfy the triangle inequality. It also gives rise to the *quantum mutual Information* I. Given a finite graph $\Gamma = ABC$ the mutual information of a state $\rho \in \mathcal{D}(\mathcal{H}_\Gamma)$ between the reduced state on the region A and the one on the region C is defined as

$$I_\rho(A : C) := D(\rho_{AC}\|\rho_A \otimes \rho_C). \tag{2.23}$$

It is a measure of mutual information between these two regions.
The operational interpretation of the relative entropy as an information theoretic measure is further underlined by a very important property it satisfies, called the *data processing inequality* (DPI) (2.24). This property is that no quantum channel, i.e. (CPTP) map Φ_*, can increase the relative entropy between any two states.

$$D(\Phi_*(\rho)\|\Phi_*(\sigma)) \leq D(\rho\|\sigma). \tag{2.24}$$

The core part of the the entropic-inequalities approach to thermalization relies upon a strengthening of this inequality. We say a quantum channel Φ_* satisfies a non-trivial *strong data processing* (sDPI) with *contraction coefficient* $\eta \equiv \eta(\Phi_*) < 1$ if for any pair (ρ, σ) of states, s.t. $\rho \neq \sigma$, it holds that

$$D(\Phi_*(\rho)\|\Phi_*(\sigma)) \leq \eta(\Phi_*) D(\rho\|\sigma). \tag{2.25}$$

More formally, we define the contraction coefficient for a GNS symmetric QMS $\Phi_* \equiv \Phi_{t_0 *}$ as

$$\eta(\Phi_*) := \sup_{\rho \in \mathcal{D}(\mathcal{H})} \frac{\inf_{\sigma \in \Sigma} D(\Phi_*(\rho) \| \Phi_*(\sigma))}{\inf_{\sigma \in \Sigma} D(\rho \| \sigma)},$$

where Σ is the set of *stationary states* of Φ_*. These are all the density operators which are left invariant under the action of the channel.[8] Assume we have some channel Φ_* which has a contraction coefficient $\eta(\Phi_*) < 1$ and a unique invariant state σ, i.e. $\Phi_*(\sigma) = \sigma$. Then sDPI immediately induces an exponential decay of the relative entropy in the number of times the channel is applied.

$$D(\Phi_*^n(\rho) \| \sigma) = D(\Phi_*^n(\rho) \| \Phi_*^n(\sigma)) \leq \eta^n D(\rho \| \sigma).$$

2.8 Relative entropy decay via the complete modified logarithmic Sobolev Inequality

We can establish the sDPI, not only for time-discrete, but also time-continous QMS $\{e^{t\mathcal{L}}\}_{t \geq 0}$. Here we assume that our QMS has at least one full rank invariant state, say σ, w.r.t. which it is GNS-symmetric. This turns out to be the case for all Davies Lindbladians we will be considering in this work. The way to do this is via a differential version of the strong data processing inequality (2.25) for the channel $\Phi_{t*} := e^{t\mathcal{L}_*}$ in which we set $\eta(e^{t\mathcal{L}_*}) = e^{-t\alpha}$, yielding

$$-\frac{d}{dt} D(e^{t\mathcal{L}_*}(\rho) \| E_*(\rho))\Big|_{t=0} =: \mathrm{EP}_{\mathcal{L}}(\rho) \geq \alpha D(\rho \| E_*(\rho)). \tag{2.26}$$

Here $E_* := \lim_{t \to \infty} e^{t\mathcal{L}_*}$ is the projection onto the stationary states (2.18), see also Example 1. We call this inequality (2.26) the *modified logarithmic Sobolev inequality* (MLSI) and $\mathrm{EP}_{\mathcal{L}}(\rho)$ the *entropy production* of the QMS $\{e^{t\mathcal{L}}\}_{t \geq 0}$. The optimal constant α satisfying the MLSI is called the *modified logarithmic Sobolev constant* (MLSI constant) $\alpha(\mathcal{L})$. It is hence given by

$$\alpha(\mathcal{L}) := \inf_{\rho \in \mathcal{D}(\mathcal{H})} \frac{\mathrm{EP}_{\mathcal{L}}(\rho)}{D(\rho \| E_*(\rho))}.$$

Thus essentially by construction, formally by integration and use of Gronwall's inequality, it follows that any QMS $\{e^{t\mathcal{L}}\}_{t \geq 0}$ which satisfies the MLSI with strictly positive MLSI constant $\alpha \equiv \alpha(\mathcal{L}) > 0$ induces exponential convergence in relative entropy to its stationary states, i.e.

$$D(e^{t\mathcal{L}_*}(\rho) \| E_*(\rho)) \leq e^{-\alpha t} D(\rho \| E_*(\rho)).$$

One important way to establish the existence of such constants in the classical setting is to exploit its stability under tensorization. This allows us to describe the dynamics of large composite systems via their dynamics on small subregions. This is, however, not in general given in the quantum setting, i.e. if we have two QMS $\{e^{t\mathcal{L}}\}_{t \geq 0}, \{e^{t\mathcal{K}}\}_{t \geq 0}$, then the joint evolution, given by $\{e^{t\mathcal{L}} \otimes e^{t\mathcal{K}} = e^{t(\mathcal{L}+\mathcal{K})}\}$ is not necessarily as quickly mixing as the slower individual one, i.e. $\alpha(\mathcal{L} + \mathcal{K}) \ngeq \min\{\alpha(\mathcal{L}), \alpha(\mathcal{K})\}$. [30] In order to recover the stability under tensorization we can introduce the so called *complete MLSI* (cMLSI) and the *cMLSI constant* $\alpha_c(\mathcal{L})$

$$\alpha_c(\mathcal{L}) := \inf_{k \in \mathbb{N}} \alpha(\mathcal{L} \otimes \mathrm{id}_k), \tag{2.27}$$

[8] In the case of a general quantum channel (instead of QMS) we would need to replace Σ by the decoherence free subalgebra. If $\Phi_*(X) = \sum_k A_k X A_k^*$ is the Kraus representation of Φ_*, then the decoherence free subalgebra is $\Sigma := \bigcap_{k \in \mathbb{N}} \mathcal{N}(\Phi_*^k)$, where $\mathcal{N}(\Phi_*) := \mathrm{Alg}\{X \in \mathcal{B}(\mathcal{H}) | [X, A_i^* A_j] = 0 \ \forall i, j\}$

where $\mathrm{id}_n : \mathcal{B}(\mathbb{C}^n) \to \mathcal{B}(\mathbb{C}^n)$ is the identity channel [31]. Hence we say that the QMS $\{e^{t\mathcal{L}}\}_{t\geq 0}$ satisfies the cMLSI if the QMS $\{e^{t\mathcal{L}} \otimes \mathrm{id}_n\}_{t\geq 0}$ satisfies the MLSI for all ancilla system of arbitrary dimension with the same constant. Indeed in [30] it was shown that for two KMS-symmetric QMS with commuting generators $\mathcal{L}, \mathcal{K}$, respectively, it holds that

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

Next the following important result from [8], [32] guarantees the existence of positive cMLSI constants for a sufficiently large class of QMS.

Theorem 2.1 ([8])**.** For any GNS-symmetric QMS $\{e^{t\mathcal{L}}\}_{t\geq 0}$ over the algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators over some finite dimensional Hilbert space $\mathcal{H}$, $\alpha_c(\mathcal{L}) > 0$. In particular $\alpha_c(\mathcal{L}) \geq \frac{\lambda(\mathcal{L})}{\log \dim \mathcal{H}}$, i.e. for many-body quantum lattice systems the cMLSI constant is deceasing as $\Omega(|\Lambda|^{-1})$.

Here $\lambda(\mathcal{L})$ is the spectral gap of the generator. Local existence of a strictly positive cMLSI constant is a great starting point, however, on its own not very helpful for systems in the thermodynamic limit, since it does not give good bounds on the mixing time, as is discussed below. A common way this result is used, when showing existence of a cMLSI constant which scales better than $O(|\Lambda|^{-1})$, is to use approximate tensorization results to geometrically break down the lattice into logarithmic or finite size parts and apply Theorem 2.1 to these regions and then put these small regions back together into the whole lattice. Such an approach is sometimes called a divide-and-conquer technique, or global-to-local reduction. This is also the overall strategy we are following in this work.

2.9 Variance Decay and Gap

A non information theoretic inspired and simpler approach to mixing of QMS $\{e^{t\mathcal{L}}\}$ is via the *Poincaré inequality* instead of the MLSI. Assume for simplicity, that the QMS is primitive with fixed point σ and GNS symmetric, and write $X_t := e^{t\mathcal{L}}(X)$. Then the *Poincaré inequality* is

$$\lambda \operatorname{Var}_\sigma(X_t) \leq -\frac{d}{dt}\bigg|_{t=0} \operatorname{Var}_\sigma(X) = -\langle X, \mathcal{L}(X)\rangle_\sigma^{\mathrm{KMS}},$$

where $\operatorname{Var}_\sigma(X)$ is the KMS variance defined in Section 2.3. It turns out that the largest constant λ which satisfies this inequality for all $X \in \mathcal{B}(\mathcal{H})$ is the spectral gap of the Lindbaldian $\mathcal{L}$, that is the absolute value of the greatest non-zero eigenvalue of $\mathcal{L}$[9] [3], [33], [34].

$$\lambda(\mathcal{L}) := \inf_{X\in\mathcal{B}(\mathcal{H})} \frac{-\langle X, \mathcal{L}(X)\rangle_\sigma^{\mathrm{KMS}}}{\operatorname{Var}_\sigma(X)} \tag{2.28}$$

By Grönwall's inequality, this directly implies exponential decay of the variance, i.e.

$$\operatorname{Var}_\sigma(X_t) \leq e^{-\lambda(\mathcal{L})t} \operatorname{Var}_\sigma(X). \tag{2.29}$$

2.10 Rapid Thermalization

A natural figure of merit to quantitatively describe mixing of QMS $\{e^{t\mathcal{L}}\}_{t\geq 0}$ and hence thermalization is the so-called *mixing* or *return time* of the QMS. Write $\rho_t := e^{t\mathcal{L}_*}(\rho)$. Then for $\epsilon > 0$,

$$t_{\mathrm{mix}}(\epsilon) := \inf\{t \geq 0 | \forall \rho \in \mathcal{D}(\mathcal{H}) \|\rho_t - E_*(\rho)\|_1 \leq \epsilon\}. \tag{2.30}$$

[9] Due to primitivity the Lindbadian has one eigenvalue 0, and all others strictly smaller.

It tells us how long we have to let the system evolve, s.t. under any initial state ρ_0, the time-evolved state ρ_t is ϵ close in trace-distance to the stationary state $E_*(\rho)$. For a primitive quantum Gibbs sampler, like the Davies evolution,the mixing time tells us how quickly the system approaches the thermal equilibrium state σ^Λ.
We are interested in how this mixing time scales with system size $|\Lambda|$. In the previous section it was shown that the spectral gap $\lambda \equiv \lambda(\mathcal{L}_\Lambda^D)$ induces exponential decay of the KMS-variance. This directly implies

$$\|\rho_t - \sigma\|_1 \leq \|\sigma^{-1}\|^{-\frac{1}{2}} e^{-\lambda t} \implies t_{\text{mix}}(\epsilon) = \frac{1}{\lambda} \log(\epsilon^{-1} \|\sigma^{-\frac{1}{2}}\|) = \frac{1}{\lambda} O\left(\ln \frac{1}{\epsilon} + |\Lambda|\right), \tag{2.31}$$

where in the implication we assumed σ was the thermal state of some bounded geometrically-local Hamiltonian, hence $\|\sigma^{-1}\|^{-1} = e^{O(|\Lambda|)}$ is the smallest eigenvalue of the Gibbs state of a uniform family of Hamiltonians [34]. Thus proving thermalization via gap, which may still depend on the system size, gives a mixing time which scales in the best case as $O(|\Lambda|)$.
We can do better though, in fact if the primitive QMS satisfies the cMLSI with constant $\alpha \equiv \alpha_c(\mathcal{L}) > 0$, then via Pinsker's inequality we get

$$\begin{aligned} &\|\rho_t - \sigma\|_1 \leq \sqrt{2D(\rho_t\|\sigma)} \leq e^{-\frac{t\alpha}{2}} \sqrt{2D(\rho\|\sigma)} \leq e^{-\frac{t\alpha}{2}} \sqrt{2 \log \|\sigma^{-1}\|} \\ \implies &t_{\text{mix}}(\epsilon) \leq \frac{1}{\alpha} O\left(\ln \frac{1}{\epsilon} + \ln |\Lambda|\right). \end{aligned} \tag{2.32}$$

Hence if the cMLSI constant of a QMS is system size independent, then cMLSI implies a scaling of the mixing time with $O(\ln |\Lambda|)$. We call such a logarithmic or poly-logarithmic scaling of the mixing time in system size *rapid mixing*. Hence a primitive uniform family of Lindbladians $\mathcal{L} = \{\mathcal{L}_\Gamma\}_{\Gamma \subset\subset \Lambda}$ satisfies rapid mixing if $0 < \alpha(\mathcal{L}_\Gamma) = \Omega\left((\text{polylog}(|\Gamma|))^{-1}\right)_{|\Gamma| \to \infty}$ and obviously especially if it is constant in system size.
For Davies evolution of 1 dimensional systems with uniform geometrically-local, commuting, and translation invariant Hamiltonians, it was shown in [7] that there exists a strictly positive cMLSI constant $\alpha(\mathcal{L}_\Lambda^D) = \Omega((\ln |\Lambda|)^{-1})$ at any temperature. For hypercubic latices in dimensions $D \geq 2$, the Schmidt generators, see Section 4.2 of a system with uniform nearest neighbour commuting Hamiltonian was shown to satisfy a cMLSI with constant $\alpha(\mathcal{L}_\Lambda^D) = O(1)$ under certain mixing condition. Hence for these systems this directly implies rapid thermalization, see also [22].

Remark. The main result of this work establishes rapid thermalization, informally speaking, of nearest neighbour quantum spin systems on 2-colorable lattices with finite growth constant whenever their generator is gapped. For exponential lattices, such a $b-$ary trees (under suitable conditions) this is a novel result and although in the hypercubic lattice case this was already known, the main result of this work constitutes an improvement over literature in these cases.

3 Clustering and Strong local indistinguishability

In this section we deal with some static properties of quantum spin systems. Namely various types of spacial clustering and spacial mixing properties. Among these rather weak notions, such as exponential decay of correlation, to stronger ones such as exponential decay of mutual information. Clustering properties and the often from these derived mixing conditions are central to the geometric divide-and-conquer arguments that establish entropic decay for quantum spin systems. Both in the quantum setting, see e.g [9], [11], [12] and also the classical setting, see e.g. [35]. In this section we will show that exponential decay of correlations implies an a priori stronger version of clustering and a new stronger form of a spacial mixing condition, which we coin *strong local indistinguishability*. We will use this notion and the results from this section to derive an even stronger clustering result in Chapter 5, however, they are also of independent interest. The results in this section are a crucial step in the proof of the system-size independence of the MLSI constant of the quantum Gibbs samplers introduced above in Chapter 4. We first introduce a useful relation to simplify notations. Then in Theorem 3.4 we establish, for uniform geometrically-local, bounded, commuting Hamiltonians on a graph with finite growth constant, that uniform exponential $\mathbb{L}_\infty$ decay of correlations (2) implies, one, strong local indistinguishability, two, something referred to as '(strong) mixing condition' (or strong tensorization) [7], [12], and three, is equivalent to exponential decay of the mutual information. In the 1-dimensional setting we show these implications qualitatively[1] without the commutativity requirement on the Hamiltonian in Theorem 3.5.

The version of clustering of correlations we will be looking at is the usual following notion of exponential decay of correlations, following the nomenclature of [11], denoted as as L_∞-clustering.

Definition 2 ($\mathbb{L}_\infty$-clustering)**.** *We call a pair of Potential Φ on Λ, and inverse temperature β, uniformly exponentially L_∞-clustering if for any subregion $\Gamma \subset\subset \Lambda$ and any $A, B \subset \Gamma$, s.t.* $\operatorname{dist}(A, B) = l$ *there exists an exponentially decaying function $l \mapsto \epsilon(l)$, s.t.*

$$\operatorname{Cov}^{(0)}_{\sigma^\Gamma}(f, g) \leq \|f\|\|g\||\Gamma|\epsilon(l), \tag{3.1}$$

for any self-adjoint $f, g \in \mathcal{B}(\mathcal{H}_\Gamma)$ with support on A, B, respectively. Here σ^Γ is the Gibbs state of H_Γ to inverse temperature β and $\operatorname{Cov}^{(0)}_{\sigma}(f, g)$ *the GNS-covariance defined in Section 2.3. The decay length of the function $\epsilon(l)$ is called the correlation length ξ, that is the standard decay rate of thermal two-point correlation functions. I.e.* $-\ln \epsilon(l) = \mathcal{O}\left(\frac{l}{\xi}\right)_{l\to\infty}$.

Remark. This is a rather weak notion of clustering, hence sometimes refereed to as *weak clustering*. It is known to imply local indistinguishability [4], [6] and a mixing condition [7], something elaborated in in the next section. It is for example known to hold for steady states of rapidly mixing QMS [3], [5], [15] and in particular for steady states of gapped primitive QMS, as in the

[1]However, in the 1 dimensional setting only the exponential strong local indistinguishability is a novel result. And the decay rate may not be ξ

following theorem, whose proof makes use of the detectibility lemma [36]. This is a very important implication, which we henceforth shall refer to as "gap implies exponential decay of correlations". The prefactor $|\Gamma|$ is a in general too cautious choice and can in many cases actually be relaxed or omitted, see [6] or [7][1D case]. Intuitively, one could assume that it should never occur in a physically relevant definition of exponential decay of correlations. Since it won't make a qualitative difference in this work, we will, for sake of generality, keep it as $|\Gamma|$, though. If we would remove it, then all the dependencies on sizes of regions in the theorems of this section could be dropped. However, the scalings w.r.t the sizes of boundaries of regions would be unaffected by this.

Theorem 3.1 (Gap implies $\mathbb{L}_\infty$ clustering [3](Corollary 27), adapted)**.** Let Λ be a graph with finite growth constant. Let $\{H_\Gamma\}_{\Gamma\subset\subset\Lambda}$ be a uniform, bounded, geometrically-local, commuting family of Hamiltonians with Gibbs states $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$ to some inverse temperature β. We say that a family of Lindbladians $\mathcal{L}_\Lambda := \{\mathcal{L}_\Gamma\}_{\Gamma\subset\subset\Lambda}$ is gapped if

$$\inf_{\Gamma\subset\subset\Lambda} \lambda(\mathcal{L}_\Gamma) > 0.$$

If a local Gibbs sampler[2] is gaped, then the Gibbs states (invariant states) satisfy $\mathbb{L}_\infty$-clustering, where the exponentially decaying function ϵ does not carry the $|\Gamma|$ factor. [3]

In fact, the authors of [3] show a somewhat stronger statement namely $\mathbb{L}_2$-clustering in form of $\mathrm{Cov}_\sigma(f,g) \leq \|f\|_{2,\sigma}\|g\|_{2,\sigma}\epsilon(l)$, where $\epsilon(l)$ is an exponentially decaying function in l the distance of the supports of f and g. By the monotonicity of the $\mathbb{L}_{p,\sigma}$−norms in p this directly implies $\mathbb{L}_\infty$-clustering. Note also, that there is no dependence on $|\Gamma|$ in this decay of correlations! Hence the assumption of a gapped generator will give us notions of clustering where we can omit the system and subsystem sizes in the exponential bounds. However, it is notably not as strong as a a form of clustering which we will be requiring and considering in Chapter 5, Section 5.1. There, as the first big result of that section, we will, hence, establish a stronger implication under appropriate conditions.

3.1 A useful relation

We first define the following relation to simplify the notation in the rest of this work.

Definition 3 (A strong similarity relation)**.** *Introduce the relation $\sim$ on $\mathcal{D}(\mathcal{H}_\Lambda)$, s.t. we write for two states ω,τ with the same support* $\mathrm{supp}(\omega) = \mathrm{supp}(\tau)$*:*

$$\omega \overset{\epsilon}{\sim} \tau :\Leftrightarrow \|\omega^{\frac{1}{2}}\tau^{-1}\omega^{\frac{1}{2}} - \mathbb{1}\| \leq \epsilon < 1, \tag{3.2}$$

where the identity $\mathbb{1} \equiv \mathbb{1}_{\mathrm{supp}(\omega)} = \mathbb{1}_{\mathrm{supp}(\tau)}$ is on the support of the states. The inverse represents the generalized inverse here, i.e. the inverse on the support times the support projection.

Note, that by Hölder's inequality it follows immediately that $\omega \overset{\epsilon}{\sim} \tau \implies \|\omega - \tau\|_1 \leq \epsilon$, but the converse is in general not true. Hence, this relation quantifies a stronger form of similarity between a pair of states. This turns out to be a natural and powerful notion when working with Gibbs states of local Hamiltonians.
We will often have ϵ be some (exponentially decaying) function depending on the supports of

[2]This is a uniform, locally primitive, locally GNS-symmetric and reversible, frustration-free family of Lindbladians with the Gibbs states $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$ as unique invariant states.

[3]This holds hence in particular for the Davies and Schmidt generators, which are introduced in 4.1 and 4.2, respectively.

ω, τ and may sloppily write $\omega \sim \tau$ when meaning that there exists some exponentially decaying function $\epsilon(l)$ s.t. $\omega \overset{\epsilon}{\sim} \tau$, where $l = \mathrm{dist}(\mathrm{supp}\,\omega, \mathrm{supp}\,\tau)$ and the exact function is not relevant. In this sense the mathematical terminology *relation* is justified, as per the following proposition.

Proposition 3.2 (Properties of $\overset{\epsilon}{\sim}$)**.** Let $A, B, C, D, E\tilde{A}, \tilde{B} \in B_1(\mathcal{H})$ be self-adjoint and full rank and $P \in \mathcal{B}(\mathcal{K})$ a projection. The above defined relation $\overset{\epsilon}{\sim}$ is reflexive (0), symmetric (1), transitive (2), and tensor multiplicative (3) in the following senses. Additionally it satisfies the natural localization and normalization properties 4) and 4′).

0) $A \sim A$	reflexive
1) $A \overset{\epsilon}{\sim} B \implies B \overset{\epsilon(1-\epsilon)^{-1}}{\sim} A$	symmetric
2) $A \overset{\epsilon_1}{\sim} B, B \overset{\epsilon_2}{\sim} C \implies A \overset{\eta}{\sim} C$	transitive
3) $A \overset{\epsilon_1}{\sim} \tilde{A}, B \overset{\epsilon_2}{\sim} \tilde{B} \implies A \otimes B \overset{\eta}{\sim} \tilde{A} \otimes \tilde{B}$	tensor multiplicative
4) $D \overset{\epsilon}{\sim} E \implies \mathrm{tr}_{\mathcal{K}}(\mathbb{1} \otimes P)D(\mathbb{1} \otimes P) \overset{\epsilon}{\sim} \mathrm{tr}_{\mathcal{K}}(\mathbb{1} \otimes P)E(\mathbb{1} \otimes P)$	locally preserved
4′) $D \overset{\epsilon}{\sim} E \implies \frac{\mathrm{tr}_{\mathcal{K}}(\mathbb{1}\otimes P)D(\mathbb{1}\otimes P)}{\mathrm{Tr}[(\mathbb{1}\otimes P)D(\mathbb{1}\otimes P)]} \overset{\epsilon(2+\epsilon)}{\sim} \frac{\mathrm{tr}_{\mathcal{K}}(\mathbb{1}\otimes P)E(\mathbb{1}\otimes P)}{\mathrm{Tr}[(\mathbb{1}\otimes P)E(\mathbb{1}\otimes P)]}$	normalization preserved

where $\eta = \epsilon_1(1 + \epsilon_2) + \epsilon_2$.

For notational simplicity, we may write $A \overset{\epsilon_1}{\sim} B \overset{\epsilon_2}{\sim} C$, implying transitivity, when we mean $A \overset{\epsilon_1}{\sim} B$, $B \overset{\epsilon_2}{\sim} C$.

Corollary 3.3. If $A_i \overset{\epsilon}{\sim} A_{i+1}$ for i= 0,...,K-1, then $A_0 \overset{\eta}{\sim} A_K$ with $\eta = (1 + \epsilon)^K - 1$.

A proof of the proposition and the corollary is given in Section A.1.

3.2 Local and Strong local indistinguishability

Local indistinguishability[4], [6] pertains to observing quantum many-body states on finite subregions $A \subset \Lambda$ and quantifying the influence of spatially far away (from A) regions on the marginal on subregion A.
We say that a family of Gibbs states $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$ satisfies *exponential uniform local indistinguishability* [4] if, for any subregion $\Gamma \subset\subset \Lambda$ and any partition thereof into disjoint regions $\Gamma = ABC$, the effect of subregion C on A with $\mathrm{dist}(A, C) = l$ is exponentially decaying in l. I.e. there exists an exponentially decreasing function $l \mapsto \epsilon(l)$, s.t.

$$\| \mathrm{tr}_{BC}[\sigma^{ABC}] - \mathrm{tr}_B[\sigma^{AB}]\|_1 \equiv \|\sigma_A - \mathrm{tr}_B[\sigma^{AB}]\|_1 \le |\partial C|\epsilon(l).$$

Here, B here shields A away from B and the function ϵ, as defined above, is independent of the regions (and their sizes) A, B, C. In [4] it was shown that Gibbs states of geometrically-local, bounded, possibly non-commuting Hamiltonians, which satisfy universal exponential-$\mathbb{L}_\infty$-decay of correlations, satisfy universal exponential local indistinguishability. Furthermore they show that, if the Hamiltonian is commuting, then the decay length of the function in the local indistinguishability can be controlled by the thermal correlation length ξ. It implies that expectation values of local observables can be evaluated quasi-locally, a property that plays an important role in establishing stability of gapped ground state phases [4], [6].

Similarly, Uuing the strong similarity relation defined above in Definition 3, we may now, in analogy to the above, define the stronger notion of *strong local indistinguishability*.

Definition 4 (Strong local indistinguishability). *We say the family of Gibbs states $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$ satisfies (exponential uniform) strong local indistinguishability if for any finite subregion $\Gamma \subset\subset \Lambda$ and any partition of it into $\Gamma = ABC$, s.t. B shields A away from C and* $\mathrm{dist}(A, C) = l$*, there exists an exponentially decreasing function $l \mapsto \epsilon(l)$, s.t. the following holds.*

$$\mathrm{tr}_{BC}\,\sigma^{ABC} \overset{\epsilon(l)}{\sim} \mathrm{tr}_B\,\sigma^{AB}. \tag{3.3}$$

Note that here $\epsilon(l)$ may depend on the sizes of the regions A, B, C.

We will show that the Gibbs state of a geometrically local, bounded, and commuting Hamiltonian H at inverse temperature β satisfies strong local indistinguishability on any lattice with finite growth constant ν if the pair (H, β) satisfies uniform exponential $\mathbb{L}_\infty$-clustering, and that the commuting property for 1D systems may be dropped. For commuting Hamiltonians we will also give quantitative results in terms of the correlation length ξ and inverse temperature β in Theorem 3.4. Before we come to this though, let's consider one more important property of Gibbs states.

Definition 5 (Strong tensorization/mixing condition from [7]). *We say the family of Gibbs states $\{\sigma^\Gamma\}_{\Gamma\subset\subset\Lambda}$ is uniformly exponentially strongly mixing if for any finite subregion $\Gamma \subset\subset \Lambda$ and any partition of it into $\Gamma = ABC$, s.t. B shields A away from C and* $\mathrm{dist}(A, C) = l$*, there exists an exponentially decreasing function $l \mapsto \epsilon(l)$, s.t. the following holds.*

$$\sigma_{AC} \overset{\epsilon(l)}{\sim} \sigma_A \otimes \sigma_C, \tag{3.4}$$

Note that here $\epsilon(l)$ may depend on the sizes of the regions A, B, C.

Strong mixing of Gibbs states of geometrically-local, possibly non-commuting, bounded, and translation invariant Hamiltonians on a 1D lattice was shown to hold qualitatively under the condition of uniform exponential $\mathbb{L}_\infty$-clustering [7][Proposition 8.1]. This was crucial in order to establish the existence of a log-decreasing MLSI constant $\alpha = \Omega(\ln|\Lambda|)^{-1}$ for commuting quantum spin chain systems.
We build upon this work to extend this result to hold for geometrically-local, commuting, bounded Hamiltonians on any lattice with finite growth constant. Moreover, we make our statement depend explicitly on the inverse temperature β, as well as the thermal correlation length ξ in the following Theorem 3.4.

Theorem 3.4 (Implications of $\mathbb{L}_\infty$-clustering: Strong local indistinguishability and more; commuting case). Let Λ be any graph with finite growth constant ν. Let $\Gamma = ABC \subset\subset \Lambda$, with $l := \mathrm{dist}(A, C) \geq 2r$ and let $\{H_\Lambda\}_{\Lambda\subset\subset\Gamma}$ be a uniform, bounded, commuting, geometrically-r-local family of Hamiltonians on Λ that satisfy universal exponential $\mathbb{L}_\infty$-clustering at inverse temperature β with correlation length ξ. Then its Gibbs state on Γ at inverse temperature β, satisfies

1) strong local indistinguishability with decay length ξ:

$$\mathrm{tr}_{BC}(\sigma^{ABC}) \overset{\epsilon(l)}{\sim} \mathrm{tr}_B(\sigma^{AB}) \text{ with } \epsilon(l) := e^{O(\beta\min\{|\partial A|,|\partial B|\})}O(|\partial C||ABC|)\exp\left(-\frac{l-r}{\xi}\right). \tag{3.5}$$

Hence this family of Gibbs states satisfies uniform strong local indistinguishability.

2) strong tensorization with decay length ξ:

$$\sigma_{AC} \overset{\eta(l)}{\sim} \sigma_A \otimes \sigma_C \text{ with } \eta(l) := e^{O(\beta(|\partial A|+|\partial C|))}O(\mathrm{poly}(|A|,|B|,|C|))\exp\left(-\frac{l-2r}{\xi}\right). \tag{3.6}$$

Hence this family of Gibbs states satisfies uniform strong mixing.

3) exponential decay of mutual information with decay length ξ, i.e.

$$I_{\sigma^{ABC}}(A:C) \leq \eta(l) = e^{O(\beta(|\partial A|+|\partial C|))} O(\operatorname{poly}(|A|,|B|,|C|)) \exp\left(-\frac{l-2r}{\xi}\right). \tag{3.7}$$

Hence this family of Gibbs states satisfies uniform exponential decay of mutual information.

Remark: Note that in order to guarantee the universality of these three properties we need a finite growth constant, under which the bounds $|\partial A|, |\partial B|, |\partial C| \leq \nu^r < \infty$ hold. Since by Theorem 3.1 the existence of a gap implies uniform exponential $\mathbb{L}_\infty$-clustering, we immediately have strong local indistinguishability, strong mixing, and exponential decay of mutual information from the gap property. Explicitly, this implies that 1 dimensional quantum spin chains satisfy these properties at any temperature for geometrically-local, commuting, bounded Hamiltonians and in d-dimensional regular latices if the temperature is high enough. With this in mind, implication 1), i.e. gap implies strong local indistinguishability, is a strict strengthening of the local indistinguishability result in [4], [6] in the case of commuting Hamiltonians. Implication 2) and 3) can be viewed as extensions of the results in [7] to any lattice with finite growth constant under the additional condition of commutativity of the Hamiltonian. By a standard use of Pinsker's and Hölder's inequalities, exponential decay of the mutual information directly implies exponential decay of GNS-Correlations with halved decay rate, since

$$I_\sigma(A:C) = D(\sigma_{AC}\|\sigma_A \otimes \sigma_C) \geq \frac{1}{2}\|\sigma_{AC} - \sigma_A \otimes \sigma_C\|_1^2.$$

Thus what we are showing here is equivalence of exponential decay of mutual information, a seemingly strictly stronger type of decay; and a more basic exponential decay of correlations. In summary, to establish mutual information decay for the above considered systems it is enough to establish decay of correlations. For an exemplary visualization of a splitting of $\Gamma = ABC$ see Figure 3.1 and for a graphical representation of the implications and relation between the different notions of clustering dealt with there, see the static properties section part in Figure 5.1.

Before proving this theorem we note that in the case of 1 dimensional quantum spin chains we can also establish the above results and implications for non-commuting Hamiltonians. In this case 2) and 3) are the main results of [7].

Theorem 3.5 (Strong local indistinguishability in 1D)**.** Let $I = ABC \subset\subset \mathbb{Z}$ be a convex, where B shields A away from C, s.t. $2l := |B| = \operatorname{dist}(A,C)$ and let H be a geometrically-r-local and $J-$bounded Hamiltonian on Λ which satisfies universal exponential clustering.

$$\|(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B \sigma^{AB})^{-1} - \mathbb{1}\| \leq Ke^{-al}, \tag{3.8}$$

with some $K, a > 0$ depending only on the interaction range r and $J\beta$, where we recall that J is the interaction strength and $\beta > 0$ the inverse temperature.

Theorem 3.5 will be proven in Appendix A.1. Its proof is however essentially the same as the one for Theorem 3.4 using some additional technical prerequisites from [7]. For the latter we will first need the following technical Lemma. Recall that $E_{A,B} := \exp(-\beta H_{AB})\exp(\beta(H_A + H_B))$ denote Araki's expansionals.

Lemma 3.6. Let Φ be a geometrically$-r-$local, $J-$bounded, commuting potential on a quantum spin system Λ with finite growth constant ν. Let $\Gamma = ABC \subset\subset \Lambda$. Then the following bounds hold with $K = \exp(O(\beta|\partial B|))$ independent of l and ξ:

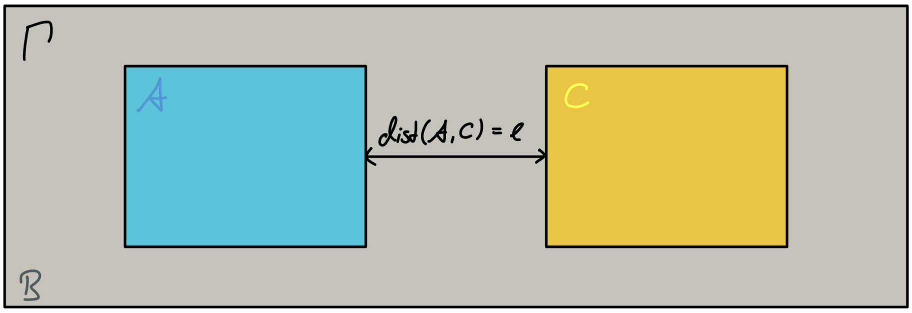

Figure 3.1: The graph region Γ, here depicted as a colored region without the edges is partitioned into the three sub-regions A, C and B, such s.t. dist(A,C)=l. This is an example of a partitioning in Theorem 3.4. If the Graph Γ is in a state which satisfies $\mathbb{L}_\infty$-clustering, then via this theorem, the reduced state on AC is approximately a product state between the reduced states on regions A and C. Approximately in the sense of the strong similarity relation $\overset{\epsilon}{\sim}$ from 3, where ϵ is exponentially decaying in l.

0) $\|E^{\pm 1}_{A,B}\| \le K = \exp(O(\beta|\partial B|))$

1) $\|Q^{\mp 1}\|^{-1} \le \|\operatorname{tr}_B[\sigma^B Q]^{\pm 1}\| \le \|Q^{\pm 1}\|$ for any strictly positive $Q \in \mathcal{B}(\mathcal{H}_\Lambda)$

1′) $\|\operatorname{tr}_B[\sigma^B E^{\pm 1}_{A,B}]^{\pm 1}\| \le K$, $\|\operatorname{tr}_B[\sigma^B E^{\pm 1}_{B,C}]^{\pm 1}\| \le K$, $\|\operatorname{tr}_{AB}[\sigma^{AB} E^{\pm 1}_{A,B}]^{\pm 1}\| \le K$, $\|\operatorname{tr}_B[\sigma^B E^{\pm 1}_{A,B} E^{\pm 1}_{AB,C}]^{\pm 1}\| \le K^2$.

The big-O notation refers to the dependence in β and $|\partial B|$ and omits dependence on J, d, r, ν.

Proof. To show 1) consider the map $Q \mapsto \operatorname{tr}_B[\sigma^B Q] = \frac{1}{\operatorname{Tr}[e^{-H_B}]} \operatorname{tr}_B[e^{-\frac{1}{2}H_B} Q e^{-\frac{1}{2}H_B}] : \mathcal{B}(\mathcal{H}_\Lambda) \to \mathcal{B}(\mathcal{H}_{\Lambda\setminus B})$ which is evidently positive and unital. Note that if $Q > 0$ is strictly positive, then so is $Q^{-1} > 0$ and $\lambda_{\min}(Q)\mathbb{1} = \|Q^{-1}\|^{-1}\mathbb{1} \le Q \le \|Q\|\mathbb{1} = \lambda_{\max}(Q)\mathbb{1}$. Applying the aforementioned map to this inequality immediately gives that

$$\|Q^{-1}\|^{-1}\mathbb{1}_{B^c} \le \operatorname{tr}_B[\sigma^B Q] \le \|Q\|\mathbb{1}_{B^c},$$
$$\|Q\|^{-1}\mathbb{1}_{B^c} \le \operatorname{tr}_B[\sigma^B Q]^{-1} \le \|Q^{-1}\|\mathbb{1}_{B^c},$$

since inversion of two commuting operators is order reversing. Taking norms gives 1).
For 0), if Φ is chosen as in the statement of proposition, then there exists a constant $c_{r,\nu}$ depending only on r, ν s.t.

$$\begin{aligned}
\|E^{\pm 1}_{A,B}\| &= \|e^{\mp\beta H_{AB}} e^{\pm\beta H_A \pm \beta H_B}\| \\
&= \|\exp\big(\mp\beta \sum_{X\cap A\neq\emptyset, X\cap B\neq\emptyset} \Phi_X\big)\| \\
&\le \exp\big(\beta J \sum_{\substack{X\cap A\neq\emptyset, X\cap B\neq\emptyset \\ \operatorname{diam}(X)\le r}} 1\big) \\
&\le \exp(\beta J c_{r,\nu} \min\{|\partial A|, |\partial B|\}) =: K = \exp\big(O(\beta|\partial B|)\big),
\end{aligned}$$

and

$$\|E^{\pm 1}_{A,B}E^{\pm 1}_{AB,C}\| \leq \|E^{\pm 1}_{A,B}\|\|E^{\pm 1}_{AB,C}\| \leq K^2 = \exp\left(O(\beta|\partial B|)\right).$$

Note that 1′) are just special cases of 1) given 0), since each of the $E_{A,B}$ are strictly positive, as Φ is commuting, e.g. by self-adjointness and the spectral theorem. ◻

Remark. Note that the proof of Lemma 3.6 requires the commutativity of the Hamiltonian, since we require $E_{A,B} \geq 0$. If a proof of it, which does not require commutativity can be found then we are hopeful that we can establish Theorem 3.4 without the additional assumption of commutativity. For more details on this see the discussion in Chapter 7 on this.

By the combined use of Lemma 3.6, clever rewritings inspired by the proofs in [7], and repeated application of local indistinguishability, we can prove the main theorem of this section.

Proof of Theorem 3.4. We first note that the following holds: $\|A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} - \mathbb{1}\| \leq \|AB^{-1} - \mathbb{1}\|$. Set $l := \operatorname{dist}(A, C) > r$. Assume uniform exponential $\mathbb{L}_\infty$-clustering with correlation length ξ, hence we may write $l \mapsto \tilde{K}\exp\left(-\frac{l}{\xi}\right)$ for the exponentially decaying function, for some constant $K > 0$.
1) To show strong local indistinguishability (3.5) we start by rewriting

$$(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B \sigma^{AB})^{-1} = \operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\operatorname{tr}_B[\sigma^B E_{A,B}]^{-1}\lambda^{-1}_{ABC},$$

where $\lambda^{-1}_{ABC} = \frac{\operatorname{Tr}[e^{-\beta H_{AB}}]\operatorname{Tr}[e^{-\beta H_{BC}}]}{\operatorname{Tr}[e^{-\beta H_{ABC}}]\operatorname{Tr}[e^{-\beta H_B}]} = \frac{\operatorname{Tr}[\sigma^{ABC}E^{-1}_{A,BC}]}{\operatorname{Tr}[\sigma^{AB}E^{-1}_{A,B}]}$.
Claim 1: $|\lambda^{\mp 1}_{ABC} - 1|$ is exponentially decaying in l with decay length ξ.
Proof of Claim 1:

$$\begin{aligned}|\lambda^{-1}_{ABC} - 1| &= \frac{1}{\operatorname{Tr}[\sigma^{ABC}E^{-1}_{A,BC}]}\left|\operatorname{Tr}[\sigma^{AB}E^{-1}_{A,B} - \operatorname{Tr}[\sigma^{ABC}E^{-1}_{A,BC}]]\right| \\ &\overset{\text{Lemma 3.6 1)}}{\leq} \|E^{-1}_{A,BC}\|\left|\operatorname{Tr}[\sigma^{AB}E^{-1}_{A,B} - \operatorname{Tr}[\sigma^{ABC}E^{-1}_{A,BC}]]\right|.\end{aligned}$$

Now set $B = B_1B_2$ with $B_1 := A\partial$, $B_2 = B \setminus B_1$, s.t. $\operatorname{dist}(A, B_2) = r, \operatorname{dist}(B_1, C) = l - r$. Then $E_{A,BC} = E_{A,B} = E_{A,B_1}$. So now

$$\begin{aligned}\left|\operatorname{Tr}[\sigma^{AB}E^{-1}_{A,B} - \operatorname{Tr}[\sigma^{ABC}E^{-1}_{A,BC}]]\right| &= \left|\operatorname{Tr}_{AB_1}[\operatorname{tr}_{B_2}(\sigma^{AB}E^{-1}_{A,B})] - \operatorname{Tr}_{AB_1}[\operatorname{tr}_{B_2C}(\sigma^{ABC}E^{-1}_{A,BC})]\right| \\ &= |\operatorname{Tr}_{AB_1}[(\operatorname{tr}_{B_2}\sigma^{AB} - \operatorname{tr}_{B_2C}\sigma^{ABC})E^{-1}_{A,B_1}]| \\ &\overset{\text{Hölder}}{\leq} \|(\operatorname{tr}_{B_2}\sigma^{AB} - \operatorname{tr}_{B_2C}\sigma^{ABC})\|_1\|E^{-1}_{A,B_1}\| \\ &\leq K|\partial C||ABC|\tilde{K}\exp\left(\frac{-1}{\xi}\operatorname{dist}(B_1, C)\right) \\ &= K|\partial C||ABC|\tilde{K}\exp\left(-\frac{l-r}{\xi}\right).\end{aligned}$$

Thus $|\lambda^{-1}_{ABC} - 1| \leq K^2|\partial C||ABC|\tilde{K}\exp\left(-\frac{l-r}{\xi}\right)$. The same holds for $|\lambda_{ABC} - 1|$ by the same argument as above. Now we can rewrite

$$\begin{aligned}\|(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B\sigma^{AB})^{-1}_{\mathbb{1}}\| \leq& \|(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B\sigma^{AB})^{-1} - \operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}](\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\| \\ &+ \|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}](\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1} - \mathbb{1}\| \\ \leq& \|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\|\|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\||\lambda^{-1}_{ABC} - 1| \\ &+ \|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\|\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\|.\end{aligned}$$

Claim 2: $\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}] \|$ is exponentially decaying in l with decay rate ξ.
Proof of Claim 2: Again set $B = B_1 B_2$, $B_1 := A\partial$, $B_2 := B \setminus B_1$, thus $\operatorname{dist}(B_1, C) = l - r$ and $E_{A,BC} = E_{A,B} = E_{A,B_1}$. Then by local indistinguishability

$$\begin{aligned}
\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}] \| &= \| \operatorname{tr}_{B_1}[(\operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C}) E_{A,B_1}] \| \\
&\leq \| \operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C} \|_1 \| E_{A,B_1} \| \\
&\leq K |\partial C| |BC| \tilde{K} \exp\left(-\frac{\operatorname{dist}(B_1, C)}{\xi} \right) \\
&= K |\partial C| |BC| \tilde{K} \exp\left(-\frac{l - r}{\xi} \right),
\end{aligned}$$

where the first inequality follows similar to Lemma 3.6 1), since $Q \mapsto \operatorname{tr}_{B_1}[| \operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C} | Q]$ is positive and unital up to a scalar factor of $\operatorname{Tr}[| \operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C} |]$, i.e.

$$\begin{aligned}
\| \operatorname{tr}_{B_1}[(\operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C}) E_{A,B_1}] \| &\leq \| \operatorname{tr}_{B_1}[| \operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C} | E_{A,B_1}] \| \\
&\overset{\text{Lemma 3.6 1)}}{\leq} \| E_{A,B_1} \| \operatorname{Tr}_{B_1}[| \operatorname{tr}_{B_2} \sigma^{B_1 B_2} - \operatorname{tr}_{B_2 C} \sigma^{B_1 B_2 C} |].
\end{aligned}$$

Putting everything together, we get the desired result:

$$\begin{aligned}
\| (\operatorname{tr}_{BC} \sigma^{ABC})(\operatorname{tr}_B \sigma^{AB})^{-1}_{\mathbb{1}} \| &\leq 2K^2 K^2 |\partial C| |ABC| \tilde{K} \exp\left(-\frac{l - r}{\xi} \right) + CC |\partial C| |BC| \tilde{K} \exp\left(-\frac{l - r}{\xi} \right) \\
&= \mathcal{O}(K^4) |\partial C| |ABC| \tilde{K} \exp\left(-\frac{l - r}{\xi} \right) \\
&= \exp \mathcal{O}(\beta |\partial A|) \mathcal{O}(|\partial C| |ABC|) \exp\left(-\frac{\operatorname{dist}(A\partial, C)}{\xi} \right).
\end{aligned}$$

2) Assume $l \geq 2r$. To prove the strong tensorization (3.6), we can, similar to above (or see e.g. [7][Cor 8.3]), rewrite

$$\begin{aligned}
\| \sigma_{AC} (\sigma_A \otimes \sigma_C)^{-1} - \mathbb{1} \| \leq \;& \| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}]^{-1} \| \| \operatorname{tr}_{AB}[\sigma^{AB} E_{AB,C}]^{-1} \| \| \operatorname{tr}_B[\sigma^B E_{A,B} E_{AB,C}] \| \underbrace{|\lambda_{ABC} - 1|}_{\textit{Claim 1}} \\
&+ \| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}]^{-1} \| \| \operatorname{tr}_{AB}[\sigma^{AB} E_{AB,C}]^{-1} \| \\
&\cdot \underbrace{\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}] \operatorname{tr}_{AB}[\sigma^{AB} E_{AB,C}] - \operatorname{tr}_B[\sigma^B E_{A,B} E_{AB,C}] \|}_{\textit{Claim 3}} \\
\leq \;& K^4 K^2 |\partial C| |ABC| \tilde{K} \exp\left(-\frac{l - r}{\xi} \right) + K^2 (\textit{Claim 3}).
\end{aligned}$$

Claim 3: $\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}] \operatorname{tr}_{AB}[\sigma^{AB} E_{AB,C}] - \operatorname{tr}_B[\sigma^B E_{A,B} E_{AB,C}] \|$ is exponentially decaying in l with correlation length ξ.
Proof of Claim 3: Set $B = B_1 B_2 B_3$ with $B_1 := \partial A, B_3 := \partial C, B_2 := B \setminus (B_1 \cup B_3)$, then

$\mathrm{dist}(B_1, B_3) = l - 2r$ and $E_{A,BC} = E_{A,B} = E_{A,B_1}$ and $E_{AB,C} = E_{B,C} = E_{B_3,C}$ and consequently

$$\begin{aligned}
&\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}] \operatorname{tr}_{AB}[\sigma^{AB} E_{AB,C}] - \operatorname{tr}_B[\sigma^B E_{A,B} E_{AB,C}] \| \\
=&\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,B_1}] \operatorname{tr}_{AB}[\sigma^{AB} E_{B_3,C}] - \operatorname{tr}_B[\sigma^B E_{A,B_1} E_{B_3,C}] \| \\
\le& \underbrace{\| \operatorname{tr}_B[\sigma^B E_{A,B_1} E_{B_3,C}] - \operatorname{tr}_B[\sigma^B E_{A,B_1}] \operatorname{tr}_B[\sigma^B E_{B_3,C}] \|}_{\text{(I)}} \\
+& \underbrace{\| \operatorname{tr}_B[\sigma^B E_{A,B_1}] \operatorname{tr}_B[\sigma^B E_{B_3,C}] - tr_{BC}[\sigma^{BC} E_{A,B_1}] \operatorname{tr}_{AB}[\sigma^{AB} E_{B_3,C}] \|}_{\text{(II)}} .
\end{aligned}$$

To bound (II) we use that by the proof of *Claim 2*:

$$\begin{aligned}
\| \operatorname{tr}_B[\sigma^B E_{A,B_1}] - \operatorname{tr}_{BC}[\sigma^{BC} E_{A,B_1}] \| &\le \|E_{A,B_1}\| \| \operatorname{tr}_{B_2B_3} \sigma^B - \operatorname{tr}_{B_2B_3C} \sigma^{BC} \|_1 \\
&\le K^2 |\partial C| |BC| \tilde{K} \exp\left(-\frac{1}{\xi} \mathrm{dist}(B_1, C)\right), \\
\| \operatorname{tr}_B[\sigma^B E_{B_3,C}] - \operatorname{tr}_{AB}[\sigma^{AB} E_{B_3,C}] \| &\le \|E_{B_3,C}\| \| \operatorname{tr}_{B_1B_2} \sigma^B - \operatorname{tr}_{AB_1B_2} \sigma^{AB} \|_1 \\
&\le K^2 |\partial A| |AB| \tilde{K} \exp\left(-\frac{1}{\xi} \mathrm{dist}(A, B_3)\right).
\end{aligned}$$

Together

$$\begin{aligned}
\text{(II)} \le& \| \operatorname{tr}_B[\sigma^B E_{A,B_1}] \| \| \operatorname{tr}_B[\sigma^B E_{B_3,C}] - \operatorname{tr}_{AB}[\sigma^{AB} E_{B_3,C}] \| \\
&+ \| \operatorname{tr}_B[\sigma^B E_{B_3,C}] \| \| \operatorname{tr}_B[\sigma^B E_{A,B_1}] - \operatorname{tr}_{BC}[\sigma^{BC} E_{A,B_1}] \| \\
&\le KK^2 |\partial A| |AB| \tilde{K} \exp\left(-\frac{l-r}{\xi}\right) + KK^2 |\partial C| |BC| \tilde{K} \exp\left(-\frac{l-r}{\xi}\right) \\
&= \exp(O(\beta|\partial A| + \beta|\partial C|))(|\partial A||AB| + |\partial C||BC|) \exp\left(-\frac{l-r}{\xi}\right).
\end{aligned}$$

To bound (I) we use that the $\mathbb{L}_\infty$-clustering directly implies $\|\sigma_{AC} - \sigma_A \otimes \sigma_C\|_1 \le |ABC| \tilde{K} \exp\left(-\frac{\mathrm{dist}(A,C)}{\xi}\right)$ for $\sigma \equiv \sigma^{ABC}$, by use of Hölder duality. Thus

$$\begin{aligned}
\text{(I)} &= \| \operatorname{tr}_{B_1B_3}[\sigma_{B_1B_3}(E_{A,B_1} \otimes E_{B_3,C})] - \operatorname{tr}_{B_1}[\sigma_{B_1} E_{A,B_1}] \otimes \operatorname{tr}_{B_3}[\sigma_{B_3} E_{B_3,C}] \| \\
&= \| \operatorname{tr}_{B_1B_3}[\sigma_{B_1B_3}(E_{A,B_1} \otimes E_{B_3,C}) \mp (\sigma_{B_1} \otimes \sigma_{B_3})(E_{A,B_1} \otimes E_{B_3,C})] - \operatorname{tr}_{B_1}[\sigma_{B_1} E_{A,B_1}] \otimes \operatorname{tr}_{B_3}[\sigma_{B_3} E_{B_3,C}] \| \\
&= \| \operatorname{tr}_{B_1B_3}[(\sigma_{B_1B_3} - \sigma_{B_1} \otimes \sigma_{B_3})(E_{A,B_1} \otimes E_{B_3,C})] + \\
&+ \underbrace{\operatorname{tr}_{B_1B_3}[(\sigma_{B_1} \otimes \sigma_{B_3})(E_{A,B_1} \otimes E_{B_3,C})] - \operatorname{tr}_{B_1}[\sigma_{B_1} E_{A,B_1}] \otimes \operatorname{tr}_{B_3}[\sigma_{B_3} E_{B_3,C}]}_{=0} \| \\
&\overset{\text{Proof of Claim 2}}{\le} \|E_{A,B_1} \otimes E_{B_3,C}\| \|\sigma_{B_1B_3} - \sigma_{B_1} \otimes \sigma_{B_3}\|_1 \\
&\le KK|B|\tilde{K} \exp\left(-\frac{\mathrm{dist}\, B_1, B_3}{\xi}\right) = K^2 |B| \tilde{K} \exp\left(-\frac{l-2r}{\xi}\right).
\end{aligned}$$

So together $\| \operatorname{tr}_{BC}[\sigma^{BC} E_{A,BC}] \operatorname{tr}_{AB}[\sigma^{AB} E_{AB,C}] - \operatorname{tr}_B[\sigma^B E_{A,B} E_{AB,C}] \| \le \exp(O(\beta|\partial A| + \beta|\partial C|))(|\partial A||AB| + |B| + |\partial C||BC|)\tilde{K} \exp\left(-\frac{l-2r}{\xi}\right)$ and altogether

$$\|\sigma_{AC}(\sigma_A \otimes \sigma_C)^{-1} - \mathbb{1}\| \le \exp O(\beta(|\partial A| + |\partial C|)) O(\mathrm{poly}(|A|, |B|, |C|)) \exp\left(-\frac{l-2r}{\xi}\right).$$

3) The result for the mutual information follows directly from the one for the mixing condition 2) (3.6) via the following inequality [7][Lemma 3.1]

$$I_\sigma(A:C) \leq \|\sigma_{AC}(\sigma_A \otimes \sigma_C)^{-1} - \mathbb{1}_{AC}\|.$$

. $\square$

4 Local Generators and conditional expectations

The previous section was on statics of quantum spin systems. Next, we want to consider the dynamics of these spin systems. In this work we consider mainly two classes of dynamics. The first one, commonly known as Davies dynamics, is a particularly physical motivated dynamic. It is often used to model thermalization of finite dimensional quantum systems from weak coupling to its environment. The main result is formulated w.r.t it. See Section 4.1 for Details of the Davies evolution. In contrast, the second class is the is the *Heat-bath dynamics* w.r.t the *Schmidt conditional expectation*, see Section 4.2 for its definition and the notation used. It serves as a mathematically simpler model, but lacks physical interpretation.
In this work we will use the latter as a proxy to derive rigorous bounds on the mixing time of the former through establishment of the modified logarithmic Sobolev inequality.

4.1 Davies Evolution and Local Davies Generators

The Davies evolution is a Markovian approximation of the reduced state dynamics of a many-body spin system weakly-coupled to an infinite-dimensional environment in thermal equilibrium. Although it is known that reduced evolutions of quantum systems are never exactly Markovian, this approximation is very powerful. In general, the open system dynamics described by a master equations, which always has a QMS as a solution, is of high interest to the quantum optics, condensed matter, chemical physics, statistical physics, quantum information, and mathematical physics communities. The interest in Markovian description of open system dynamics has especially followed a the rise of interest in quantum information theory and decoherence phenomena. Moreover, Davies evolutions appear widely in the literature concerning thermalization of quantum systems, both from physical and computational view points [2], [3], [7], [11]. It was originally studied by Davies in [14]. For a great overview of general open system dynamics, including a derivation of the weak coupling limit (in Section 6) see [13]. A sketch follows. We assume that our spin system $\mathcal{H}_\Lambda$ is in contact with an environment $\mathcal{H}_E$, where the Hamiltonian of the combined closed system $\mathcal{H}_\Lambda \otimes \mathcal{H}_E$ is given by

$$H_{\Lambda E} = H_\Lambda \otimes \mathbb{1}_E + \mathbb{1}_\Lambda \otimes H_E + \lambda \sum_{\alpha(x), x\in\Lambda} A_{\alpha(x)} \otimes B_{\alpha(x)},$$

where $\{\alpha(x)\}$ labels a set of operators acting on site $x \in \Lambda$ and the operators $\{A_{\alpha(x)}\}_{\alpha(x),x\in\Lambda}$ span all of $\mathcal{B}(\mathcal{H}_\Lambda)$. A typical example of these would be the generalized Pauli matrices indexed by α. λ is the interaction strength between the environment and the system, where we have that each spin is individually coupled to the bath. We further assume that the environment system is in its thermal state σ_E given by H_E at inverse temperature β, and the initial state of system and environment is a product state, however, this is a easily satisfiable assumption. The weak coupling limit is first the limit where we take the coupling constant $\lambda \to 0$ and the interaction time $t \to \infty$, s.t. $\lambda^2 t = \tau$ is held constant and then use the Born-Markov approximation. It now turns out that under a certain convergence condition [13][Theorem 6.1], which amongst others requires the environment system to be infinite dimensional, the resulting reduced dynamics of the system is given by the following Davies evolution

$$\lim_{\lambda\to 0, \tau=\lambda t^2} \mathrm{tr}_E[e^{-i\tau H_{\Lambda E}}(\rho\otimes\sigma_E)e^{i\tau H_{\Lambda E}}] = e^{\tau \mathcal{L}^D_{\Lambda *}}(\rho), \tag{4.1}$$

where the limit converges in trace norm [13]. The corresponding Lindbladian is often called the *Davies generator* and is given by

$$\mathcal{L}^D_\Lambda(X) = i[H_\Lambda, X] + \sum_{x\in\Lambda} \mathcal{L}^D_x(X), \tag{4.2}$$

with

$$\mathcal{L}^D_x(X) = \sum_{\omega,\alpha(x)} \chi_{\alpha(x)}(\omega)\left(A^*_{\alpha(x)}(\omega)XA_{\alpha(x)}(\omega) - \frac{1}{2}(A^*_{\alpha(x)}(\omega)A_{\alpha(x)}(\omega)X + XA^*_{\alpha(x)}(\omega)A_{\alpha(x)}(\omega))\right), \tag{4.3}$$

with $\omega \in \sigma(H_\Lambda) - \sigma(H_\Lambda)$ the Bohr frequencies, $A_\alpha(x)(\omega)$ the operator Fourier coefficients of $e^{-itH_\Lambda}A_\alpha(x)e^{itH_\Lambda} = \sum_\omega e^{-it\omega}A_{\alpha(x)}(\omega)$, and $\chi_{\alpha(x)}(\omega) = 2\pi\,\mathrm{Tr}[B_{\alpha(x)}(\omega)B_\alpha(x)\sigma_E]$ the discrete Fourier transform of the two-point correlations function of the environment [3]. If we now assume that we have a uniformly bounded, geometrically-r-local commuting family of Hamiltonians $\{H_\Gamma\}_{\Gamma\subset\subset\Lambda}$, as in (2.14), then the above generator reduces to a *local Davies generators*

$$\mathcal{L}^D_\Gamma(X) = i[H_\Gamma, X] + \sum_{x\in\Gamma} \mathcal{L}^D_x(X), \tag{4.4}$$

for any $\Gamma \subset\subset \Lambda$. These correspond to a uniformly bounded, geometrically-local family of Lindbladians satisfying the following properties:

Proposition 4.1. ([3][Lemma 11]) For a finite graph Λ, a subset $\Gamma\subset\Lambda$ and $\{H_\Gamma\}_{\Gamma\subset\Lambda}$ a uniformly bounded, geometrically$-r$-local, commuting family of Hamiltonians, then the associated local Davies generators defined in (4.2-4.4) satisfy

1. For any subset $\Gamma\subset\Lambda$, $\{e^{t\mathcal{L}^D_\Gamma}\}_{t\geq 0}$ is a QMS with generator $\mathcal{L}^D_\Gamma$.
2. The family $\mathcal{L}^D = \{\mathcal{L}^D_\Gamma\}_\Gamma$ is geometrically-local, in the sense that each individual term $\mathcal{L}^D_x$ acts only nontrivially on the region $B_{\tilde r}(x)$ for some fixed $r \leq \tilde r \leq 2r$.
3. The family $\mathcal{L}^D = \{\mathcal{L}^D_\Gamma\}_\Gamma$ is locally primitive, locally reversible, and satisfies detailed balance w.r.t the global Gibbs state σ^Λ.[1]
4. The family $\mathcal{L}^D = \{\mathcal{L}^D_\Gamma\}_\Gamma$ is frustration free.

For a proof see e.g. [3]. Recall Section 2.6 Hence $E^D_\Gamma(X) := \lim_{t\to\infty} e^{t\mathcal{L}^D_\Gamma}(X)$ is a conditional expectation, called the *Davies conditional expectation*.

[1] i.e. the unique local full rank invariant states are σ^Γ, and $e^{\mathcal{L}^D_\Gamma}$ is KMS and GNS symmetric w.r.t these.

4.2 Schmidt Conditional Expectations

In order to get the main result, we will require a quite strong approximate tensorization statement from [11], which, however, is a statement on the so called *Schmidt conditional expectation*, also introduced in [11]. Since it is of central importance we will explicitly give their quite technical construction and some properties in the following. The construction and results in this section work for any 2-colorable graph of finite growth constant.

Let $\Lambda = (V, E_V)$ be a quantum spin system with nearest-neighbour, bounded, commuting Hamiltonian H, which hence can be writen as

$$H = \sum_{(i,j)\in E_V} h_{i,j},$$

where each term $h_{i,j}$ acts only non-trivially on vertices i and j. Then the Gibbs state of H with inverse temperature β is given by

$$\sigma = \frac{e^{-\beta H}}{\operatorname{Tr}[e^{-\beta H}]} = \frac{\prod_{\{i,j\}\in E_V} e^{-\beta h_{i,j}}}{\operatorname{Tr}[\prod_{\{i,j\}\in E_V} e^{-\beta h_{i,j}}]}.$$

Given some $A \subset \Lambda$ we will define a suitable *-algebra $\mathcal{N}_A$ and conditional expectation E_A^S onto it, which has the Gibbs state as an invariant state. For simplicity of notations, we do this for a singelton $A = \{a\}$, however, this construction works similarly for all $A \subset\subset \Lambda$. Given some $a \in \Lambda$, we enumerate the sets

$$\begin{aligned}\partial\{a\} &:= \{x \in \Lambda |\operatorname{dist}(x,a) = 1\} = \{b_i\}_{i\in I_a},\\ \partial\{b_i\} &:= \{x \in \Lambda |\operatorname{dist}(x,b_i) = 1\} = \{c_{i,j}\}_{j\in J^{(i)}},\ \text{s.t. } a = c_{i,0}\ \forall i.\end{aligned}$$

Hence $\partial(\partial\{a\}) \setminus \{a\} = \{y \in \Lambda |\operatorname{dist}(a,y) = 2\} = \{c_{i,j}\}_{i\in I_a, j\in J^{(i)}\setminus\{0\}}$. See Figure 4.1 for a graphical example of these definitions on a section of a 3-ary tree graph.

We will drop the index a of $I \equiv I_a$, the labeling of all the neighbours of a in the following. Now we Schmidt-decompose

$$e^{-\beta h_{b_i c_{ij}}} = \sum_s X_{b_i}^{j,s} \otimes X_{c_{ij}}^s$$

for $i \in I$, where the operators $\left\{X_{b_i}^{j,s}\right\}_{j,s} \subset \mathcal{B}(\mathcal{H}_{b_i})$ and for $j \in J^{(i)}$, $\left\{X_{c_{ij}}^s\right\}_s \subset \mathcal{B}(\mathcal{H}_{c_{ij}})$. We now define the *-algebras $\mathcal{A}_{b_j}^j$ to be generated by all $\left\{X_{b_i}^{j,s}\right\}_s$ [37].

Proposition 4.2. Any two non-identical of these algebras $\left\{\mathcal{A}_{b_i}^j\right\}_{i\in I, j\in J^{(i)}}$ commute.

Proof. [37] Consider $\mathcal{A}_{b_i}^j$ and $\mathcal{A}_{b_m}^n$. If $i \neq m$, then the statement is obvious, since their generators act on different Hilbert spaces $\mathcal{H}_{b_i}, \mathcal{H}_{b_m}$, respectively. If $i = m$, see that

$$0 = \left[e^{-\beta h_{b_i c_{ij}}}, e^{-\beta h_{b_i c_{in}}}\right] = \left[\sum_s X_{b_i}^{j,s} \otimes X_{c_{ij}^s}, \sum_r X_{b_i}^{n,r} \otimes X_{c_{in}}^r\right] \overset{j\neq n}{=} \sum_{s,r} X_{c_{ij}}^s \otimes X_{c_{in}}^r \otimes \left[X_{b_i}^{j,s}, X_{b_i}^{n,r}\right]$$
$$\implies \left[X_{b_i}^{j,s}, X_{b_i}^{n,r}\right] = 0\ \forall s, r.$$

Where the last implication follows since $\{X_{c_{ij}}^s\}_s, \{X_{c_{in}}^r\}_r$ form a set of linear independent operators by Schmidt decomposition. □

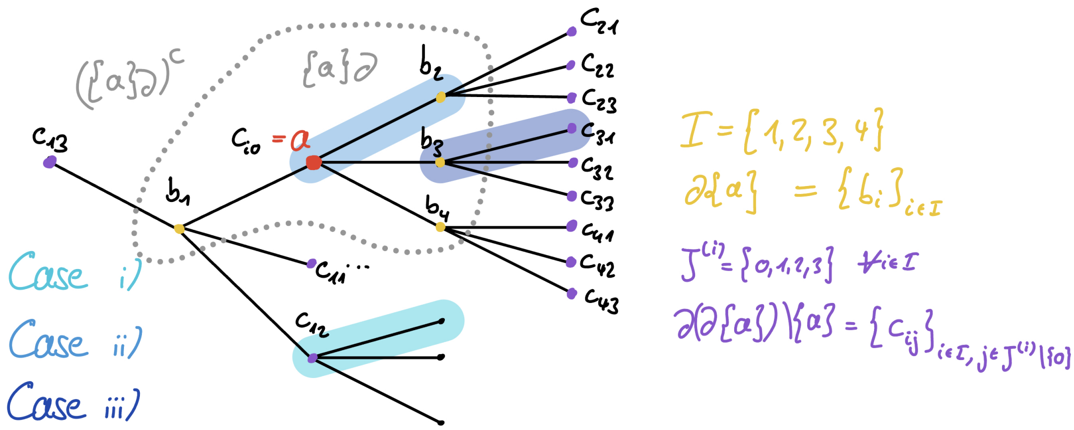

Figure 4.1: Simple example of the notation introduced in Section 4.2 to define the local algebras $\mathcal{N}_a$ and thus the Schmidt conditional expectation E_a^S. Depicted is a small region of a 3-ary tree, with a (red) vertex labeled a=$c_{i,0}$. Its neighbours (yellow) are labeled with $\{b_i\}_{i\in I} = \partial\{a\}$, where $I \equiv I_a = \{1,2,3,4\}$. The next-nearest neighbours (purple) are $\partial(\partial\{a\}) \setminus \{a\} = \{c_{i,j}\}_{i\in I, j\in J^{(i)}\setminus\{0\}}$. The central vertex a is logically the same as $c_{i,j0}$ for each $i \in I$. In the proof of proposition 4.3 the case i) considers for example the turquiose shaded edge, case ii) for example the blue shaded edge and case iii) for example the dark blue shaded edge. The boundary between the subsets $\{a\}\partial$ and $(\{a\}\partial)^c$ is marked with a gray dotted line.

Therefore these algebras and the underlying Hilbert spaces admit the following joint decomposition

$$\mathcal{H}_{b_i} := \bigoplus_{\alpha_i} \bigotimes_{j\in J^{(i)}} \mathcal{H}_j^{\alpha_i} \otimes \mathcal{H}_c^{\alpha_i} =: \bigoplus_{\alpha_i} P^{\alpha_i}\mathcal{H}_{b_i},$$

where P^{α_i} are orthogonal projectors such that

$$P^{\alpha_i}\mathcal{H}_{b_i} = \bigotimes_{j\in J^{(i)}} \mathcal{H}_j^{\alpha_i} \otimes \mathcal{H}_c^{\alpha_i},$$

where $\{\mathcal{H}_j^{(\alpha_i)}, \mathcal{H}_c^{(\alpha_i)}\}_{\alpha_i}$ are such that

$$\mathscr{A}_{b_i}^j = \bigoplus_{\alpha_i} \mathcal{B}(\mathcal{H}_j^{\alpha_i}) \otimes \mathbb{1}_{\bigotimes_{k\in J^{(i)}\setminus\{j\}} \mathcal{H}_k^{\alpha_i}\otimes\mathcal{H}_c^{\alpha_i}} \qquad \forall j \in J^{(i)}.$$

Definition 6. *Define the* $*$*-subalgebra* $\mathcal{N}_a := \mathbb{1}_{\{a\}} \otimes \bigotimes_{i\in I} \bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathscr{A}_{b_i}^j \otimes \mathcal{B}(\mathcal{H}_{(\{a\}\partial)})^c \subset \mathcal{B}(\mathcal{H}_\Lambda)$ *for any* $\{a\} \in \Lambda$.

Proposition 4.3. The modular group of the Gibbs state σ leaves this algebra invariant, for any $a \in \Lambda$, i.e.

$$\Delta_\sigma^{it}(\mathcal{N}_a) \subset \mathcal{N}_a \qquad \forall t \in \mathbb{R},$$

where $\Delta_\sigma^{it}(X) := \sigma^{it} X \sigma^{-it}\ \forall X \in \mathcal{B}(\mathcal{H}_\Lambda)$.

Proof. wlog $\beta = 1$. Fix $a \in \Lambda$. It is enough to show that $e^{ith_{kl}} \mathcal{N}_a e^{-ith_{kl}} \subset \mathcal{N}_a$ holds for any pair $(k,l) \in E_V$.
For examples of the following cases see also Figure 4.1. *Case i)* $(k,l) \subset (\{a\}\partial)^c$, then this is obvious, since $\mathcal{N}_a\big|_{\mathcal{H}_{(\{a\}\partial)^c}} = \mathcal{B}(\mathcal{H}_{(\{a\}\partial)^c})$.
Case ii) $(k,l) \subset \{a\}\partial$ with WLOG $k = a$, hence $l = b_i$ for some $i \in I$. Let $Y \in \mathcal{N}_a$ then via the spectral theorem $[e^{-ith_{ab_i}}, Y] = 0 \Leftrightarrow [e^{-h_{ab_i}}, Y] = 0$. It is enough to show this for the generators $Y = \mathbb{1}_{\{a\}} \otimes X^{j,s}_{b_i}$ for all s and $j \in J^{(0)} \setminus \{0\}$, by closeness of the algebra.

$$\left[e^{-h_{ab_i}}, \mathbb{1}_{\{a\}} \otimes X^{j,s}_{b_i}\right] = \left[\sum_r X^r_a \otimes X^{0,r}_{b_i}, \mathbb{1}_{\{a\}} \otimes X^{j,s}_{b_i}\right] = \sum_r X^r_a \otimes \left[X^{0,r}_{b_i}, X^{j,s}_{b_i}\right] = 0,$$

$\forall s, j \in J^{(i)} \setminus \{0\}$ since the algebras $\mathcal{A}^0_{b_i}$ and $\mathcal{A}^j_{b_i}$ commute for $j \in J^{(i)} \setminus \{0\}$.
Case iii) $(k,l) \subset \partial\{a\} \cup (\{a\}\partial)^c$. WLOG $k = b_i$ for some $i \in I$, hence $l = c_{ij}$ for some $j \in J^{(i)} \setminus \{0\}$. Then $e^{-\beta h_{kl}} = e^{-\beta h_{b_i c_{ij}}} \otimes \mathbb{1}_{\mathcal{H}_{\{k,l\}^c}} = \sum_s X^{j,s}_{b_i} \otimes X^s_{c_{ij}} \otimes \mathbb{1}_{\mathcal{H}_{\{k,l\}^c}} \in \mathcal{N}_a$. Thus by the spectral theorem $e^{\pm ith_{kl}} \otimes \mathbb{1} \in \mathcal{N}_a$ and hence by closedness of the algebra $e^{ith_{kl}} \mathcal{N}_a e^{-ith_{kl}} \subset \mathcal{N}_a$. □

Thus by Takesaki's theorem [27], see also [11][Proposition 10] due to Proposition 4.3, there exists a conditional expectation

$$E^S_a : B(\mathcal{H}_\Lambda) \to \mathcal{N}_a, \tag{4.5}$$

which has the Gibbs state σ, which is faithful, as an invariant state, i.e. $E^S_{a*}(\sigma) = \sigma$. It is called *Schmidt conditional expectation*.

Remark. Note that the above construction works exactly the same for any subsection $A \subset \Lambda$ in place of $\{a\} \subset \Lambda$, yielding a *-subalgebra $\mathcal{N}_A$. Hence we equally define the family of conditional expectations $\{E^S_A\}_{A \subset \Lambda}$ on $\mathcal{B}(\mathcal{H}_\Lambda)$. Similarly for these it holds that $E^S_{A*}(\sigma) = \sigma$. We can think of these conditional expectations as sort of replacing any given observable on the local subset with the identity, in such a way that is consistent with the invariance of the Gibbs state under its pre-adjoint. Hence, the family of Schmidt conditional expectations still have the desirable properties of the Davies expectations, i.e. that the Gibbs state is invariant, but their structure is easier to analyze since we can give an explicit expression for these which does not depend on system environment couplings. [11] We highlight one other important property of the Schmidt conditional expectations before we give an explicit form for these.

Proposition 4.4. For any two subsets $A_1, A_2 \subset \Lambda$, s.t. $\mathrm{dist}(A_1, A_2) \geq 2$, the Schmidt conditional expectations $E^S_{A_1}$ and $E^S_{A_2}$ satisfy

$$E^S_{A_1} \circ E^S_{A_2} = E^S_{A_2} \circ E^S_{A_1} = E^S_{A_1 \cup A_2},$$
$$E^S_{A_1} \circ E^S_{A_2} = E^S_{A_2} \circ E^S_{A_1} = E^S_{A_1}.$$

This follows from the fact that the conditional expectation E^S_A is a local map, acting only non trivially on $A\partial$ and the following Lemma.

Lemma 4.5. For any two subsets $A_1, A_2 \subset \Lambda$, s.t. $\mathrm{dist}(A_1, A_2) > 1$, or such that one is a subset of the other it holds that
1) $\mathcal{N}_{A_1} \cap \mathcal{N}_{A_1} = \mathcal{N}_{A_1 \cap A_2}$
2) $\mathcal{N}_{A_1} \cup \mathcal{N}_{A_2} = \mathcal{N}_{A_1 \cup A_2}$.
Here $\mathcal{N}_{A_1} \cup \mathcal{N}_{A_2}$ denotes the *-algebra generated by $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$. $\mathcal{N}_{A_1} \cap \mathcal{N}_{A_2}$ denotes the *-algebra generated by all elements in both $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$.

Proof. The proof is quite elementary from the definition of the algebras $\mathcal{N}$and may be found in Appendix A.1. □

For a subset $A \subset \Lambda$, we call a set $(\alpha) := \{\alpha_i\}_{i\in I_A}$ a *boundary condition.*

Proposition 4.6 (Explicit Form of Schmidt conditional Expectation, [11])**.** For $A \subset \Lambda$, let $(\alpha) := \{\alpha_i\}_{i\in I_A}$ be a fixed boundary condition for the subset A and denote $P^{(\alpha)} := \bigotimes_{i\in I_A} P^{\alpha_i}$. We set

$$\mathcal{H}^{(\alpha)}_{A_{\text{in}}} := \mathcal{H}_A \otimes \bigotimes_{i\in I_A} \mathcal{H}^{\alpha_i}_0 \otimes \mathcal{H}^{\alpha_i}_c \equiv \mathcal{H}_A \otimes \mathcal{H}^{(\alpha)}_{\partial_{\text{in}}A},$$
$$\mathcal{H}^{(\alpha)}_{A_{\text{out}}} := \mathcal{H}_{(A\partial)^c} \otimes \bigotimes_{i\in I_A} \bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathcal{H}^{\alpha_i}_j \equiv \mathcal{H}_{(A\partial)^c} \otimes \mathcal{H}^{(\alpha)}_{\partial_{\text{out}}A},$$

i.e. $\mathcal{H}_\Lambda = \bigoplus_{(\alpha)} \mathcal{H}^{(\alpha)}_{A_{\text{in}}} \otimes \mathcal{H}^{(\alpha)}_{A_{\text{out}}}$, and write $\text{tr}_{\mathcal{H}_{A^{(\alpha)}_{\text{in}}}} \equiv \text{tr}_{A_{\text{in}},(\alpha)}$, respectively, $\text{tr}_{\mathcal{H}_{A^{(\alpha)}_{\text{out}}}} = \text{tr}_{A_{\text{out}},(\alpha)}$ for simplicity.
Since every element of the algebra $\mathcal{N}_A$ is block diagonal w.r.t the sets (α), we can decompose the Schmidt conditional expectation E^S_A and its pre-adjoint E^S_{A*} along those blocks as well, yielding:

$$E^S_A(X) := \bigoplus_{(\alpha)} E^{S,(\alpha)}_A(X), \tag{4.6}$$
$$E^S_{A*}(\rho) := \bigoplus_{(\alpha)} E^{S,(\alpha)}_{A*}(\rho), \tag{4.7}$$

for any $X \in \mathcal{B}(\mathcal{H}_\Lambda)$ and $\rho \in \mathcal{D}(\mathcal{H}_\Lambda)$. $E^{S,(\alpha)}_A$ and $E^{S,(\alpha)}_A$ have the following expressions [11], defined on the block Hilbert space $P^{(\alpha)}\mathcal{H}_\Lambda$:

$$E^{S,(\alpha)}_A(X) = P^{(\alpha)} \left(\text{tr}_{A_{\text{in}},(\alpha)}[\tau^{(\alpha)}_{A_{\text{in}}} X] \otimes \mathbb{1}_{A_{\text{in}},(\alpha)}\right) P^{(\alpha)}, \tag{4.8}$$
$$E^{S,(\alpha)}_{A*}(\rho) = \text{tr}_{A_{\text{in}},(\alpha)}[P^{(\alpha)}\rho P^{(\alpha)}] \otimes \tau^{(\alpha)}_{A_{\text{in}}}, \tag{4.9}$$

where the state $\tau^{(\alpha)}_{A_{\text{in}}}$ is given by

$$\tau^{(\alpha)}_{A_{\text{in}}} := \frac{1}{\text{Tr}[\dots]} \text{tr}_{A_{\text{out}},(\alpha)}[P^{(\alpha)}\sigma^{(A\partial)}P^{(\alpha)}] = \frac{1}{\text{Tr}[\dots]} \text{tr}_{\partial_{\text{out}}A,(\alpha)}[P^{(\alpha)}\sigma^{(A\partial)}P^{(\alpha)}], \tag{4.10}$$

where the prefactors $\frac{1}{\text{Tr}[\dots]}$ contains the trace on $\mathcal{B}(\mathcal{H}^(\alpha)_{A_{\text{in}}})$ and ensure proper trace normalization $\text{Tr}[\tau^{(\alpha)}_{A_{\text{in}}}] = 1$ and the partial trace in the last expression traces out the Hilbert space $\mathcal{H}^{(\alpha)}_{\partial_{\text{out}}A} = \bigotimes_{i\in I_A} \bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathcal{H}^{\alpha_i}_j = \mathcal{H}^{(\alpha)}_{A_{\text{out}}\setminus(A\partial)^c}$,.

It is easy to check that these expressions (4.8) and (4.9) are dual to each other w.r.t. the Hilbert Schmidt inner product on $\mathcal{B}(\mathcal{H}_\Lambda)$. The expression (4.10) follows from the invariance of the local Gibbs states $\sigma^{(A\partial)}$ via $P^{(\alpha)}\sigma^{(A\partial)}P^{(\alpha)} = E^{S(\alpha)}_{A*}\left(P^{(\alpha)}\sigma^{(A\partial)}P^{(\alpha)}\right) = \text{tr}_{A_{\text{in}},(\alpha)}\left[P^{(\alpha)}\sigma^{(A\partial)}P^{(\alpha)}\right] \otimes \tau^{(\alpha)}_{A_{\text{in}}}$. Taking the partial trace of this expression on $\mathcal{H}^{(\alpha)}_{A_{\text{out}}}$ now gives the above expression.

Remark. We can think of the α_i as labeling the boundary conditions at site $i \in I_A$ and all $(\alpha) := \{\alpha_i\}_{i\in I_A}$ give a complete labeling of all boundary conditions of some subset A. Hence, we can think of the effect of the Schmidt-conditional expectation on states as effectively replacing the state ρ locally on A with the Gibbs state σ, where the boundary conditions to $(\alpha) = \{\alpha_i\}_{i\in I_A}$.

Equally to the Davies evolution, there exists a uniform family of Lindbaldians, s.t. the Schmidt conditional expectations are given as the local Lindbladian projectors of this family. For the Schmidt conditional expectation one corresponding family of Lindbladians is

$$\mathcal{L}_A^S(X) := \sum_{x \in A} E_x^S(X) - X. \tag{4.11}$$

We call them *Schmidt generators*[2] [11]. It is straightforward to see that the projection onto their kernel is given by the Schmidt conditional expectation. From the properties established above, they are uniform families of locally primitive, locally GNS-symmetric, frustration free Lindbladians. And via this definition we immediately get the additivity in the region.

4.3 Relating Davies and Schmidt dynamics

An important observation is that, since the Schmidt- and Davies- families of conditional expectations almost have the same fixed-point algebras, we can relate the relative entropy distance of any given state to the fixed point subalgebra of one to the other:

Lemma 4.7. Let $X \subset \Lambda$, $\rho \in \mathcal{D}(\mathcal{H}_\Lambda)$, then for $E_{X*}^S(\rho)$ the Schmidt conditional expectation of ρ and $E_{X*}^D(\rho)$ its Davies conditional expectation, it holds that

$$D(\rho \| E_{X*}^D(\rho)) \leq D(\rho \| E_{X*}^S(\rho)) \leq D(\rho \| E_{X\partial *}^D(\rho)). \tag{4.12}$$

Proof. First recall that for some region $X \subset \Lambda$, E_X^D is the projection onto the kernel of $\mathcal{L}_X^D$, call it $\mathcal{F}_X^D$:=Fix(E_X^D). It is also the projection onto the largest *-subalgebra of $\mathcal{B}(\mathcal{H}_{X^c}) \otimes \mathbb{1}_X \subset \mathcal{B}(\mathcal{H}_\Lambda)$ which is invariant under the modular group of the Gibbs state σ $\{\Delta_\sigma^{it}\}_{t\in\mathbb{R}}$ [11], [38]. Now E_X^S is a projection onto, say, $\mathcal{F}_X^S$:=Fix(E_X^S). This is by construction a *-subalgebra of $\mathcal{B}(\mathcal{H}_{X^c}) \otimes \mathbb{1}_X \subset \mathcal{B}(\mathcal{H}_\Lambda)$ invariant under the modular group $\{\Delta_\sigma^{it}\}_{t\in\mathbb{R}}$, see Section 4.2. Thus $\mathcal{F}_X^S \subset \mathcal{F}_X^D$. This implies that $D(\rho \| E_{X*}^D(\rho)) \leq D(\rho \| E_{X*}^S(\rho))$ for any state $\rho \in \mathcal{D}(\mathcal{H}_\Lambda)$, since

$$\begin{aligned} D(\rho \| E_{X*}^S(\rho)) - D(\rho \| E_{X*}^D(\rho)) &= \operatorname{Tr}[\rho \underbrace{(\log E_{X*}^D(\rho) - \log E_{X*}^S(\rho))}_{\in \mathcal{F}_X^D}] \\ &= \operatorname{Tr}[E_{X*}^D(\rho)(\log E_{X*}^D(\rho) - \log E_{X*}^S(\rho))] \\ &= D(E_{X*}^D(\rho) \| E_{X*}^S(\rho)) \geq 0. \end{aligned}$$

Where we used that if ω_1, ω_2 are fixed points of some conditional expectation, then so is $\log \omega_1 - \log \omega_2$. Now since $\mathcal{F}_{X\partial}^D \subset \mathcal{F}_X^S$ holds by frustration freeness, it equally follows form the calculation above that $D(\rho \| E_{X*}^S(\rho)) \leq D(\rho \| E_{X\partial *}^D(\rho))$. More explicitly $\mathcal{F}_{X\partial}^D \subset \mathcal{F}_X^S$ follows from the fact that the Gibbs state $\sigma^{X\partial}$, the unique fixed point of the Davies evolution on $X\partial$, is also a stationary under the Schmidt conditional expectation $E_{X\partial *}^S$ by construction. And hence by frustration freeness also stationary w.r.t E_{X*}^S. □

Remark. This is a crucial lemma that allows us to analyze the MLSI for Davies generators in terms of entropic inequalities associated to Schmidt, which are easier to analyze. The Schmidt generators hence serves as a proxy QMS to the Davies. Such a comparison is a well known technique for classical Markov chains.

[2] They are sometimes also referred to as *Heat bath generators w.r.t. the Schmidt conditional expectation.*

5 Main results: Gap implies cMLSI

This section contains the main result of this work, stated in Theorem 5.1. It is followed by some immediate corollaries and a discussion thereof. The rest of the section is then devoted to its proof. There in Section 5.1 we first establish an important implication of weak clustering in Theorem 5.4 and then a connection between the static (clustering) and dynamic properties of spin systems of non-zero temperature. In particular in Theorem 5.6 we establish a necessary approximate tensorization statement, which with a novel geometric argument in Section 5.2, in particular Lemma 5.8, we use to proof the main result in Section 5.3. For a graphical visualization of some of the implications that are proved in this thesis, see Figure 5.1.

Definition 7 (Geometric condition on correlation length of fixed point)**.** *For an infinite graph Λ, recall $N(l) := \sup_{x\in\Lambda} |B_l(x)|$, where $B_l(x) := \{v \in \Lambda | \operatorname{dist}(x, v) \leq l\}$ is the ball around x of radius l. We will require the correlation length ξ of the Gibbs state to satisfy*

$$\xi < \frac{l}{2\ln N(l)} \tag{5.1}$$

eventually in l, i.e. for $l \geq l_0$ for some l_0.

Remark. First note that this correlation length is the one of the Gibbs state and hence not just a function of the Graph alone, but also of the Hamiltonian and the inverse temperature.
Moreover, if Condition (5.1) holds, then $N(l)\exp\left(-\frac{l}{\xi}\right)$ is exponentially decaying in l. Recall that hypercubic lattices are sub-exponential under this definition, whereas $b-$ary trees are exponential. Hence for infinite hypercubic lattices of dimension D we have that $N(l) \propto l^D$ and hence this condition is trivially fulfilled for any $\xi > 0$. For b-ary trees we have that $N(l) = \sum_{k=0}^{l} b^k = \frac{b^{l-1}-1}{b-1} \propto b^l$, and hence this condition becomes $\xi < \frac{1}{\ln b}$.
This gives an implicit condition on the temperature β^{-1}.

We may now state the main theorem of this work.

Theorem 5.1. Let Λ be a 2-colorable graph with finite growth constant. Then the Davies generator $\mathcal{L}^D_\Lambda := \{\mathcal{L}^D_\Gamma\}_{\Gamma\subset\subset\Lambda}$ corresponding to a uniform, nearest neighbour, commuting family of Hamiltonians acting on the locally-finite dimensional quantum spin system $\mathcal{H}_\Lambda$, i.e. it satisfies a MLSI with constant

1) $\alpha(\mathcal{L}^D_\Gamma) = \Omega(1)_{|\Gamma|\to\infty}$ independent of system size, when Λ is a sub-exponential graph, or

2) $\alpha(\mathcal{L}^D_\Gamma) = \Omega\left((\ln|\Gamma|)^{-1}\right)_{|\Gamma|\to\infty}$ logarithmically decreasing in system size, when Λ is an exponential graph and the correlation length of the fixed point (Gibbs state) satisfies condition (5.1).

whenever the Lindbladian is gapped, i.e. $\inf_{\Gamma\subset\subset\Lambda}\lambda(\mathcal{L}^D_\Gamma) > 0$, or the thermal states satisfy $\mathbb{L}-\infty$-clustering.

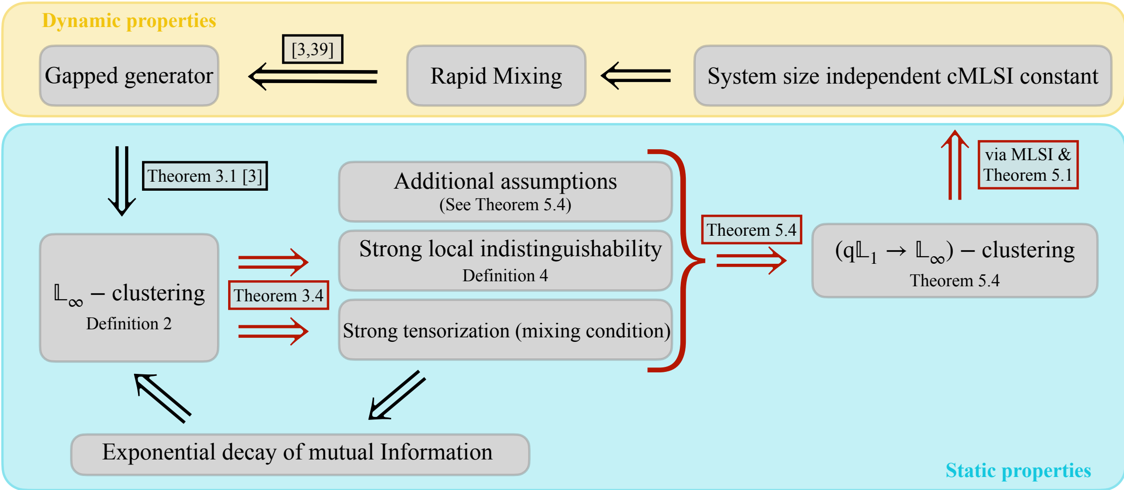

Figure 5.1: Relation between the different notions of clustering (static properties) and their relation to thermalization and gap (dynamic properties), in the setting of this thesis. The red implications signify novel implications, wheres the black ones signify results from literature. The main result of Chapter 3 is Theorem 3.4, whereas the main result of this thesis in Chapter 5 is Theorem 5.1. In effect we show that (system size invariant gap) is equivalent to a system size independent MLSI constant via the visualized chain of implications. For details on the necessary assumptions and requirements on the systems, see the respective theorems.

Remark. Recall that by Theorem 3.1 the existence of a system size independent gap implies $\mathbb{L}_\infty$-clustering. Hence the, a priori, slightly weaker assumption required for the main result to hold is is uniform exponential decay of correlations ($\mathbb{L}_\infty$-clustering) of the uniform family of thermal states.
This implies, that for the above considered systems, MLSI implies existence of a strictly positive gap of the generator. Hence we have, for these systems equivalence of strictly positive gap of generators, (weak) $\mathbb{L}_\infty$-clustering of the invariant state, MLSI with strictly positive constant, and rapid thermalization. Hence in order to establish rapid thermalization under the Davies evolution of a system under the conditions in the above theorem, it suffices to check for a strictly positive gap, or exponential decay of clustering. Hence it also partially[1] affirmatively answers an open question from [3], that for nearest neighbour commuting systems, existence of exponential decay of correlations (in the form of $\mathbb{L}_\infty$-clustering) is sufficient to prove existence of a spin-system size invariant strictly positive spectral gap for $\mathcal{L}_\Gamma^D$. This is since we have the implication that $\mathbb{L}_\infty$-clustering of a thermal state of a uniform geometrically-local, commuting Hamiltonian implies a strictly positive and system size invariant MLSI constant, which implies rapid thermalization, which in turn implies the existence of a strictly positive and spin-system invariant spectral gap of the Davies generators [3]. See also [39](Lemma 6). As seen in the remark before, the geometric condition on ξ is trivially fulfilled for a hypercubic lattice of dimension $D \in \mathbb{N}$, and becomes $\xi < \frac{1}{2\ln b}$ for the $b-$ary tree. We also recall that the correlation length of the Gibbs state is the exponential decay rate of many clustering properties, as established in Chapter 3 and Section 5.1. Hence the smaller the correlation length, the faster correlation between subsystems decay in their

[1]For nearest neighbour interactions we show that gap implies MLSI. Since in [3] MLSI implies gap is shown for uniform, geometrically-local, commuting systems. See also [39](Lemma 6).

distance and hence the closer some distinct subsystems can be s.t. their correlations are still smaller than some prescribed bound.
Note also that this result works for graphs Λ which are not two-colorable, but which can be brought into a two-colorable graph after finite size coarse-graining, i.e. by assigning a color to subsets of some finite size. See Chapter 7 for a more thorough discussion of this point, where we conjecture that this may include a very large class of systems.
For sake of readability we now consider two special cases of Theorem 5.1, namely 1D systems, $\Lambda = \mathbb{Z}$ in Theorem 5.2 and $b-$ary trees $\Lambda = \mathbb{T}_b$ in Theorem 5.3.

Corollary 5.2 (1D constant MLSI at any temperature)**.** Let $\Lambda = \mathbb{Z}$ be a 1 dimensional quantum spin chain endowed with a uniform family of bounded, geometrically$-r-$local, commuting, translation invariant Hamiltonians for some $r \in \mathbb{N}$ and let $\beta > 0$ be any inverse temperature. Then there exists a strictly positive spin-chain length independent MLSI constant $\alpha(\mathcal{L}_\Gamma^D)$ of any element of the uniform family of Davies generators $\mathcal{L}_\Lambda^D = \{\mathcal{L}_\Gamma^D\}_{\Gamma\subset\subset\Lambda}$ corresponding to this family of Hamiltonians to the inverse temperature β. I.e. $\alpha(\mathcal{L}_\Gamma^D) = \Omega(1)_{|\Gamma|\to\infty}$. Hence $\alpha := \inf_{\Gamma\subset\subset\Lambda} \alpha(\mathcal{L}_\Gamma^D) > 0$ and for any finite $\Gamma \subset\subset \Lambda$ we have

$$D(\rho_t\|\sigma^\Gamma) \leq e^{-t\alpha} D(\rho_0\|\sigma^\Gamma). \tag{5.2}$$

Remark. This follows from the fact that the Davies generators to such a uniform family of Hamiltonians are known to be gaped with a gap independent of spin-chain length. [3] Note, that here we do not require $r = 2$, i.e. nearest neighbour interaction, since by coarse-graining, we can always map a geometrically$-r-$local system to a nearest neighbour one. This only works in the 1D case, however. This result is a strict improvement over the previous best known one [12], where it was shown that $\alpha(\mathcal{L}_\Gamma^D) = \Omega(\ln|\Gamma|^{-1})$ is logarithmically decreasing in the system size. Moreover, we obtain the same scaling as of the so called LSI constant in the classical setting, which is known to be optimal.

Corollary 5.3 (b-ary Trees)**.** Let $\Lambda = \mathbb{T}_b$ be the infinite $b-$ary tree with a uniform family of bounded, nearest-neighbour, commuting Hamiltonians. Let the inverse temperature β be such that the correlation length ξ of the Gibbs state to this inverse temperature satisfies $\xi < (2\ln b)^{-1}$ and the corresponding, uniform family of Davies generators $\mathcal{L}_{\mathbb{T}_b}^D = \{\mathcal{L}_\Gamma^D\}_{\Gamma\subset\subset\mathbb{T}_b}$ have a system size independent gap, $\inf_{\Gamma\subset\subset\mathbb{T}_b} \lambda(\mathcal{L}_\Gamma^D) > 0$. Then $\inf_{\Gamma\subset\subset\mathbb{T}_b, |\Gamma|\leq l} \alpha(\mathcal{L}_\Gamma^D) \geq \alpha_l = \Omega((\log(l))^{-1})$. Hence for any finite $\Gamma \subset\subset \mathbb{T}_b$ with $|\Gamma| \leq l$ we have

$$D(\rho_t\|\sigma^\Gamma) \leq e^{-t\alpha_l} D(\rho_0\|\sigma^\Gamma) \tag{5.3}$$

and thus rapid thermalization.

Remark. This is a novel result which implies that the mixing time $t_{\text{mix}}(\epsilon) = O(\text{polylog}(|\Gamma|))$ scales poly-logarithmically with the system size, i.e. rapid thermalization holds. Classically it is known, however, that the exponential decay rate of the relative entropy towards the equilibrium is tree-size independent [35]. We suspect that this should also hold in the quantum case. Moreover, the exponential decay found in [35] also hinges upon a very analogous condition on the temperature as we require here, implicitly through $\xi < (2\ln b)^{-1}$.

In the rest of this section we will only be considering bounded, nearest-neighbour, commuting Hamiltonians.
The proof of the main result will be split into two parts, first in Theorem 5.4 showing that the clustering properties established above imply a stronger form of clustering, which are known to

be equivalent to an approximate tensorization statement for the state ω in Theorem 5.6. This and the work in [11] will already be enough to establish the main result for quantum systems on hypercubic latices of dimension $D > 1$. The second part in Section 5.2 will then consist of a geometric argument to show the main result for quantum spin chains and $b-$ary trees explicitly. The derivations generalize straightforwardly to the case of sub-exponential and exponential graphs respectively.

5.1 Approximate tensorization for an almost classical state ω via weak clustering

From this section onward we only consider graphs Λ which are two-colorable up to a finite size coarse-graining. As examples, any connected complete subtree of an infinite hypercubic lattice or infinite b-ary tree are two-colorable. See Chapter 7 for a more detailed discussion on the reason behind our restriction to two-colorable graphs. Given a two-colorable graph Λ and a fixed 2-coloring, we denote the set of vertices with labels 0 as $\Lambda_0 =: A$ and the set of vertices with labels 1 as $\Lambda_1 = \Lambda \setminus A$.

Definition 8. *Given a 2-colorable graph Λ and a quantum state $\rho \in \mathcal{D}(\mathcal{H}_\Lambda)$, we fix some 2-coloring, write $A = \Lambda_0$ and define*

$$\omega := E^S_{A*}(\rho) = (\bigcirc_{a\in\Lambda_0} E^S_{a*})(\rho), \tag{5.4}$$

where the second equality and well definiteness on the r.h.s follows from Proposition 4.4.

Next we show that in the setting which we are considering exponential decay of correlations (weak clustering) implies a stronger form of decay of correlations, namely $q\mathbb{L}_1 \to \mathbb{L}_\infty$-decay of correlations, see Theorem 5.4 for a formulation here (or see [11][Definition 8] for a definition in the general case). This is a first big result of this work which is also of independent interest since exponential decay of correlations (weak-clustering) is the type of clustering usually considered in the physics literature and much easier to prove in general. It will also be an important step in the proof of the main result.

Remark. This $q\mathbb{L}_1 \to \mathbb{L}_\infty$-decay of correlations was used and required to prove the main result of [11](Section 4). As a simple corollary of the next theorem, we can therefore establish the main result of [11] under the seemingly weaker assumption of gaped Lindbladians.

Theorem 5.4 ($\mathbb{L}_\infty$-clustering is equivalent to $q\mathbb{L}_1 \to \mathbb{L}_\infty$-clustering)**.** Let Λ be a 2-colorable graph with finite growth constant and $\Gamma \subset\subset \Lambda$ a finite subgraph. Let H_Λ be a uniformly bounded, commuting, nearest-neighbour Hamiltonian and let $\{E^S_X\}_{X\subset\subset\Lambda}$ be the Schmidt conditional expectation Then $\mathbb{L}_\infty$-clustering implies $q\mathbb{L}_1 \to \mathbb{L}_\infty$-clustering, that is for any overlapping subregions $C, D \subset \Gamma$ it holds that

$$\max_{\alpha=\{\alpha_i\}_{i\in I_{C\cup D}}} \|E^{S,(\alpha)}_C \circ E^{S,(\alpha)}_D - E^{S,(\alpha)}_{C\cup D} : \mathbb{L}_1(\tau^\alpha_{C\cup D}) \to \mathbb{L}_\infty\| \leq |C\cup D|\epsilon(l),$$

holds for an exponentially decaying function $\epsilon(l)$, where $l = \mathrm{dist}(C\setminus D, D\setminus C)$.

Remark. An example of subsets C, D of a 3-ary tree can be found in Figure 5.2. With this statement we effectively show that gap, which implies $\mathbb{L}_\infty$-clustering, implies this seemingly stronger notion of clustering, called $q\mathbb{L}_1 \to \mathbb{L}_\infty$-clustering. Before giving its proof, we discuss two corollaries.

Corollary 5.5 (Strengthening of Main Result of [11] under weaker assumption)**.** [2] The main result of [11] holds for the Schmidt and Davies conditional expectations[3] under the weaker assumption of a gaped Davies Lindbladian[4]. Informally, nearest neighbour quantum systems with a uniformly bounded, commuting Hamiltonian on hypercubic lattices of dimension $D \geq 2$, which evolve under Davies evolution, satisfy strong exponential convergence in relative entropy towards equilibrium. I.e. they thermalize rapidly under Davies evolution whenever the Davies Lindbladian is gaped.

The statement of this corollary is essentially the main theorem in case of hypercubic lattices of dimension $D > 1$. Hence we will focus in the rest of this work on the 1D and b-ary tree settings. For a thorough discussion of this result, see [11]. Another corollary of Theorem 5.4 is the following approximate tensorization statement for the state ω, which we will require to prove our main result.

Theorem 5.6 (Approximate tensorization for ω)**.** Let Λ be an infinite, 2-colorable graph of finite growth constant. Let H be a nearest neighbour, bounded, commuting Hamiltonian on a family of finite subgraphs $\{\Gamma\}_{\Gamma\subset\Lambda}$ which, on any given Γ, satisfies uniform $\mathbb{L}_\infty$-clustering at inverse temperature β with correlation length ξ. Let $C, D, R \subset \Gamma$ be three connected subsets, s.t. $R := C \cup D, \partial C, \partial D, \partial R \cap \Lambda_0 = \emptyset$, and let $l := \operatorname{dist}(C \setminus D, D \setminus C) > 1$. Fix some state $\rho \in \mathcal{D}(\mathcal{H}_\Gamma)$ and define ω w.r.t it as above. Then

$$D(\omega \| E^S_{C\cup D*}(\omega)) \leq \frac{1}{1-2\epsilon(l)}[D(\omega\|E^S_{C*}(\omega)) + D(\omega\|E^S_{D*}(\omega))], \tag{5.5}$$

where $\epsilon(l) = \exp\left(O(\beta)\right)O(|C|, |D|)O\left(\exp\left(-\frac{l}{\xi}\right)\right)$.

The proof of this is, given Theorem 5.4, just an application of [11][Theorem 8]. This theorem is repeated below for convenience.

Theorem 5.7 ([11][Theorem 8)**.**] If there exists an exponentially decaying function $l \mapsto \epsilon(l) \leq \frac{1}{2}$, s.t.

$$\max_{\alpha=\{\alpha_i\}_{i\in I_{C\cup D}}} \|E^{S,(\alpha)}_C \circ E^{S,(\alpha)}_D - E^{S,(\alpha)}_{C\cup D} : \mathbb{L}_{1,\tau^\alpha_{C\cup D}} \to \mathbb{L}_\infty\| \leq |C \cup D|\epsilon(l),$$

then

$$D(\omega \| E^S_{C\cup D*}(\omega)) \leq \frac{1}{1-2\epsilon(l)}[D(\omega\|E^S_{C*}(\omega)) + D(\omega\|E^S_{D*}(\omega))].$$

Here $\tau^{(\alpha)}_{C\cup D}$, with $(\alpha) = \{\alpha_i\}_{i\in I_{CD}}$ and is $E^{S,(\alpha)}_X$ is just as defined in Section 4.2.

Proof of Theorem 5.4. We show that, under the assumption of the theorem, $\mathbb{L}_\infty$-clustering implies

$$\max_{(\alpha)} \|E^{S,(\alpha)}_C \circ E^{S,(\alpha)}_D - E^{S,(\alpha)}_{C\cup D} : \mathbb{L}_{1,\tau^{C\cup D}_\alpha} \to \mathbb{L}_\infty\| \leq |C \cup D|\epsilon(l)$$

explicitly for the Schmidt conditional expectations. Here (α) is a boundary condition of the subset $CD := C \cup D \subset \Lambda$ and $\epsilon(l)$ is exponentially decaying with decay length ξ. We will do this in two steps, first establish that what we need to show is algebraically equivalent to a statement $\sigma_1 \overset{\epsilon(l)}{\sim} \sigma_2$

[2] The Main result of [11] concerns only the Schmidt conditional expectation, but with the geometric argument we give in Section 5.2 we can extend this to the Davies as well.

[3] For the Davies with the geometric argument in Section 5.2, for the Schmidt this is immediate from [11].

[4] Or the condition of (q$\mathbb{L}_\infty$-clustering), which is slightly stronger than the gap condition

for two states σ_1, σ_2 and then employ results of Theorem 3.4 to show this statement from the assumptions of the theorem.
Before we continue let us first establish some nomenclature for the subregions we are considering. We split the region CD into the following disjoint subsets $\tilde{l}_{\text{out}}\tilde{l}_{\text{in}}El_{\text{out}}l_{\text{in}}Fr_{\text{in}}r_{\text{out}}G\tilde{r}_{\text{in}}\tilde{r}_{\text{out}}$, where $C = ElF, C_{\text{in}} = \tilde{l}_{\text{in}}ElFr_{\text{in}}, D = FrG, D_{\text{in}} = l_{\text{in}}FrG\tilde{r}_{\text{in}}$. The projectors $P^{(\alpha_{\tilde{l}})}, P^{(\gamma_l)}, P^{(\beta_r)}, P^{(\alpha_{\tilde{r}})}$ act on the regions $\tilde{l}, l, r, \tilde{r}$ respectively. For a graphical representation of this see Figure 5.2.

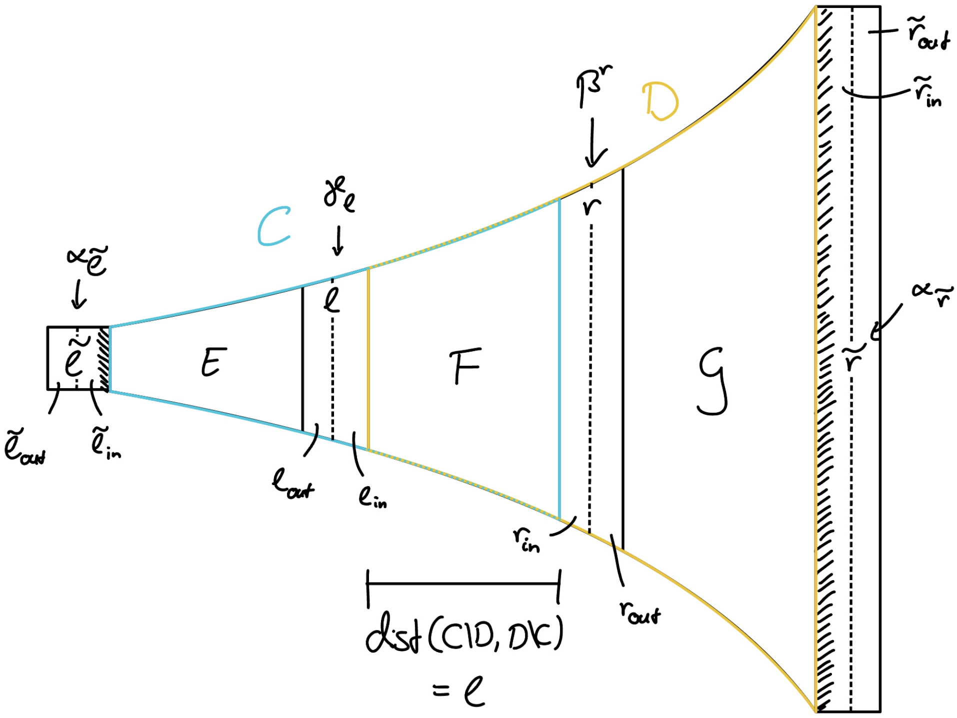

Figure 5.2: Partition of a subregion CD of a tree into two overlapping subregions C and D. We have $E := C\backslash(D\partial), F := C\cap D, G := D\backslash(C\partial)$. The splitting of the boundary Hilbert spaces corresponding to a boundary site $\{b_i\}$ in the boundary of a region $A \in \{E, F, G\}$ into $\bigotimes_{i\in I_A} P^{\alpha_i}\mathcal{H}_{b_i} = \mathcal{H}^{\alpha_i}_{\partial_{\text{in}}A} \otimes \mathcal{H}^{\alpha_i}_{\partial_{\text{out}}A} = \left(\mathcal{H}^{\alpha_i}_0 \otimes \mathcal{H}^{\alpha_i}_c\right) \otimes \left(\bigotimes_{j\in J^{(i)}\backslash\{0\}} \mathcal{H}^{\alpha_i}_j\right)$ is represented by a dotted line. Hence e.g. the Hilbert space of the region $\tilde{l}_{\text{in}}$ is $\mathcal{H}^{(\alpha_{\tilde{l}})}_{\partial^{\text{left}}_{\text{in}}E}$ and of $\tilde{l}_{\text{out}}$ it is $\mathcal{H}^{(\alpha_{\tilde{l}})}_{\partial^{\text{left}}_{\text{out}}E}$. Here the superskript *left*, refers to the part of $\partial_{\text{in}}E$, which is located in the geometric region $\tilde{l}$, i.e. the in the figure left boundary of the region E. Respectively, this is the same with the other boundaries. We fix the boundary conditions $(\alpha_{\tilde{l}})$ on region $\tilde{l}$, (γ_l) on l, (β_r) on r, and $(\alpha_{\tilde{r}})$ on $\tilde{r}$.

Recall the notation for the Schmidt conditional expectation established in Section 4.2 and equation (4.10). Fix some boundary condition $(\alpha) = (\alpha_{\tilde{l}}, \alpha_{\tilde{r}}) = \{\alpha_i\}_{i\in I_{CD}}$ for CD, where $(\alpha_{\tilde{l}}) := \{\alpha_i\}_{i\in I_{\partial C\backslash D}}$ labels the boundary of C not in D and $(\alpha_{\tilde{r}}) := \{\alpha_i\}_{i\in I_{\partial D\backslash C}}$ labels the boundary of D not in C. Similarly denote with $(\beta_r) := \{\beta_i\}_{i\in I_{\partial C\cap D}}$ and $(\gamma_l) := \{\gamma_i\}_{i\in I_{\partial D\cap D}}$ the boundaries of C in D and of D in C respectively. For visualization, see also Figure 5.2. Let

$0 \leq X \equiv X^{(\alpha)} \in \mathbb{L}_{1,\tau^{(\alpha)}_{CD_{\text{in}}}}$ s.t. $\|X\|_{1,\tau^{(\alpha)}_{CD_{\text{in}}}} = 1$. Set

$$N \equiv N^{(\alpha_{\tilde{r}})} := E_D^{S,(\alpha_{\tilde{r}})}(X) = \bigoplus_{(\gamma_l)} E_D^{S,(\gamma_l,\alpha_{\tilde{r}})}(X) \in \mathcal{B}(\mathcal{H}_{D^{(\alpha_{\tilde{r}})}_{\text{out}}}) \otimes \mathbb{1}_{D^{(\alpha_{\tilde{r}})}_{\text{in}}}.$$

By construction it holds that $P^{(\alpha_{\tilde{r}})} N P^{(\alpha_{\tilde{r}})} = N$. Recall that since $D \subset CD$, it follows that $E^S_{CD} \circ E^S_C = E^S_{CD}$ and hence

$$\begin{aligned}
&\left(E_C^{S,(\alpha)} \circ E_D^{S,(\alpha)} - E_{CD}^{S,(\alpha)}\right)(X) \\
&\quad = \left(\bigoplus_{(\beta_r)} E_C^{S,(\alpha_{\tilde{l}},\beta_r)} - E_{CD}^{S,(\alpha_{\tilde{l}},\alpha_{\tilde{r}})}\right)(N) \\
&\quad = \bigoplus_{(\beta_r)} \left(P^{(\alpha_{\tilde{l}})} P^{(\beta_r)} \left(\operatorname{tr}_{C_{\text{in}},(\alpha_{\tilde{l}},\beta_r)}\left[\tau_{C_{\text{in}}}^{(\alpha_{\tilde{l}},\beta_r)} N\right] \otimes \mathbb{1}_{C_{\text{in}}}\right) P^{(\alpha_{\tilde{l}})} P^{(\beta_r)}\right) \\
&\quad - P^{(\alpha_{\tilde{l}})} P^{(\alpha_{\tilde{r}})} \left(\operatorname{tr}_{CD_{\text{in}},(\alpha_{\tilde{l}},\alpha_{\tilde{r}})}\left[\tau_{CD_{\text{in}}}^{(\alpha_{\tilde{l}},\alpha_{\tilde{r}})} N\right] \otimes \mathbb{1}_{CD_{\text{in}}}\right) P^{(\alpha_{\tilde{l}})} P^{(\alpha_{\tilde{r}})} \\
&\quad = P^{(\alpha_{\tilde{l}})} \left[\bigoplus_{(\beta_r)} P^{(\beta_r)} \left(\operatorname{tr}_{C_{\text{in}},(\alpha_{\tilde{l}},\beta_r)}\left[\tau_{C_{\text{in}}}^{(\alpha_{\tilde{l}},\beta_r)} P^{(\alpha_{\tilde{r}})} (N \otimes \mathbb{1}_{D_{\text{in}}}) P^{(\alpha_{\tilde{r}})}\right] \otimes \mathbb{1}_{C_{\text{in}}}\right) P^{(\beta_r)}\right. \\
&\quad - \left(\bigoplus_{(\beta_r)} P^{(\beta_r)}\right) P^{(\alpha_{\tilde{r}})} \left(\operatorname{tr}_{CD_{\text{in}},(\alpha_{\tilde{l}},\alpha_{\tilde{r}})}\left[\tau_{CD_{\text{in}}}^{(\alpha_{\tilde{l}},\alpha_{\tilde{r}})} (N \otimes \mathbb{1}_{D_{\text{in}}})\right] \otimes \mathbb{1}_{CD_{\text{in}}}\right) P^{(\alpha_{\tilde{r}})} P^{(\alpha_{\tilde{l}})}
\end{aligned}$$

$$\begin{aligned}
&= \left(P^{(\alpha_{\tilde{l}})} P^{(\alpha_{\tilde{r}})} \bigoplus_{(\beta_r)} P^{(\beta_r)}\right) \operatorname{tr}_{C_{\text{in}} \setminus D_{\text{in}},(\alpha_{\tilde{l}})}\left[\left(\underbrace{\operatorname{tr}_{C_{\text{in}} \cap D_{\text{in}},(\beta_r)}\left[\tau_{C_{\text{in}}}^{(\alpha_{\tilde{l}},\beta_r)}\right]}_{=:\sigma_2^{(\alpha_{\tilde{l}},\beta_r)} \equiv \sigma_2} - \underbrace{\operatorname{tr}_{D_{\text{in}},(\alpha_{\tilde{r}})}\left[\tau_{CD_{\text{in}}}^{(\alpha_{\tilde{l}},\alpha_{\tilde{r}})}\right]}_{=:\sigma_1^{\alpha_{\tilde{l}}} \equiv \sigma_1}\right) N\right] \times \\
&\left(P^{(\alpha_{\tilde{l}})} P^{(\alpha_{\tilde{r}})} \bigoplus_{(\beta_r)} P^{(\beta_r)}\right) \equiv \bigoplus_{(\beta_r)} P^{(\alpha_{\tilde{l}},\beta_r,\alpha_{\tilde{r}})} \left[\operatorname{tr}_{C_{\text{in}} \setminus D_{\text{in}}}\left[(\sigma_2 - \sigma_1) N\right]\right] \bigoplus_{(\beta_r)} P^{(\alpha_{\tilde{l}},\beta_r,\alpha_{\tilde{r}})},
\end{aligned}$$

where $\sigma_2 \equiv \sigma_2^{(\alpha_{\tilde{l}},\beta_r)} := \operatorname{tr}_{C_{\text{in}} \cap D_{\text{in}},(\beta_r)}\left[\tau_{C_{\text{in}}}^{(\alpha_{\tilde{l}},\beta_r)}\right]$ and $\sigma_1 \equiv \sigma_1^{(\alpha_{\tilde{l}})} := \operatorname{tr}_{D_{\text{in}},(\alpha_{\tilde{r}})}\left[\tau_{CD_{\text{in}}}^{(\alpha_{\tilde{l}},\alpha_{\tilde{r}})}\right]$. In the first line here we used the definition of X, in the second the explicit expressions for the Schmidt conditional expectation from Section 4.2. Then in the third we factor out a common $P^{\alpha_{\tilde{l}}}$ and introduce an identity $\mathbb{1} = \bigoplus_{(\beta_r)} P^{(\beta_r)}$, which commutes with the term just after. We also employ the fact that $N \equiv P^{\alpha_r} N \otimes \mathbb{1}_{D_{\text{in}}} P^{\alpha_r}$. In the forth equality we factor out the projections and rearrange the partial traces suitably. The last equality is then just introducing a simplifying notation. For simplicity we suppress the boundary conditions $(\alpha_{\tilde{l}})$ index on the states. Hence

$$\begin{aligned}
\| \left(E_C^{S(\alpha)} \circ E_D^{S(\alpha)} - E_{CD}^{S(\alpha)}\right)(X)\| &\leq \max_{(\beta_r)} \| \operatorname{tr}_{C_{\text{in}}\setminus D_{\text{in}},(\alpha_{\tilde{l}})} [(\sigma_2 - \sigma_1)N] \| \\
&= \max_{(\beta_r)} \left|\operatorname{Tr}_{C_{\text{in}}\setminus D_{\text{in}},(\alpha_{\tilde{l}})} [(\sigma_2 - \sigma_1)N]\right| \\
&= \max_{(\beta_r)} | \operatorname{Tr}_{C_{\text{in}}\setminus D_{\text{in}},(\alpha_{\tilde{l}})} [(\mathbb{1} - \sigma_1^{-\frac{1}{2}}\sigma_2\sigma_1^{-\frac{1}{2}})(\sigma_1^{\frac{1}{2}} N \sigma_1^{\frac{1}{2}})]| \\
&\leq \max_{(\beta_r)} \|(\mathbb{1} - \sigma_1^{-\frac{1}{2}}\sigma_2\sigma_1^{-\frac{1}{2}})(\sigma_1^{\frac{1}{2}} N \sigma_1^{\frac{1}{2}})]\|_1 \\
&\leq \max_{(\beta_r)} \|\sigma_1^{-\frac{1}{2}}\sigma_2\sigma_1^{-\frac{1}{2}} - \mathbb{1}\|_\infty \|N\|_{\mathbb{L}_{1,\sigma_1}},
\end{aligned}$$

where the equality in second line follows since $N = E_D^{S(\gamma_l,\alpha_{\tilde{r}})}(X)$ is the identity on the complementary Hilbert space to $(\mathcal{H}^{(\alpha_{\tilde{l}})}_{C_{\text{in}}\setminus D_{\text{in}}}$ The last inequality we used Hölder and the definition of the $\mathbb{L}_{1,\sigma_1}$ norm. By definition of N, we have

$$\begin{aligned}
\|N\|_{\mathbb{L}_{1,\sigma_1}} &= \|\sigma_1^{\frac{1}{2}} N \sigma_1^{\frac{1}{2}}\|_1 \overset{N\geq 0}{=} \operatorname{Tr}[\sigma_1 N] = \operatorname{Tr}[\tau_{CD_{\text{in}}}^{(\alpha_{\tilde{l}},\alpha_{\tilde{r}})} N] = \operatorname{Tr}[\tau_{CD_{\text{in}}}^{(\alpha)} E_D^{S(\alpha)}(X)] \\
&= \operatorname{Tr}[E_{D*}^{S(\alpha)}(\tau_{CD_{\text{in}}}^{(\alpha)})X] = \operatorname{Tr}[\tau_{CD_{\text{in}}}^{(\alpha)} X] \overset{X\geq 0}{=} \|X\|_{\mathbb{L}_{1,\tau_{CD_{\text{in}}}^{(\alpha)}}} = 1.
\end{aligned}$$

Hence to prove the theorem, we need to establish that $\sigma_1 \overset{\epsilon}{\sim} \sigma_2$ for any boundary condition $(\alpha_{\tilde{l}}, \beta_r, \alpha_{\tilde{r}})$. We will do with with $l \mapsto \epsilon(l)$ an exponentially decreasing function in $l = \operatorname{dist}(C \setminus D, D \setminus C)$ with decay length ξ.
Thus the two states can be written as

$$\begin{aligned}
\sigma_1 &:= \operatorname{tr}_{D_{\text{in}},(\alpha_{\tilde{r}})}\left[\tau_{CD_{\text{in}}}^{(\alpha_{\tilde{l}},\alpha_{\tilde{r}})}\right] = \frac{1}{\operatorname{Tr}[\dots]} \operatorname{tr}_{\partial_{\text{out}}(CD)\cup D_{\text{in}},(\alpha_{\tilde{r}})}\left[P^{(\alpha_{\tilde{l}})} P^{(\alpha_{\tilde{r}})} \sigma^{(CD\partial)} P^{(\alpha_{\tilde{l}})} P^{(\alpha_{\tilde{r}})}\right] \\
&= \frac{1}{\operatorname{Tr}[\dots]} \operatorname{tr}_{\tilde{l}_{\text{out}} l_{\text{in}} D \tilde{r}}\left[P^{(\alpha_{\tilde{r}})} P^{(\alpha_{\tilde{l}})} \sigma^{(CD\partial)} P^{(\alpha_{\tilde{l}})}\right] \in \mathcal{B}(\mathcal{H}^{(\alpha_{\tilde{l}})}_{C_{\text{in}}\setminus D_{\text{in}}}) \\
\sigma_2 &:= \operatorname{tr}_{F_{\text{in}},(\beta_r)}\left[\tau_{C_{\text{in}}}^{(\alpha_{\tilde{l}},\beta_r)}\right] = \frac{1}{\operatorname{Tr}[\dots]} \operatorname{tr}_{\partial_{\text{out}}(C)\cup F_{\text{in}},(\beta_r)}\left[P^{(\alpha_{\tilde{l}})} P^{(\beta_r)} \sigma^{(C\partial)} P^{(\alpha_{\tilde{l}})} P^{(\beta_r)}\right] \\
&= \frac{1}{\operatorname{Tr}[\dots]} \operatorname{tr}_{\tilde{l}_{\text{out}} l_{\text{in}} F r}\left[P^{(\beta_r)} P^{(\alpha_{\tilde{l}})} \sigma^{(C\partial)} P^{(\alpha_{\tilde{l}})}\right] \in \mathcal{B}(\mathcal{H}^{(\alpha_{\tilde{l}})}_{C_{\text{in}}\setminus D_{\text{in}}})
\end{aligned}$$

Note that both are fullrank states on $\mathcal{H}^{(\alpha_{\tilde{l}})}_{C_{\text{in}}\setminus D_{\text{in}}}$ and thus have the same support. The intuition now is, since we have a Gibbs state that satisfies exponential decay of correlations, local indistinguishability, and the mixing condition tensorization (see Theorem 3.4) these two states should be approximately the same in the bulk, where we compare them. This is since they only differ on $D \setminus C$, but we look at them on $C \setminus D$. From the assumptions of the Theorem we have that the family of Gibbs states satisfy $\mathbb{L}_\infty$-clustering and hence by Theorem 3.4 also strong local indistinguishability and strong tensorization (mixing condition). We start by applying the strong tensorization property (3.6) to each of the states, which gives the existence of two exponentially decaying functions ϵ_1, ϵ_2 in $\operatorname{dist}(l, \tilde{r}) > \operatorname{dist}(C \setminus D, D \setminus C)$ and $\operatorname{dist}(l, r) = \operatorname{dist}(C \setminus D, D \setminus C)$, respectively, s.t.

$$\begin{aligned}
&\operatorname{tr}_{l_{\text{in}} D} \sigma^{CD\partial} \overset{\epsilon_1}{\sim} \operatorname{tr}_{l_{\text{in}} D \tilde{r}} \sigma^{CD\partial} \otimes \operatorname{tr}_{\tilde{l} E l D} \sigma^{CD\partial} \\
&\overset{\text{Theorem 3.2}}{\Longrightarrow} \operatorname{tr}_{\tilde{r}} P^{(\alpha_{\tilde{r}})} \operatorname{tr}_{l_{\text{in}} D} \sigma^{CD\partial} P^{(\alpha_{\tilde{r}})} \overset{\epsilon_1}{\sim} \operatorname{tr}_{l_{\text{in}} D \tilde{r}} \sigma^{CD\partial} \operatorname{Tr}[P^{(\alpha_{\tilde{r}})} \sigma^{CD\partial} P^{(\alpha_{\tilde{r}})}], \\
&\operatorname{tr}_{l_{\text{in}} F} \sigma^{C\partial} \overset{\epsilon_2}{\sim} \operatorname{tr}_{l_{\text{in}} F r} \sigma^{C\partial} \otimes \operatorname{tr}_{\tilde{l} E l F} \sigma^{C\partial} \\
&\overset{\text{Theorem 3.2}}{\Longrightarrow} \operatorname{tr}_r P^{(\beta_r)} \operatorname{tr}_{l_{\text{in}} F} \sigma^{C\partial} P^{(\beta_r)} \overset{\epsilon_2}{\sim} \operatorname{tr}_{l_{\text{in}} F r} \sigma^{C\partial} \operatorname{Tr}[P^{(\beta_r)} \sigma^{C\partial} P^{(\beta_r)}].
\end{aligned}$$

Here the implication follows from Proposition 3.2 4) applied to the projections $P^{(\alpha_{\tilde{r}})}, P^{(\beta_r)}$ respectively. Now by strong local indistinguishability (Theorem 3.4) there exists an exponentially decaying function ϵ_3 in $\mathrm{dist}(l,r) = \mathrm{dist}(C \setminus D, D \setminus C)$, s.t.

$$\mathrm{tr}_{l_{\mathrm{in}}D\tilde{r}}\, \sigma^{CD\partial} = \mathrm{tr}_{l_{\mathrm{in}}FrG\tilde{r}}\, \sigma^{CD\partial} \overset{\epsilon_3}{\sim} \mathrm{tr}_{l_{\mathrm{in}}Fr}\, \sigma^{C\partial}.$$

Hence by transitivity and symmetry of the strong similarity relation, see Theorem 3.2 1),2), it follows that

$$\begin{aligned}\mathrm{tr}_{l_{\mathrm{in}}D\tilde{r}}\left[P^{(\alpha_{\tilde{r}})}\sigma^{CD\partial}P^{(\alpha_{\tilde{r}})}\right] &\overset{\epsilon_1}{\sim} \mathrm{tr}_{l_{\mathrm{in}}D\tilde{r}}\, \sigma^{CD\partial}\, \mathrm{Tr}[P^{(\alpha_{\tilde{r}})}\sigma^{CD\partial}P^{(\alpha_{\tilde{r}})}] \overset{\epsilon_3}{\sim} \mathrm{tr}_{l_{\mathrm{in}}Fr}\, \sigma^{C\partial}\, \mathrm{Tr}[P^{(\alpha_{\tilde{r}})}\sigma^{CD\partial}P^{(\alpha_{\tilde{r}})}] \\ &\overset{\epsilon_2(1-\epsilon_2)^{-1}}{\sim} \mathrm{tr}_{l_{\mathrm{in}}Fr}\left[P^{(\beta_r)}\sigma^{C\partial}P^{(\beta_r)}\right] \frac{\mathrm{Tr}[P^{(\alpha_{\tilde{r}})}\sigma^{CD\partial}P^{(\alpha_{\tilde{r}})}]}{\mathrm{Tr}[P^{(\beta_r)}\sigma^{C\partial}P^{(\beta_r)}]},\end{aligned}$$

hence

$$\widetilde{\sigma_1} := \frac{\mathrm{tr}_{l_{\mathrm{in}}D\tilde{r}}\left[P^{(\alpha_{\tilde{r}})}\sigma^{CD\partial}P^{(\alpha_{\tilde{r}})}\right]}{\mathrm{Tr}[P^{(\alpha_{\tilde{r}})}\sigma^{CD\partial}P^{(\alpha_{\tilde{r}})}]} \overset{\eta}{\sim} \frac{\mathrm{tr}_{l_{\mathrm{in}}Fr}\left[P^{(\beta_r)}\sigma^{C\partial}P^{(\beta_r)}\right]}{\mathrm{Tr}[P^{(\beta_r)}\sigma^{C\partial}P^{(\beta_r)}]} =: \widetilde{\sigma_2},$$

where η is exponentially decreasing in $\mathrm{dist}(C \setminus D, D \setminus C)$ with decay length ξ, since all $\epsilon_1, \epsilon_2, \epsilon_3$ are. Now by Proposition 3.2 4′) it follows that this strong similarity of the states obviously also holds in each block $(\alpha_{\tilde{l}})$:

$$\sigma_1 = \frac{1}{\mathrm{Tr}[\ldots]}\, \mathrm{tr}_{\tilde{l}_{\mathrm{in}}}\, [P^{\alpha_{\tilde{l}}}\widetilde{\sigma_1}P^{\alpha_{\tilde{l}}}] \overset{\epsilon:=\eta(2+\eta)}{\sim} \frac{1}{\mathrm{Tr}[\ldots]}\, \mathrm{tr}_{\tilde{l}_{\mathrm{in}}}\, [P^{\alpha_{\tilde{l}}}\widetilde{\sigma_2}P^{\alpha_{\tilde{l}}}] = \sigma_2.$$

This establishes the bound for any boundary condition $(\alpha_{\tilde{l}}, \beta_r, \alpha_{\tilde{r}})$, which concludes the proof. Note that this also shows that the exponential decay rate of ϵ is the same as the one of the $\mathbb{L}_\infty$-clustering we assumed, hence the correlation length of the Gibbs state ξ. □

5.2 Geometric argument

In this section we will provide the final geometric part of the proof of Theorem 5.1 explicitly for $b-$ary trees with $b \in \mathbb{N}$. If $b = 1$ this will prove Corollary 5.2, since every finite complete connected subset of a $1-$ary tree is a finite length subset of a 1-dimensional quantum spin chain. If $b > 1$, this will prove Corollary 5.3. The idea of the proof will be essentially the same for the two quantitatively very different cases of exponential and sub-exponential graphs. Denote the infinite complete $b-$ary tree with $\mathbb{T}_b$. Denote with $B_{x,l}$ the subtree rooted at site $x \in \mathbb{T}_b$ of height l. For a finite subtree $\Gamma \subset\subset \mathbb{T}_b$ of height L, say $\Gamma = B_{0,L}$, let Γ_0 be all its vertices of index 0 under some fixed 2-coloring. We define the following set of subsets

$$\{R_k\}_{k\in K} := \{B_{x_k,l_0} \cap \Gamma\}_{k\in K} = \{B_{x,l_0} \cap \Gamma\}_{x\in\Gamma_0}, \tag{5.6}$$

where $K := \{k | x_k \in \Gamma_0\}$ and where we fix l_0 later on. We can think of l_0 as a suitably large constant. These will form a suitable coarse-graining into subtrees on each vertex of the same label (i.e. in Γ_0) of finite fixed height l_0. We will be considering the cMLSI constant of our evolution on these sets end extend these via a suitable Cesaro averaging argument to the whole lattice. For a set $B_{x_k,L}$ define

$$C_k^{\tilde{l}} := B_{x_k,\tilde{l}} \qquad D_k^{\tilde{l},l} := \bigcup_{\substack{m\in K \\ \mathrm{dist}(x_m,x_k=\tilde{l}-l)}} B_{x_m,L+l-\tilde{l}}. \tag{5.7}$$

Hence we cover the subtree $B_{x_k,L}$ of height L by a subtree of height $\tilde{l}$, called $C_k^{\tilde{l}}$, and a union of disjoint subtrees of height $L + l - \tilde{l}$, called $D_k^{\tilde{l},l}$, s.t. their overlap has height l and is hence s.t. we can apply the approximate tensorization result Theorem 5.6 with the function $\epsilon(l)$. Importantly we require that each of these sets *begins* and *ends* with some vertices of the same index 0, i.e. in Γ_0. In formulae this is $C_k^{\tilde{l}} \cup D_k^{\tilde{l},l} = B_{x_k,L}$ and $\operatorname{dist}(C_k^{\tilde{l}} \setminus D_k^{\tilde{l},l}, D_k^{\tilde{l},l} \setminus C_k^{\tilde{l}}) = l$ for all $0 \leq \tilde{l} \leq L$. Hence each set $B_{x_k,L}$ has the family of non-trivial partitions $\{C_k^{\tilde{l}}, D_k^{\tilde{l},l}\}_{\tilde{l}=1}^{L-1}$ and for each of these it holds, due to Theorem 5.6, that

$$D(\omega \| E^S_{B_{x_k,L}*}(\omega)) \leq \frac{1}{1 - 2\epsilon(l)} \left[D(\omega \| E^S_{C_k^{\tilde{l}}*}(\omega)) + D(\omega \| E^S_{D_k^{\tilde{l},l}*}(\omega)) \right],$$

where $\epsilon(l) = K|B_{x_k,L}| \exp(-\frac{l}{\xi})$. For an example of these regions see Figure 5.3.

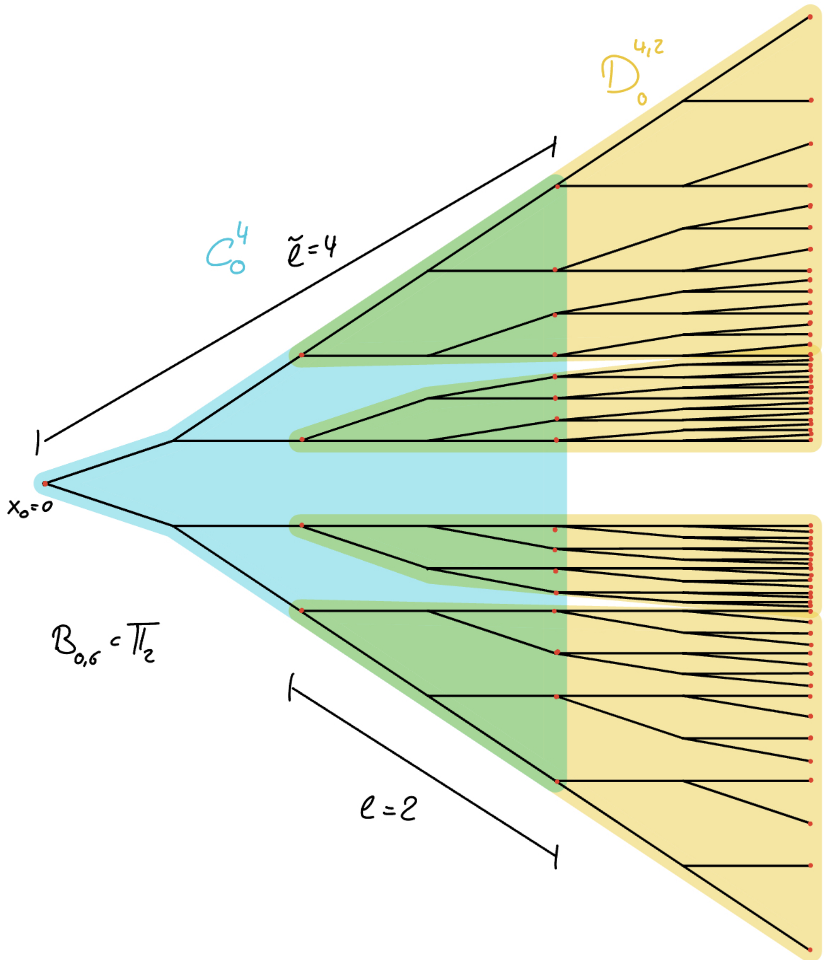

Figure 5.3: Example of a partition of $B_{0,6} \subset \mathbb{T}_2$ into the regions $C_0^{\tilde{l}}$, $D_0^{\tilde{l},l}$ with height of the regions $\tilde{l} = 4$ and $l = 6 - \tilde{l} + 2 = 4$ and the height of their overlap being $l = 2$. This is in $B_{0,6}$ as part of a 2-ary tree. The red vertices are the ones of index 0. Notice that each of these sets 'begins' and 'ends' in these sets.

We will now consider the two cases of subexponential (1D) and exponential (b-ary, $b > 1$) separately.

Case 1) The 1D quantum spin chain, i.e. $b = 1$ case where $|B_{x_k,L}| = L$.

Case 2) The $b-$ary tree case with $b > 1$, where $|B_{x_k,L}| = \frac{b^{L+1}-1}{b-1} = O(b^L)$.

Remark: Note that the analysis in the case of any sub-exponential graph will work just in case 1), however, the sets C_k, D_k will obviously be different. Similarly, the analysis of any exponential graph will work just as in case 2).

Case 1) We can pick the overlap $l = \lfloor\sqrt{L}\rfloor = O(\sqrt{L})$. Then $\epsilon(l) \propto L\exp(-\frac{\sqrt{L}}{\xi})$ is exponentially decaying in L for any $\xi > 0$. Set $l_{\min,1}$ to be the smallest L, s.t. $\epsilon(l) < \frac{1}{2}$ for all $l \geq l_{\min,1}$.

Case 2) We have to pick the overlap $l = \frac{L}{N} = O(L)$ for some $N \in \mathbb{N} \setminus \{1\}$ and require the correlation length to satisfy $\xi < \frac{1}{N\ln b}$. This suffices for $N = 2$, but we will keep it general in the following derivation. Then $\epsilon(l) = |B_{x_k,L}|\exp(-\frac{L}{N\xi}) = b^L e^{-\frac{L}{N\xi}}$ is eventually exponentially decaying in L and we set l_0 to be the smallest l, s.t. $\epsilon(l) < \frac{1}{2}$ for all $l \geq l_0$.

We define

$$D_R(\omega) := \sum_{R_k \subset R} D(\omega \| E^S_{R_k*}(\omega)), \tag{5.8}$$

where the sets $\{R_k\}_{k\in K}$ are as in Equation 5.6 defined above. Observe that $D_R(\omega)$ is monotonically increasing in R, i.e. if $A \subset B \subset \Gamma$ are two subregions, then $D_A(\omega) \leq D_B(\omega)$ by positivity of the relative entropy, and additive up to boundary terms, i.e. if $A, B \subset \Gamma$ are two disjoint subregions, then $D_{AB}(\omega) = D_A(\omega) + D_B(\omega) + \sum_{\substack{R_k\cap B\neq\emptyset \\ R_k\cap A\neq\emptyset}} D(\omega\|E^S_{R_k*}(\omega))$.
Next we define a function $L \mapsto C(L) : \mathbb{N} \to \mathbb{R}$, s.t. $C(L)$ is the smallest number, s.t. $\forall j \in \Gamma_0$

$$D(\omega\|E^S_{B_{x_j,L}*}(\omega)) \leq C(L) D_{B_{x_j,L}}(\omega) \tag{5.9}$$

This can always be done and is, trivially, monotonically non-decreasing. We will show, by inductively applying the approximate tensorization and averaging suitably, that the following Lemma holds.

Lemma 5.8. For sets of the form $R = B_{x_j,L}$ for some $j \in K$, (5.9) holds with

Case 1) $C(L) = O(1)_{|\Gamma|\to\infty}$ is uniformly upper bounded by a $C(\infty) < \infty$ in case 1).

Case 2) $C(L) = O(L)_{L\to\infty} = O(\ln|\Gamma|)_{|\Gamma|\to\infty}$ in case 2).

Remark: This Lemma is an approximate tensorization statement for the state ω. I.e. we give an upper bound on the relative entropy distance between it and its Schmidt conditional expectation on the whole lattice Γ in terms of the relative entropies of it and the Schmidt conditional expectations of the fixed finite size regions $\{R_k\}_{k\in K}$. Later on we will be able to generalize this statement rather simply to arbitrary states ρ and the Davies instead of the Schmidt conditional expectation. This will then allow us to extend the existence of a local cMLSI constant, see Theorem 2.1, to the whole lattice with the cost of $C(L)^{-1}$.

Proof of 1) in Lemma 5.8. The proof idea is to average the approximate tensorization result, Theorem 5.6, over all the above defined coverings $\{C_k^{\tilde{l}}, D_k^{\tilde{l},l}\}$ to get the relative entropy between ω and its Schmidt conditional expectation on the whole of $B_{x_j,L}$ in terms of the relative entropy between it and the Schmidt conditional expectation on subregions of height ηL, where $\eta < 1$. Repeating this inductively $O(\log(L))$ times will then give the statement of the Lemma. Hence the constant $C(L)$ will be given as a product of $O(\log(L))$ terms. Bounding this product will will then give the in the Lemma stated asymptotics.

Fix $x_j \in \Gamma_0$ and let $\eta < 1$. We enumerate all partitions of $B_{x_j,L}$ into $\{C_j^{\tilde{l}}, D_j^{\tilde{l},\lfloor\sqrt{L}\rfloor}\}$, s.t. $\tilde{l}, L + \lfloor\sqrt{L}\rfloor - \tilde{l} \leq \eta L$ and s.t. different partitions have disjoint overlaps, i.e. $\left(C_j^{\tilde{l}_1} \cap D_k^{\tilde{l}_1,\lfloor\sqrt{L}\rfloor}\right) \cap \left(C_j^{\tilde{l}_2} \cap D_k^{\tilde{l}_2,\lfloor\sqrt{L}\rfloor}\right) = \emptyset$, whenever $\tilde{l}_1 \neq \tilde{l}_2$. This works as long as $\sqrt{L} \lesssim (2\eta - 1)L$ which gives another condition on the minimal size of $l = \lfloor\sqrt{L}\rfloor \geq: l_{\min,2}$. There exist $\frac{L}{\lfloor\sqrt{L}\rfloor} = O(\sqrt{L})$ of these partitions, since their overlap is of height $\lfloor\sqrt{L}\rfloor = O(\sqrt{L})$. Call these partitions $\{C_i, D_i\}_{i=1}^{O(\sqrt{L})}$. Now we average over all the approximate tensorization results of these partitions to get

$$
\begin{aligned}
&D(\omega \| E^S_{B_{x_j,L}*}(\omega)) \\
&\leq \frac{1}{O(\sqrt{L})} \sum_{i=1}^{O(\sqrt{L})} \frac{1}{1 - 2\epsilon(\sqrt{L})} \left[D(\omega \| E^S_{C_i*}(\omega)) + D(\omega \| E^S_{D_i*}(\omega)) \right] \\
&\leq \frac{1}{1 - 2\epsilon(\sqrt{L})} \frac{1}{O(\sqrt{L})} \sum_{i=1}^{O(\sqrt{L})} C(\mathrm{height}(C_i)) D_{C_i}(\omega) + C(\mathrm{height}(D_i)) D_{D_i}(\omega) \\
&\leq C(\eta L) \frac{1}{1 - 2\epsilon(\sqrt{L})} \frac{1}{O(\sqrt{L})} \sum_{i=1}^{O(\sqrt{L})} (2 D_{C_i \cap D_i}(\omega) + D_{C_i \setminus D_i \cup D_i \setminus C_i}(\omega) \\
&+ \sum_{\substack{R_k \cap (C_i \cap D_i) \neq \emptyset \\ R_k \cap (C_i \setminus D_i) \neq \emptyset}} D(\omega \| E^S_{R_k*}(\omega)) + \sum_{\substack{R_k \cap (C_i \cap D_i) \neq \emptyset \\ R_k \cap (D_i \setminus C_i) \neq \emptyset}} D(\omega \| E^S_{R_k*}(\omega)) \\
&\leq C(\eta L) \frac{1}{1 - 2\epsilon(\sqrt{L})} \frac{1}{O(\sqrt{L})} \left(2 D_{\bigcup_{i=1}^{O(\sqrt{L})} (C_i \cap D_i)}(\omega) + \sum_{i=1}^{O(\sqrt{L})} D_{(C_i \setminus D_i)\partial \cup (D_i \setminus C_i)\partial}(\omega) \right) \\
&\leq C(\eta L) \frac{1}{1 - 2\epsilon(\sqrt{L})} \frac{\lfloor\sqrt{L}\rfloor}{L} \left(2 + \frac{L}{\lfloor\sqrt{L}\rfloor} \right) D_{B_{x_j,L}}(\omega),
\end{aligned}
$$

where in the second line we used the definition of $C(\mathrm{height}(\cdot))$ and in the third line we used that $\mathrm{height}(C_i), \mathrm{height}(D_i) \leq C(\eta L)$ since $\mathrm{height}(C_i) = \tilde{l}, \mathrm{height}(D_i) = L + \lfloor\sqrt{L}\rfloor - \tilde{l} \leq \eta L$. Hence it follows that $C(L) \leq C(\eta L) \frac{1}{1-2\epsilon(\sqrt{L})}(1 + \frac{2}{\sqrt{L}}) =: C(\eta L) f(L)$. Repeating this $M = O(\ln L)$ times, s.t. $\eta^M L = l_0 =: \max\{l_{\min,1}, l_{\min,2}\}$ then gives

$$
C(L) \leq C(l_0) \prod_{k=1}^{M} f(\eta^k L) \leq C(l_0) \prod_{k=0}^{\infty} f(l_0 \eta^{-k}) < \infty,
$$

where the infinite product converges since $\epsilon(\sqrt{L}) = L e^{-\frac{\sqrt{L}}{\xi}}$ is exponentially decaying in L and $(1 + \frac{2}{\sqrt{L}}) \to 1$ fast enough. Note that by definition $C(l_0) = 1$ and that it is independent of the x_j which we fixed. Hence the result follows. □

Proof of 2) in Lemma 5.8. This proof follows exactly the same idea as the one above, with the slight difference that we need to choose the overlap of the coverings C, D to scale as $O(L)$. Hence the number of partitions has to be constant in system size which will give us a constant multiplicative factor in each inductive step. Since we need $O(\log(L))$ steps this gives the $O(L)$ scaling of $C(L)$ in the lemma.
Let $N \in \mathbb{N} \setminus \{1\}$ and $\frac{1}{2} < \eta < 1, s.t. \frac{1}{N} \geq (2\eta - 1)$. We enumerate all partitions of $B_{x_j,L}$ into $\{C_j^{\tilde{l}}, D_j^{\tilde{l},\frac{L}{N}}\}$, s.t. $\tilde{l}, L + \frac{L}{N} - \tilde{l} \leq \eta L$ and s.t. different partitions have disjoint overlaps, i.e. $\left(C_j^{\tilde{l}_1} \cap D_k^{\tilde{l}_1,\frac{L}{N}}\right) \cap \left(C_j^{\tilde{l}_2} \cap D_k^{\tilde{l}_2,\frac{L}{N}}\right) = \emptyset$, whenever $\tilde{l}_1 \neq \tilde{l}_2$. There exist $\frac{L}{\frac{L}{N}} = N$ of these partitions, since their overlap is of height $\frac{L}{N}$. Call these partitions $\{C_i, D_i\}_{i=1}^N$. Now we average over all the approximate tensorization results of these partitions to get

$$
\begin{aligned}
&D(\omega \| E^S_{B_{x_j,L}*}(\omega)) \\
&\qquad \leq \frac{1}{N}\sum_{i=1}^{N} \frac{1}{1-2\epsilon(\frac{L}{N})}\left[D(\omega\|E^S_{C_i*}(\omega)) + D(\omega\|E^S_{D_i*}(\omega))\right] \\
&\qquad \leq \frac{1}{1-2\epsilon(\frac{L}{N})}\frac{1}{N}\sum_{i=1}^{N} C(\text{height}(C_i))D_{C_i}(\omega) + C(\text{height}(D_i))D_{D_i}(\omega) \\
&\qquad \leq C(\eta L)\frac{1}{1-2\epsilon(\frac{L}{N})}\frac{1}{N}\sum_{i=1}^{N} (2D_{C_i\cap D_i}(\omega) + D_{C_i\setminus D_i \cup D_i\setminus C_i}(\omega) \\
&\qquad + \sum_{\substack{R_k\cap(C_i\cap D_i)\neq\emptyset \\ R_k\cap(C_i\setminus D_i)\neq\emptyset}} D(\omega\|E^S_{R_k*}(\omega)) + \sum_{\substack{R_k\cap(C_i\cap D_i)\neq\emptyset \\ R_k\cap(D_i\setminus C_i)\neq\emptyset}} D(\omega\|E^S_{R_k*}(\omega)) \\
&\qquad \leq C(\eta L)\frac{1}{1-2\epsilon(\frac{L}{N})}\frac{1}{N}\left(2D_{\cup_{i=1}^N(C_i\cap D_i)}(\omega) + \sum_{i=1}^{N} D_{(C_i\setminus D_i)\partial\cup(D_i\setminus C_i)\partial}(\omega)\right) \\
&\qquad \leq C(\eta L)\frac{1}{1-2\epsilon(\frac{L}{N})}\frac{1}{N}\,(2+N)\,D_{B_{x_j,L}}(\omega),
\end{aligned}
$$

where in the second line we used the definition of $C(\text{height}(\cdot))$ and in the third line we used that $\text{height}(C_i), \text{height}(D_i) \leq C(\eta L)$ since $\text{height}(C_i) = \tilde{l}, \text{height}(D_i) = L + \frac{L}{N} - \tilde{l} \leq \eta L$. Hence it follows that $C(L) \leq C(\eta L)\frac{1}{1-2\epsilon(\frac{L}{N})}(1+\frac{2}{N}) =: C(\eta L)\tilde{f}(L)(1+\frac{2}{N})$. Repeating this $M = O(\ln L)$ times s.t. $\eta^M L = l_0 =: L_{\min,1}$ then gives

$$
\begin{aligned}
C(L) &\leq C(l_0)(1+\frac{2}{N})^M \prod_{k=1}^{M} \tilde{f}(\eta^k L) \leq C(l_0)(1+\frac{2}{N})^M \prod_{k=0}^{\infty} \tilde{f}(l_0\eta^{-k}) \\
&= O\left(\left(1+\frac{2}{N}\right)^{O(\ln L)}\right) = O(L) = O(\ln|B_{x_j,L}|),
\end{aligned}
$$

where the infinite product converges since $\epsilon\left(\frac{L}{N}\right) = b^L e^{-\frac{L}{N\xi}}$ is exponentially decaying in L under the condition on ξ. We used that $|B_{x_j,L}| = b^L$ and hence $L = O(\ln|B_{x_j,L}|)$. Note that it is independent of the x_j which we fixed and hence the result follows. □

5.3 Putting everything together

Now we can put everything together to prove the main result.

Proof of Theorem 5.1. Let Λ be as in the main Theorem 5.1. Let $\Gamma \subset\subset \Lambda$ be a complete connected finite subgraph. For the $b-$ary tree, $b \in \mathbb{N}$, wlog $\Gamma = B_{0,L}$. Let H_Γ be an element of the uniform, commuting, nearest-neighbour family from the theorem and $\mathcal{L}_\Gamma^D$ the corresponding Davies Lindbaldian with conditional expectations $\{E_\Gamma^D\}_{\Gamma\subset\subset\Lambda}$. Fix a 2-coloring Λ_0 and let $\{x_k\}_{k\in K} = \Gamma_0 = \Lambda_0 \cap \Gamma$ be the corresponding one on Γ. Set $\omega := E^S_{\Gamma_0}(\rho)$ for a state $\rho \in \mathcal{D}(\mathcal{H}_\Gamma)$. We first apply the chain rule for the relative entropy (2.19), with $\sigma \equiv \sigma^\Gamma = E^D_{\Gamma*}(\rho) = E^S_{\Gamma_0*}(\sigma)$:

$$D(\rho\|E^D_{\Gamma*}(\sigma)) = D(\rho\|\sigma) = D(\rho\|E^S_{\Gamma_0*}(\rho)) + D(E^S_{\Gamma_0*}(\rho)\|\sigma) = D(\rho\|\omega) + D(\omega\|\sigma).$$

The first summand $D(\rho\|\omega)$ satisfies exact tensorization (a form of strong subadditivity), since the Schmidt conditional expectations of two sets with distance two between each other commute, $E^S_{\{x_k\}*} \circ E^S_{\{x_j\}*} = E^S_{\{x_j\}*} \circ E^S_{\{x_k\}*} = E^S_{\{x_k\}\cup\{k_j\}*}$ for all $x_k, x_j \in \Gamma_0$, see Proposition 4.4 [8], [40]. That is

$$D(\rho\|\omega) = D(\rho\|E^S_{\Gamma_0*}(\rho)) \leq \sum_{x_k\in\Gamma_0} D(\rho\|E^S_{\{x_k\}*}(\rho)) \overset{(4.12)}{\leq} \sum_{x_k\in\Gamma_0} D\left(\rho\middle\|E^D_{\{x_k\}\partial*}(\rho)\right),$$

The second summand $D(\omega\|\sigma)$ is, using Lemma 5.8, the DPI for the relative entropy, and then Lemma 4.12, bounded by

$$\begin{aligned} D(\omega\|\sigma) = D(\omega\|E^S_{\Gamma*}(\omega)) &\overset{\text{Lemma 5.8}}{\leq} C\sum_{x_k\in\Gamma_0} D(\omega\|E^S_{R_k*}(\omega)) = C\sum_{x_k\in\Gamma_0} D(E^S_{\Gamma_0*}(\rho)\|(E^S_{R_k*}\circ E^S_{\Gamma_0})(\rho)) \\ &\overset{DPI}{\leq} C\sum_{x_k\in\Gamma_0} D(\rho\|E^S_{R_k*}(\rho)) \overset{\text{Lemma 4.12}}{\leq} C\sum_{x_k\in\Gamma_0} D(\rho\|E^D_{R_k\partial*}(\rho)), \end{aligned}$$

where in the second line we used that $E^S_{R_k*}\circ E^S_{\Gamma_0} = E^S_{\Gamma_0}\circ E^S_{R_k*}$, which holds by the construction of the sets R_k in Equation 5.6.[5] Recall, that by Lemma 5.8 in the case of sub-exponential (e.g. the 1d dimensional spin chain) the constant C is independent of system size, whereas in the exponential setting, e.g. the $b-$ary tree it scales logarithmically with system size. Now importantly, the regions $\{\{x_k\}\partial, R_j\partial\}_{k,j}$ are of fixed finite size, hence by Theorem 2.1 there exists cMLSI constants $\alpha_0, \alpha_1 > 0$, s.t. for any j, k

$$\alpha_0 D\left(\rho\|E^D_{\{x_k\}\partial*}(\rho)\right) \leq \mathrm{EP}_{\mathcal{L}^D_{\{x_k\}\partial}}(\rho), \qquad \alpha_1 D\left(\rho\|E^D_{R_j\partial*}(\rho)\right) \leq \mathrm{EP}_{\mathcal{L}^D_{R_j\partial}}(\rho).$$

Hence, putting everything above together, we have

$$\begin{aligned} D(\rho\|E^D_{\Gamma*}(\rho) = D(\rho\|\sigma) &\leq \frac{1}{\min\{\alpha_0,\alpha_1\}}\left(\sum_{x_k\in\Gamma_0}\mathrm{EP}_{\mathcal{L}^D_{\{x_k\}\partial}}(\rho) + C\sum_{k|x_k\in\Gamma_0}\mathrm{EP}_{\mathcal{L}^D_{R_k\partial}}(\rho)\right) \\ &\leq \frac{2(1+mC)}{\min\{\alpha_0,\alpha_1\}}\mathrm{EP}_{\mathcal{L}^D_\Gamma}(\rho), \end{aligned}$$

[5] This is since $\partial R_k \cap \Gamma_0 = \emptyset$. Recall that Γ_0 is the union of single vertices each with distance 2 from each other. Hence the claim follows from Proposition 4.4.

where in the last inequality we used the positivity and additivity of the entropy production and the fact that, by construction, each site $x \in \Lambda$ is contained in at most a constant number $2m$ of regions $R_k\partial$, since they are of fixed finite size. The same holds for $\{x_k\}\partial$ with 2 regions. Thus it follows that

$$\alpha(\mathcal{L}_\Gamma^D) \geq \frac{\min\{\alpha_0, \alpha_1\}}{2(1+mC)} > 0.$$

Therfore in the sub-exponential setting from Lemma 5.8 we get $\alpha(\mathcal{L}_\Gamma^D) = \mathcal{O}(1)_{|\Gamma|\to\infty}$ whereas in the exponential we get $\alpha(\mathcal{L}_\Gamma^D) = \Omega((\ln|\Gamma|)^{-1})_{|\Gamma|\to\infty}$. □

6 Applications

In the following let $\Lambda, \Gamma; \rho, \sigma \in \mathscr{D}(\mathcal{H}_\Gamma)$, and $\mathcal{L}_\Lambda^D$ be as in the main Theorem 5.1, such that $\alpha(\mathcal{L}_\Gamma^D) > 0$ is independent of $|\Gamma|$.

6.1 Exponential convergence to Gibbs states in the thermodynamic limit

A direct consequence of the spin-system size $|\Gamma|$ independence of $\alpha(\mathcal{L}_\Gamma^D)$ is that in the thermodynamic limit we have exponential decay of the relative entropy density between some state at time t and the thermal state, with decay rate α. This gives meaning to exponential convergence in the thermodynamic limit and is a strict improvement to the trivial statement we would get when taking the limit with a spin-chain length decreasing MLSI constant. It is quite the important statement since most materials in condensed matter systems are very very large. E.g. for the case of 1D spin chains this is a novel result.

Corollary 6.1. Note that for local Hamiltonians $D(\rho_t\|\sigma^\Gamma) = O(|\Gamma|)$, and hence

$$\lim_{\Gamma\uparrow\Lambda} \frac{1}{|\Gamma|} D(\rho_t\|\sigma^\Gamma) \leq e^{-\alpha t} \lim_{\Gamma\uparrow\Lambda} \frac{1}{|\Gamma|} D(\rho_t\|\sigma^\Gamma).$$

Proof. For a local Hamiltonian we have $\|H_\Gamma\| = \sum_{\substack{X\subset\Gamma \\ \mathrm{diam}(X)\leq r}} \|\Phi_X\| = O(|\Gamma|)$ and hence

$$d^{|\Gamma|} e^{-\beta O(|\Gamma|)} = \mathrm{Tr}[\mathbb{1} e^{-\beta\|H\|}] \leq \mathrm{Tr}[e^{-\beta H}] \leq \mathrm{Tr}[\mathbb{1} e^{\beta\|H\|}] = d^{|\Gamma|} e^{\beta O(|\Gamma|)},$$

hence taking logarithms gives $Z_\Gamma = \log \mathrm{Tr}[e^{-\beta H_\Gamma}] = O(|\Gamma|)$ and thus we can bound

$$D(\rho\|\sigma^\Gamma) \leq -\mathrm{Tr}\left[\rho \log \frac{e^{-\beta H}}{Z_\Gamma}\right] = Z_\Gamma + \beta\,\mathrm{Tr}[\rho H_\Gamma] = O(|\Gamma|).$$

Now the result follows directly from the main theorem by dividing through $|\Gamma|$ and taking the limit. □

Some more corollaries follow the use of a quantum transport inequality. These are upper bounds on the W_1 distance in terms of the relative entropy. In the quantum case[1] they have connections to eigenstate thermalization, concentration bounds, and many more. See e.g. [1]. From the existence of a MLSI constant α independent of the spin-system size $|\Gamma|$ one can show that the following transport cost inequality holds with a transport cost $c' = c\frac{|\Gamma|}{\alpha}$ linear in spin-system size $|\Gamma|$ [1].

$$\|\rho - \sigma\|_{W_1} \leq \sqrt{c' D(\rho\|\sigma)}. \tag{6.1}$$

For a proof of this see [1](Proposition 16, Theorem 5) and [41](Theorem 3,4). The fact that the transport cost scales linearly in $|\Gamma|$ has amongst others the following important consequences.

[1] and also the classical case [NEED SOURCE]

6.2 Tighter bounds on the entropy difference and convergence via relative entropy

The first application of this concerns bounding the von-Neumann entropy difference between two quantum states ρ, σ by their relative entropy.

Corollary 6.2. If ρ, σ are as assumed above, then the following two bounds hold

$$
\begin{aligned}
i)\ |S(\rho) - S(\sigma)| &\le g\left(c\sqrt{\frac{|\Gamma|}{\alpha}}\sqrt{D(\rho||\sigma)}\right) + c\sqrt{\frac{|\Gamma|}{\alpha}}\ln\left(d^2|\Gamma|\right)\sqrt{D(\rho||\sigma)} \\
&= O(\sqrt{|\Gamma|}\log|\Gamma|)_{|\Gamma|\to\infty}\sqrt{D(\rho||\sigma)}, \qquad (6.2) \\
ii)\ |S(\rho_t) - S(\sigma)| &= O\left(|\Gamma|\log|\Gamma|, \frac{e^{-\frac{\alpha t}{2}}}{\sqrt{\alpha}}\right), \qquad (6.3)
\end{aligned}
$$

where $g(t) = (t+1)\log(t+1) - t\log t = o(t)_{t\to\infty}$, $\rho_t := e^{t\mathcal{L}_{\Lambda^*}}(\rho)$, and c is some constant independent of $|\Gamma|$ depending only on the locality of the Lindbladian $\mathcal{L}_\Lambda^D$.

Remark: These inequalities represent a $O\left(\frac{\sqrt{|\Gamma|}}{\log|\Gamma|}\right)$ improvement compared to the use of Pinsker's inequality, see [26], as $\|\rho-\sigma\|_{W_1} \le \frac{|\Gamma|}{2}\|\rho-\sigma\|_1 \le \frac{|\Gamma|}{\sqrt{2}}\sqrt{D(\rho||\sigma)}$, or to the following inequality from [42] $|S(\rho)-S(\sigma)| \le \sqrt{3}\log(d^{|\Gamma|})\sqrt{D(\rho||\sigma)} = O(|\Gamma|)\sqrt{D(\rho||\sigma)}$. However, these inequalities apply to arbitrary quantum states ρ, σ, whereas (6.2) requires σ to be the Gibbs state of a suitable Hamiltonian. It is also an $O(\log|\Gamma|)$ improvement over the until now best known bound under the above assumptions, derived by employing a MLSI constant that scales as $\Omega(\log|\Gamma|)^{-1}$ [12]. And comparing to [1] these inequalities constitute an extension of the $O(|\Gamma|\log|\Gamma|)$ scaling of the entropy difference from product states to more general Gibbs states satisfying the assumptions of the Corollary. Note that these entropy difference bounds are optimal in their scaling in $|\Gamma|$ up to logarithmic correction, since the entropy difference is an extensive quantity, i.e. it scales as $O(|\Gamma|)$.

Proof. We first use the following continuity bound from [26](Theorem 1), that states that for any two states $\rho, \sigma \in \mathcal{D}(\mathcal{H}_\Gamma)$

$$|S(\rho) - S(\sigma)| \le g(\|\rho-\sigma\|_{W_1}) + \|\rho-\sigma\|_{W_1}\ln(d^2|\Gamma|),$$

where d is the local Hilbert space dimension, $\|\cdot\|_{W_1}$ the quantum Wasserstein distance of order 1 (see (2.13) for the definition), and $g(t) = (t+1)\log(t+1) - t\log t$. Now by the transport cost inequality from [1](Prop 16, Theorem 5) and [41](Theorem 3,4)[2], which holds under the assumptions of the Corollary we get Inequality (6.1). Combining this one with the just above gives i). For ii) use i), the bound $D(\rho||\sigma^\Gamma) \le O(|\Gamma|)$ from above, and the MLSI in its integrated form $D(\rho_t||\sigma) \le e^{-\alpha t}D(\rho||\sigma)$. □

[2]Note that in that paper a different normalization convention for the quantum Wasserstein distance was used as compared to here or the other references in this section

6.3 Optimal 1D Gaussian concentration bound

Corollary 6.3. Let $O \in \mathcal{B}(\mathcal{H}_\Lambda)$ be a k-local (but not necessarily geometrically-k-bounded) observable, i.e.

$$O = \sum_{\substack{X\subset\Lambda\\|X|\le k}} o_X, \tag{6.4}$$

such that for all $i \in \Lambda$ $\sum_{X\subset\Lambda:X\ni i}\|o_X\| \le g$, where each o_X acts only non trivially on sites $X \subset \Lambda$. Let σ be the Gibbs state of some geometrically-local, uniformly bounded, commuting Hamiltonian with uniform exponential decay of correlations, at any fixed inverse temperature $\beta > 0$. Then for $r \ge 0$ it holds that

$$\mathbb{P}_\sigma(|O - \mathrm{Tr}[O\sigma]| \ge r) \le 2\exp\left[-\frac{\alpha r^2}{\mathcal{O}(|\Gamma|)}\right]. \tag{6.5}$$

Remark: This means that 1D spin-systems with geometrically-local, commuting Hamiltonians with uniform exponential decay of correlations give rise to a sub-Gaussian random variable in their thermal equilibrium states for any observables of the above form, notably including long-range observables with a power-law decay. Thus inequality (6.5) constitutes a tightening in terms of its $|\Gamma|$-dependence and generalization to a larger class of observables and r-values of the until now best known 1D Gaussian concentration bound in [43][3]. This bound is optimal in its scaling in $|\Gamma|$ by the Gärtner-Ellis Theorem, see e.g. [44] which is applicable by [45](Theorem 3.2).

Proof. This follows from Theorem 5.1 and [1](Theorem 7 and Lemma 7), when using the transport cost $c' = c\frac{|\Gamma|}{\alpha}$ in inequality (6.1) and the fact that $\|\Delta_\sigma^{\frac{1}{2}}(O)\|_L \le 4gC = \mathcal{O}(1)_{|\Gamma|\to\infty}$, since

$$\begin{aligned}
\|\Delta_\sigma^{\frac{1}{2}}(O)\|_L &\le 2\max_{i\in\Lambda}\left\|\Delta_\sigma^{\frac{1}{2}}(O) - \mathbb{1}_d^{(i)}\otimes tr_i\Delta_\sigma^{\frac{1}{2}}(O)\right\| \\
&= 2\max_{i\in\Lambda}\left\|\sum_{X:|X|\le k}\Delta_\sigma^{\frac{1}{2}}(o_X) - \mathbb{1}_d^{(i)}\otimes tr_i\sum_{X:|X|\le k}\Delta_\sigma^{\frac{1}{2}}(o_X)\right\| \\
&= 2\max_{i\in\Lambda}\left\|\sum_{X:X\partial\ni i,|X|\le k}\left(\Delta_\sigma^{\frac{1}{2}}(o_x) - \mathbb{1}_d^{(i)}\otimes tr_i\Delta_\sigma^{\frac{1}{2}}(o_x)\right)\right\| \\
&\le 2\max_{i\in\Lambda}\sum_{X:X\partial\ni i,|X|\le k}2\left\|\Delta_\sigma^{\frac{1}{2}}(o_x)\right\| \le 4\max_{i\in\Lambda}\sum_{X:|X|\le k,X\partial\ni i}\|o_X\|\left\|\exp\left(\frac{\beta}{2}\sum_{\substack{B:B\cap X\neq\emptyset\\(B)\le r}}\Phi_B\right)\right\| \\
&\le 4g\exp\left(\frac{\beta J c_{r,k,\nu}}{2}\right) \equiv 4gC,
\end{aligned}$$

[3] The statement there [43](Theorem 4.2) is equivalent to $\mathbb{P}_\sigma(|O - \mathrm{Tr}[O\sigma]| \ge r) \le 2\exp\left[-\frac{\alpha r}{\mathcal{O}(\sqrt{|\Gamma|})}\right]$

where *tr* is the normalized partial trace. In the last two inequalities are:

$$\|\Delta_\sigma^{\frac{1}{2}}(o_X)\| = \|\sigma^{\frac{1}{2}} o_X \sigma^{-\frac{1}{2}}\| = \|e^{-\frac{\beta}{2}\sum_{B\cap X\neq\emptyset}\Phi_B} o_X e^{\frac{\beta}{2}\sum_{B\cap X\neq\emptyset}\Phi_B}\|$$
$$\leq \|o_X\|\|\exp\left(\frac{\beta}{2}\sum_{\substack{B:B\cap X\neq\emptyset\\ |X|\leq k,(B)\leq r}}\Phi_B\right)\| \leq \|o_X\|\exp\left(\frac{\beta J c_{r,k,\nu}}{2}\right).$$

$c_{r,k,\nu}$ is a constant which depends only on the locality of O, the geometric locality r of the Hamiltonian, and the growth constant ν of the graph. □

This also implies that any operator of the form of O, i.e. locally bounded and k-local, is Lipschitz. For a reference about the Lipschitz constant see [26].

6.4 Ensemble equivalence under (long-range) Lipschitz observables

The canonical ensemble state to inverse temperature β is given by the Gibbs state σ_β, where we write the temperature dependence explicitly again.
Let $\sigma_{E,\delta}$ be the microcanonical ensemble to energy $E = \arg\max_{E\in\mathbb{R}}\left(e^{-\beta E}\mathcal{N}_{E,\delta}\right)$ and energy-shell width δ [46]. Here $\mathcal{N}_{E,\delta} := \mathrm{Tr}[P((E-\delta,E])]$ is the number of eigenstates in the energy interval $(E-\delta,E]$. i.e. if P is the spectral measure of the Hamiltonian H, i.e. $H = \sum_E EP(E) \equiv \sum_{E_m} E_m |E_m\rangle\langle E_m|$ then

$$\sigma_{E,\delta} := \frac{P((E-\delta,E])}{\mathrm{Tr}[P((E-\delta,E])]} = \frac{1}{\mathcal{N}_{E,\delta}}\sum_{E_m\in(E-\delta,E]}|E_m\rangle\langle E_m|.$$

Two ensembles represented respectively by the families of states $\{\sigma_1^\Gamma,\sigma_2^\Gamma\}_{\Gamma\subset\subset\Lambda}$ are said to be *equivalent* if, in the thermodynamic limit, they produce the same expectation values on averaged geometrically-local observable $\frac{O}{|\Gamma|} = \frac{1}{|\Gamma|}\sum_{i=1}^{|\Gamma|}O_i$, with $\|O_i\| \leq g$ [46]. I.e if for any such observable

$$\left|\mathrm{Tr}\left[\sigma_1^\Gamma\frac{O}{|\Gamma|}\right] - \mathrm{Tr}\left[\sigma_2^\Gamma\frac{O}{|\Gamma|}\right]\right| = \frac{1}{|\Gamma|}|\mathrm{Tr}[(\sigma_1^\Gamma-\sigma_2^\Gamma)O]| \overset{|\Gamma|\to\infty}{\longrightarrow} 0.$$

In [47] it was shown that the microcanonical and canonical ensembles are equivalent in this sense when the system satisfies suitable concentration bounds, such as inequality (6.5). In fact they show equivalence also for observables of form (6.4). We extend this notion of equivalence in the 1D case to a more general class of Lipschitz observables, i.e. $O \in \mathcal{B}(\mathcal{H}_\Gamma)$, s.t. $\|O\|_L < \infty$. These notably include long-range locally bounded, $k-$local observables of the form (6.4).

Corollary 6.4 ([1]Corollary 2 applied to our setting)**.** For any Lipschitz observable O, i.e. $O \in \mathcal{B}(\mathcal{H}_\Lambda)$ s.t. $\|O\|_L < \infty$ and for $\sigma_{E,\delta},\sigma_\beta$ the micro- and canonical ensemble states with the same energy $E = \arg\max_{E\in\mathbb{R}}\left(e^{-\beta E}\mathcal{N}_{E,\delta}\right)$, respectively, it holds that

$$\frac{1}{|\Gamma|}|\mathrm{Tr}[\sigma_{E,\delta}O] - \mathrm{Tr}[\sigma_\beta O]| \leq \|O\|_L o(1)_{|\Gamma|\to\infty}, \tag{6.6}$$

where we can take the energy shell width $\delta = e^{-O(|\Gamma|)}$.

Proof. This follows directly from [1] (Corollary 2) when employing the linearity of the transport cost in the system size. The idea is to employ the transport cost inequality and then bound the relative entropy between the microcanonical and canonical ensembles suitably. □

7 Summary and Outlook

7.1 Summary and future work

In this work we showed that uniform families of commuting nearest-neighbour Hamiltonians on two-colorable sub-exponential graphs satisfy an exponential decay of relative entropy towards their Gibbs states with an exponential decay rate α that is spin-system size independent. This was under the assumption of Davies evolving and the general condition that the Davies generators are gaped. For 1 dimensional system this gives a strict improvement in the scaling of the MLSI constant in system size and has an implication on the convergence to the Gibbs state in the thermodynamic limit. For exponential graphs, such as $b-$ary trees we proved existence of a positive MLSI sonstant α scaling with $\Omega((\ln|\Gamma|)^{-1})$ in system size $|\Gamma|$. For both of these types of graphs we established a weaker criterion for existence of a strictly positive MLSI and rapid thermalization to occur. This has many interesting implications such optimal Gaussian concentration bounds, exponential convergence to Gibbs states in the thermodynamic limit, and connections to ETH. If such Hamiltonians and evolutions are on exponential graphs, we showed, that the exponential decay rate α is only logarithmically decreasing in system size, hence we recover rapid thermalization under certain additional assumptions on the temperature. We showed that essentially all nearest neighbour commuting systems thermalize quickly at large enough temperatures. And rapid thermalization has implications in efficient quantum algorithm designs, Gibbs and ground state sampling, and more. For trees this is a completely new result in the quantum framework. However, for classical system it is known that for systems on trees, a system size independent exponential decay rate α can be derived [35] under very similar conditions as we impose in that setting. This leads us to the following conjecture.

Conjecture 1. There exists a system size independent, strictly positive MLSI constant for trees (and general exponential graphs) under pretty much the same conditions as we impose in this work. This may be done by following an analogous strategy to [35], just for quantum instead classical systems.

We further think it should not require many more tools than the ones established in this work or the literature. This would be very interesting, since this is a step towards a general theory of thermalization of quantum spin systems, which would explain under what exact conditions quantum spin systems thermalize rapidly and how fast if not. As a first step towards this, we already showed in this work that nearest-neighbour, commuting, translation invariant, systems on 2-colorable, subexponential graphs at high enough temperatures will thermalize rapidly. Removing the requirement on the growth constant of the graph seems physically implausible. Removing the requirement of nearest-neighbour interactions would, however, be one very interesting next research direction. This should be somewhat difficult however, since we know certain commuting non-nearest neighbour systems, which have a qualitatively different thermalization behaviour. E.g the four dimensional Toric-code is known to be robust against thermal noise for low enough temperatures [48]. Next we discuss the requirement of 2-colorability.
In Chapter 5 we had to restrict our analysis to 2-colorable graphs due to the use of the Schmidt

conditional expectation and the conditioned state ω, on which the proof of the main result relies. As we mentioned all hypercubic and loop-free graphs are two-colorable. However, we may relax this condition to graphs, which are two-colorable after finite size coarse graining. I.e. given some graph, whenever there exists a constant d, independent of graph-size, such that we may think of subsets of diameter up to size d as one vertex, such that the resulting graph is two colorable. Note that without the condition that d is graph-size independent this would be trivial. This does works for example for very simple examples, like certain infinite pentagonal tilings[1], which are not two colorable, but after coarse graining set of size up to $d = 4$ give two-colorable, and indeed hypercubic, graphs. For this specific example see appendix A.2 and Figure A.1.

Conjecture 2. We suspect this *2-colorability after finite-size-coarse-graining* to be the case for a large class of graphs, but are unaware of general theorems giving conditions on graphs in order to satisfy this. Indeed we conjecture that any regular graph can be finite-size-coarse-grained into a hypercubic, and hence two-colorable, one. Although we also suspect this to hold for certain classes of non-regular graphs.

We also showed that for nearest neighbour commuting systems on a 2-colorable lattice, weak $\mathbb{L}_\infty$-clustering is equivalent to strong $q\mathbb{L}_1 \to \mathbb{L}_\infty$ clustering. This established the existence of a gaped generator as a sufficient condition for a system-size independent MLSI constant and not only gives a rather weak condition for exponential decay of relative entropy and rapid thermalization of such systems, but also shows the equivalence between system size independent gap and rapid thermalization affirmatively answering the open question from [3] for when weak ($\mathbb{L}_\infty$-clustering) implies existence of a gap, for nearest-neighbour interacting systems. We also established that weak clustering is equivalent to exponential decay of mutual information and implies something we call strong local indistinguishability regardless of the geometry, only requiring geometric-locality and commutativity of the Hamiltonian.

Conjecture 3. Theorem 3.4 should, like in the 1 dimensional case Theorem 3.5, hold also without the constraint of commuting Hamiltonians when the temperature β^{-1} is high enough.

Most of the work in doing this is to establish Lemma 3.6 and especially property 1) in the non-commuting case. The rest essentially follows from the proof here and some additional work analogously to the proofs in [7]. The way it is done here relies on the fact that the Araki-expansionals $E_{A,B}$ are positive, which only holds if the Hamiltonian is commuting. We suspect a way around this may be via a suitable cluster expansion, or similar technique, which should converge for high enough temperatures $\beta \leq \beta^*$ and give a suitable bound on $\|E^{\pm 1}_{A,B}\|$ via quantum belief-propagation. Some of the collaborators of this work had proven essentially this under the assumption that the initial claims in [49] were correct. Since the issues there seem to be practically resolved this conjecture may essentially already be proven.

Of enormous importance to this work was that we were only considering locally finite dimensional, i.e. spin-lattice systems, since e.g. the results on the existence of a local strictly positive MLSI constant only holds in the finite dimensional setting. Also the tools we used, such as the relative entropy, entropy production, etc... are much simpler to define and work with in the finite dimensional setting. A body of work has and is currently establishing suitable generalizations of some of the tools used in this work to the von Neumann algebraic setting. This is a natural framework to describe locally infinite dimensional, i.e. bosonic systems, such as those encountered in quantum optics. These are of large interest to the physics community since they are assumed

[1] E.g the Monohedral convex pentagonal tiling with ccm symmetry

to have qualitatively different dissipative behaviours. Understanding these, in contrast to local spin systems, is of large interest in the quest for noise-robust quantum information storage. Hence establishing new, suitable tools and extending the ones from this work from the matrix algebra to the general and more abstract von Neumann algebra setting is also a interesting, broad, and quite promising future research direction.

Literatur

[1] G. D. Palma und C. Rouzé, “Quantum Concentration Inequalities,” *Annales Henri Poincaré*, Jg. 23, Nr. 9, S. 3391–3429, Apr. 2022. DOI: 10.1007/s00023-022-01181-1. Adresse: https://doi.org/10.1007%2Fs00023-022-01181-1.

[2] Chi-Fang, Chen, M. J. Kastoryano, F. G. S. L. Brandão und A. Gilyén, *Quantum Thermal State Preparation*, 2023. arXiv: 2303.18224 [quant-ph].

[3] M. J. Kastoryano und F. G. S. L. Brandao, *Quantum Gibbs Samplers: the commuting case*, 2016. arXiv: 1409.3435 [quant-ph].

[4] F. G. S. L. Brandao und M. J. Kastoryano, “Finite correlation length implies efficient preparation of quantum thermal states,” 2016. DOI: 10.48550/ARXIV.1609.07877. Adresse: https://arxiv.org/abs/1609.07877.

[5] M. J. Kastoryano und J. Eisert, “Rapid mixing implies exponential decay of correlations,” *Journal of Mathematical Physics*, Jg. 54, Nr. 10, S. 102 201, Okt. 2013. DOI: 10.1063/1.4822481. Adresse: https://doi.org/10.1063%2F1.4822481.

[6] M. Kliesch, C. Gogolin, M. Kastoryano, A. Riera und J. Eisert, “Locality of Temperature,” *Physical Review X*, Jg. 4, Nr. 3, Juli 2014. DOI: 10.1103/physrevx.4.031019. Adresse: https://doi.org/10.1103%5C%2Fphysrevx.4.031019.

[7] A. Bluhm, Á. Capel und A. Pérez-Hernández, “Exponential decay of mutual information for Gibbs states of local Hamiltonians,” *Quantum*, Jg. 6, S. 650, Feb. 2022, ISSN: 2521-327X. DOI: 10.22331/q-2022-02-10-650. Adresse: https://doi.org/10.22331/q-2022-02-10-650.

[8] L. Gao und C. Rouzé, “Complete Entropic Inequalities for Quantum Markov Chains,” *Archive for Rational Mechanics and Analysis*, Jg. 245, Nr. 1, S. 183–238, Mai 2022. DOI: 10.1007/s00205-022-01785-1. Adresse: https://doi.org/10.1007%5C%2Fs00205-022-01785-1.

[9] A. Capel, A. Lucia und D. Perez-Garcia, “Superadditivity of Quantum Relative Entropy for General States,” *IEEE Transactions on Information Theory*, Jg. 64, Nr. 7, S. 4758–4765, Juli 2018. DOI: 10.1109/tit.2017.2772800. Adresse: https://doi.org/10.1109%2Ftit.2017.2772800.

[10] Á. Capel, A. Lucia und D. Pérez-Garcıa, “Quantum conditional relative entropy and quasi-factorization of the relative entropy,” *Journal of Physics A: Mathematical and Theoretical*, Jg. 51, Nr. 48, S. 484 001, Nov. 2018. DOI: 10.1088/1751-8121/aae4cf. Adresse: https://doi.org/10.1088%2F1751-8121%2Faae4cf.

[11] Á. Capel, C. Rouzé und D. S. França, *The modified logarithmic Sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions*, 2020. DOI: 10.48550/ARXIV.2009.11817. Adresse: https://arxiv.org/abs/2009.11817.

[12] I. Bardet, Á. Capel, L. Gao, A. Lucia, D. Pérez-García und C. Rouzé, *Entropy decay for Davies semigroups of a one dimensional quantum lattice*, 2021. DOI: 10.48550/ARXIV.2112.00601. Adresse: https://arxiv.org/abs/2112.00601.

[13] A. Rivas und S. F. Huelga, *Open Quantum Systems*. Springer Berlin Heidelberg, 2012. DOI: 10.1007/978-3-642-23354-8. Adresse: https://doi.org/10.1007%2F978-3-642-23354-8.

[14] R. J. Archbold, "E. B. Davies, One-Parameter Semigroups (Academic Press, London, 1980), viii 230 pp., £19.80.," *Proceedings of the Edinburgh Mathematical Society*, Jg. 26, Nr. 1, S. 115–116, 1983. DOI: 10.1017/S0013091500028169.

[15] T. S. Cubitt, A. Lucia, S. Michalakis und D. Perez-Garcia, "Stability of Local Quantum Dissipative Systems," *Communications in Mathematical Physics*, Jg. 337, Nr. 3, S. 1275–1315, Apr. 2015. DOI: 10.1007/s00220-015-2355-3. Adresse: https://doi.org/10.1007%2Fs00220-015-2355-3.

[16] F. G. S. L. Brandã o, T. S. Cubitt, A. Lucia, S. Michalakis und D. Perez-Garcia, "Area law for fixed points of rapidly mixing dissipative quantum systems," *Journal of Mathematical Physics*, Jg. 56, Nr. 10, Okt. 2015. DOI: 10.1063/1.4932612. Adresse: https://doi.org/10.1063%2F1.4932612.

[17] C.-F. Chen und F. G. S. L. Brandão, *Fast Thermalization from the Eigenstate Thermalization Hypothesis*, 2023. arXiv: 2112.07646 [quant-ph].

[18] L. D'Alessio, Y. Kafri, A. Polkovnikov und M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics," *Advances in Physics*, Jg. 65, Nr. 3, S. 239–362, 2016. DOI: 10.1080/00018732.2016.1198134. eprint: https://doi.org/10.1080/00018732.2016.1198134. Adresse: https://doi.org/10.1080/00018732.2016.1198134.

[19] F. G. Brandao und K. M. Svore, "Quantum Speed-Ups for Solving Semidefinite Programs," in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, 2017, S. 415–426. DOI: 10.1109/FOCS.2017.45.

[20] J. van Apeldoorn, A. Gilyén, S. Gribling und R. de Wolf, "Quantum SDP-Solvers: Better upper and lower bounds," *Quantum*, Jg. 4, S. 230, Feb. 2020. DOI: 10.22331/q-2020-02-14-230. Adresse: https://doi.org/10.22331%2Fq-2020-02-14-230.

[21] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy und R. Melko, "Quantum Boltzmann Machine," *Phys. Rev. X*, Jg. 8, S. 021 050, 2 Mai 2018. DOI: 10.1103/PhysRevX.8.021050. Adresse: https://link.aps.org/doi/10.1103/PhysRevX.8.021050.

[22] I. Bardet, Á. Capel, L. Gao, A. Lucia, D. Pérez-Garcıa und C. Rouzé, "Rapid Thermalization of Spin Chain Commuting Hamiltonians," *Physical Review Letters*, Jg. 130, Nr. 6, Feb. 2023. DOI: 10.1103/physrevlett.130.060401. Adresse: https://doi.org/10.1103%2Fphysrevlett.130.060401.

[23] M. Junge, N. Laracuente und C. Rouzé, "Stability of Logarithmic Sobolev Inequalities Under a Noncommutative Change of Measure," *Journal of Statistical Physics*, Jg. 190, Nr. 2, Dez. 2022. DOI: 10.1007/s10955-022-03026-x. Adresse: https://doi.org/10.1007%2Fs10955-022-03026-x.

[24] Y. M. Miranda und G. Slade, "The growth constants of lattice trees and lattice animals in high dimensions," *Electronic Communications in Probability*, Jg. 16, Nr. none, S. 129–136, 2011. DOI: 10.1214/ECP.v16-1612. Adresse: https://doi.org/10.1214/ECP.v16-1612.

[25] G. Pisier und Q. Xu, “Chapter 34 - Non-Commutative Lp-Spaces,” in Ser. Handbook of the Geometry of Banach Spaces, W. Johnson und J. Lindenstrauss, Hrsg., Bd. 2, https://dmitripavlov.org/scans/pisier-xu.pdf, Elsevier Science B.V., 2003, S. 1459–1517. DOI: https://doi.org/10.1016/S1874-5849(03)80041-4. Adresse: https://www.sciencedirect.com/science/article/pii/S1874584903800414.

[26] G. De Palma, M. Marvian, D. Trevisan und S. Lloyd, “The Quantum Wasserstein Distance of Order 1,” *IEEE Transactions on Information Theory*, Jg. 67, Nr. 10, S. 6627–6643, 2021. DOI: 10.1109/TIT.2021.3076442.

[27] M. Takesaki, “Conditional expectations in von Neumann algebras,” *Journal of Functional Analysis*, Jg. 9, Nr. 3, S. 306–321, 1972, ISSN: 0022-1236. DOI: https://doi.org/10.1016/0022-1236(72)90004-3. Adresse: https://www.sciencedirect.com/science/article/pii/0022123672900043.

[28] D. Petz, “Quasi-entropies for finite quantum systems,” *Reports on Mathematical Physics*, Jg. 23, Nr. 1, S. 57–65, 1986, ISSN: 0034-4877. DOI: https://doi.org/10.1016/0034-4877(86)90067-4. Adresse: https://www.sciencedirect.com/science/article/pii/0034487786900674.

[29] D. Petz, “The proper formula for relative entropy and its asymptotics in quantum probability,” *Communications in Mathematical Physics*, Jg. 143, Dez. 1991. DOI: 10.1007/BF02100287.

[30] M. Brannan, L. Gao und M. Junge, *Complete Logarithmic Sobolev inequalities via Ricci curvature bounded below*, 2020. arXiv: 2007.06138 [math.OA].

[31] L. Gao, M. Junge, N. LaRacuente und H. Li, *Complete order and relative entropy decay rates*, Sep. 2022. DOI: 10.48550/arXiv.2209.11684.

[32] L. Gao, M. Junge und H. Li, *Geometric Approach Towards Complete Logarithmic Sobolev Inequalities*, 2021. arXiv: 2102.04434 [quant-ph].

[33] I. Bardet, *Estimating the decoherence time using non-commutative Functional Inequalities*, 2017. arXiv: 1710.01039 [quant-ph].

[34] M. J. Kastoryano und K. Temme, “Quantum logarithmic Sobolev inequalities and rapid mixing,” *Journal of Mathematical Physics*, Jg. 54, Nr. 5, Mai 2013. DOI: 10.1063/1.4804995. Adresse: https://doi.org/10.1063%5C%2F1.4804995.

[35] F. Martinelli, A. Sinclair und D. Weitz, “Glauber Dynamics on Trees: Boundary Conditions and Mixing Time,” *Communications in Mathematical Physics*, Jg. 250, Nr. 2, S. 301–334, Aug. 2004. DOI: 10.1007/s00220-004-1147-y. Adresse: https://doi.org/10.1007%2Fs00220-004-1147-y.

[36] D. Aharonov, I. Arad, Z. Landau und U. Vazirani, *Quantum Hamiltonian complexity and the detectability lemma*, 2011. arXiv: 1011.3445 [quant-ph].

[37] S. Bravyi und M. Vyalyi, *Commutative version of the k-local Hamiltonian problem and common eigenspace problem*, 2004. arXiv: quant-ph/0308021 [quant-ph].

[38] I. Bardet, Á. Capel und C. Rouzé, “Approximate Tensorization of the Relative Entropy for Noncommuting Conditional Expectations,” *Annales Henri Poincaré*, Jg. 23, Nr. 1, S. 101–140, Juli 2021. DOI: 10.1007/s00023-021-01088-3. Adresse: https://doi.org/10.1007%2Fs00023-021-01088-3.

[39] K. Temme und M. J. Kastoryano, *How fast do stabilizer Hamiltonians thermalize?* 2015. arXiv: 1505.07811 [quant-ph].

[40] D. Petz, “On certain properties of the relative entropy of states of operator algebras,” *Mathematische Zeitschriften*, Jg. 206, S. 351–361, Jan. 1991. DOI: 10.1007/BF02571350.

[41] C. Rouzé und N. Datta, *Concentration of quantum states from quantum functional and transportation cost inequalities*, 2017. DOI: 10.48550/ARXIV.1704.02400. Adresse: https://arxiv.org/abs/1704.02400.

[42] D. Reeb und M. M. Wolf, “Tight Bound on Relative Entropy by Entropy Difference,” *IEEE Transactions on Information Theory*, Jg. 61, Nr. 3, S. 1458–1473, März 2015. DOI: 10.1109/tit.2014.2387822. Adresse: https://doi.org/10.1109%2Ftit.2014.2387822.

[43] A. Anshu, “Concentration bounds for quantum states with finite correlation length on quantum spin lattice systems,” *New Journal of Physics*, Jg. 18, Nr. 8, S. 083 011, Aug. 2016. DOI: 10.1088/1367-2630/18/8/083011. Adresse: https://doi.org/10.1088%2F1367-2630%2F18%2F8%2F083011.

[44] H. Touchette, “The large deviation approach to statistical mechanics,” *Physics Reports*, Jg. 478, Nr. 1-3, S. 1–69, Juli 2009. DOI: 10.1016/j.physrep.2009.05.002. Adresse: https://doi.org/10.1016%2Fj.physrep.2009.05.002.

[45] M. Lenci und L. Rey-Bellet, “Large Deviations in Quantum Lattice Systems: One-Phase Region,” *Journal of Statistical Physics*, Jg. 119, Nr. 3-4, S. 715–746, Mai 2005. DOI: 10.1007/s10955-005-3015-3. Adresse: https://doi.org/10.1007%2Fs10955-005-3015-3.

[46] T. Kuwahara und K. Saito, “Eigenstate Thermalization from the Clustering Property of Correlation,” *Physical Review Letters*, Jg. 124, Nr. 20, Mai 2020. DOI: 10.1103/physrevlett.124.200604. Adresse: https://doi.org/10.1103%2Fphysrevlett.124.200604.

[47] T. Kuwahara und K. Saito, “Gaussian concentration bound and Ensemble equivalence in generic quantum many-body systems including long-range interactions,” *Annals of Physics*, Jg. 421, S. 168 278, Okt. 2020. DOI: 10.1016/j.aop.2020.168278. Adresse: https://doi.org/10.1016%2Fj.aop.2020.168278.

[48] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self und J. R. Wootton, “Quantum memories at finite temperature,” *Reviews of Modern Physics*, Jg. 88, Nr. 4, Nov. 2016. DOI: 10.1103/revmodphys.88.045005. Adresse: https://doi.org/10.1103%2Frevmodphys.88.045005.

[49] T. Kuwahara, K. Kato und F. G. Brandão, “Clustering of Conditional Mutual Information for Quantum Gibbs States above a Threshold Temperature,” *Physical Review Letters*, Jg. 124, Nr. 22, Juni 2020. DOI: 10.1103/physrevlett.124.220601. Adresse: https://doi.org/10.1103%2Fphysrevlett.124.220601.

List of Figures

3.1 The graph region Γ, here depicted as a colored region without the edges is partitioned into the three sub-regions A, C and B, such s.t. dist(A,C)=l. This is an example of a partitioning in Theorem 3.4. If the Graph Γ is in a state which satisfies $\mathbb{L}_\infty$-clustering, then via this theorem, the reduced state on AC is approximately a product state between the reduced states on regions A and C. Approximately in the sense of the strong similarity relation $\overset{\epsilon}{\sim}$ from 3, where ϵ is exponentially decaying in l. 20

4.1 Simple example of the notation introduced in Section 4.2 to define the local algebras $\mathcal{N}_a$ and thus the Schmidt conditional expectation E_a^S. Depicted is a small region of a 3-ary tree, with a (red) vertex labeled a=$c_{i,0}$. Its neighbours (yellow) are labeled with $\{b_i\}_{i\in I} = \partial\{a\}$, where $I \equiv I_a = \{1,2,3,4\}$. The next-nearest neighbours (purple) are $\partial(\partial\{a\}) \setminus \{a\} = \{c_{i,j}\}_{i\in I, j\in J^{(i)}\setminus\{0\}}$. The central vertex a is logically the same as $c_{i,j0}$ for each $i \in I$. In the proof of proposition 4.3 the case i) considers for example the turquiose shaded edge, case ii) for example the blue shaded edge and case iii) for example the dark blue shaded edge. The boundary between the subsets $\{a\}\partial$ and $(\{a\}\partial)^c$ is marked with a gray dotted line. 28

5.1 Relation between the different notions of clustering (static properties) and their relation to thermalization and gap (dynamic properties), in the setting of this thesis. The red implications signify novel implications, wheres the black ones signify results from literature. The main result of Chapter 3 is Theorem 3.4, whereas the main result of this thesis in Chapter 5 is Theorem 5.1. In effect we show that (system size invariant gap) is equivalent to a system size independent MLSI constant via the visualized chain of implications. For details on the necessary assumptions and requirements on the systems, see the respective theorems. 33

5.2 Partition of a subregion CD of a tree into two overlapping subregions C and D. We have $E := C \setminus (D\partial), F := C \cap D, G := D \setminus (C\partial)$. The splitting of the boundary Hilbert spaces corresponding to a boundary site $\{b_i\}$ in the boundary of a region $A \in \{E, F, G\}$ into $\bigotimes_{i\in I_A} P^{\alpha_i}\mathcal{H}_{b_i} = \mathcal{H}^{\alpha_i}_{\partial_{\text{in}}A} \otimes \mathcal{H}^{\alpha_i}_{\partial_{\text{out}}A} = \left(\mathcal{H}_0^{\alpha_i} \otimes \mathcal{H}_c^{\alpha_i}\right) \otimes \left(\bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathcal{H}_j^{\alpha_i}\right)$ is represented by a dotted line. Hence e.g. the Hilbert space of the region $\tilde{l}_{\text{in}}$ is $\mathcal{H}^{(\alpha_{\tilde{l}})}_{\partial_{\text{in}}^{\text{left}}E}$ and of $\tilde{l}_{\text{out}}$ it is $\mathcal{H}^{(\alpha_{\tilde{l}})}_{\partial_{\text{out}}^{\text{left}}E}$. Here the superskript *left*, refers to the part of $\partial_{\text{in}}E$, which is located in the geometric region $\tilde{l}$, i.e. the in the figure left boundary of the region E. Respectively, this is the same with the other boundaries. We fix the boundary conditions $(\alpha_{\tilde{l}})$ on region $\tilde{l}$, (γ_l) on l, (β_r) on r, and $(\alpha_{\tilde{r}})$ on $\tilde{r}$. 37

5.3 Example of a partition of $B_{0,6} \subset \mathbb{T}_2$ into the regions $C_0^{\tilde{l}}$, $D_0^{\tilde{l},l}$ with height of the regions $\tilde{l} = 4$ and $l = 6 - \tilde{l} + 2 = 4$ and the height of their overlap being $l = 2$. This is in $B_{0,6}$ as part of a 2-ary tree. The red vertices are the ones of index 0. Notice that each of these sets 'begins' and 'ends' in these sets. 41

A.1 Example of a section of the graph of an infinite pentagonal tiling. It is 3-colorable, but not 2-colorable, since it contains loops of odd length 5. It becomes 2-colorable, however, after finite size coarse graining of regions of size up to 4. A suitable coarse graining is given by the yellow and turquoise regions. 63

Appendix A

A.1 Omitted proofs from the main text

Proof of proposition 3.2. The reflexivity is trivial, for symmetry see that the spectra and thus the spectral radii of $B^{-\frac{1}{2}}AB^{-\frac{1}{2}} - \mathbb{1}$ and $A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} - \mathbb{1}$ are the same. Since they are both normal we have $\|B^{-\frac{1}{2}}AB^{-\frac{1}{2}} - \mathbb{1}\| \leq \epsilon$ which proves the second implication. Now the first implication follows via

$$\begin{aligned}\|B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}} - \mathbb{1}\| = \|(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})^{-1} - \mathbb{1}\| = \|\sum_{k=1}^{\infty}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}} - \mathbb{1})^k\| \leq \sum_{k=1}^{\infty}\|B^{-\frac{1}{2}}AB^{-\frac{1}{2}} - \mathbb{1}\|^k \\ \leq \frac{\epsilon}{1-\epsilon},\end{aligned}$$

where in the last line we used that by assumption $\|B^{-\frac{1}{2}}AB^{-\frac{1}{2}} - \mathbb{1}\| \leq \epsilon$.
For transitivity, see that

$$\begin{aligned}\|A^{\frac{1}{2}}C^{-1}A^{\frac{1}{2}} - \mathbb{1}\| &= \|A^{\frac{1}{2}}B^{-\frac{1}{2}}(B^{\frac{1}{2}}C^{-1}B^{\frac{1}{2}} - \mathbb{1} + \mathbb{1})B^{-\frac{1}{2}}A^{\frac{1}{2}} - \mathbb{1}\| \\ &\leq \|A^{\frac{1}{2}}B^{-\frac{1}{2}}(B^{\frac{1}{2}}C^{-1}B^{\frac{1}{2}} - \mathbb{1})B^{-\frac{1}{2}}A^{\frac{1}{2}}\| + \|A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} - \mathbb{1}\| \\ &\leq \|(B^{\frac{1}{2}}C^{-1}B^{\frac{1}{2}} - \mathbb{1})(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})\| + \|A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} - \mathbb{1}\| \\ &\leq \epsilon_2(1+\epsilon_1) + \epsilon_1 = \eta.\end{aligned}$$

For the second inequality note that $\|XYX^*\| = |\sigma(XYX^*)| = |\sigma(X^*XY)| \leq \|X^*XY\|$ holds, for X, Y some operators with Y self-adjoint, and $|\sigma(X)|$ denotes the spectral radius of X. Using this with $X = A^{\frac{1}{2}}B^{-\frac{1}{2}}$ and $Y = B^{\frac{1}{2}}C^{-1}B^{\frac{1}{2}} - \mathbb{1}$ then gives the second inequality. The third inequality follows from the assumptions and the fact that $\|B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\| = |\sigma(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})| = |\sigma(A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}})| = |\sigma(A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} - \mathbb{1}) + 1| \leq 1 + \epsilon_1$. See also the second implication of symmetry.
For the tensor multiplicativity 3) see that

$$\begin{aligned}\|(A\otimes B)^{\frac{1}{2}}(\tilde{A}\otimes\tilde{B})^{-1}(A\otimes B)^{\frac{1}{2}} - \mathbb{1}\| &= \|A^{\frac{1}{2}}\tilde{A}^{-1}A^{\frac{1}{2}}\otimes B^{\frac{1}{2}}\tilde{B}^{-1}B^{\frac{1}{2}} - \mathbb{1}\otimes\mathbb{1}\| \\ &\leq \|(A^{\frac{1}{2}}\tilde{A}^{-1}A^{\frac{1}{2}} - \mathbb{1})\otimes(B^{\frac{1}{2}}\tilde{B}^{-1}B^{\frac{1}{2}} - \mathbb{1})\| \\ &+ \|(A^{\frac{1}{2}}\tilde{A}^{-1}A^{\frac{1}{2}} - \mathbb{1})\otimes\mathbb{1}\| + \|\mathbb{1}\otimes(B^{\frac{1}{2}}\tilde{B}^{-1}B^{\frac{1}{2}} - \mathbb{1})\| \\ &\leq \epsilon_1\epsilon_2 + \epsilon_1 + \epsilon_2 = \eta.\end{aligned}$$

For property 4) see the following chain of implications

$$\begin{aligned}\|D^{\frac{1}{2}}E^{-1}D^{\frac{1}{2}} - \mathbb{1}\| \leq \epsilon &\Leftrightarrow 1-\epsilon \leq D^{\frac{1}{2}}E^{-1}D^{\frac{1}{2}} \leq 1+\epsilon \\ &\Leftrightarrow \frac{1}{1+\epsilon} \leq D^{-\frac{1}{2}}ED^{-\frac{1}{2}} \leq \frac{1}{1-\epsilon} \Leftrightarrow \frac{D}{1+\epsilon} \leq E \leq \frac{D}{1-\epsilon}\end{aligned}$$

Since P is self-adjoint the above implies that the following holds:

$$\frac{PDP}{1+\epsilon} \leq PEP \leq \frac{PDP}{1-\epsilon} \implies \frac{\operatorname{tr}_{\mathcal{K}} PDP}{1+\epsilon} \leq \operatorname{tr}_{\mathcal{K}} PEP \leq \frac{\operatorname{tr}_{\mathcal{K}} PDP}{1-\epsilon}.$$

When $\cdot^{-1}$ represents the generalized inverse, the above chain holds also in reverse on $\operatorname{supp}\operatorname{tr}_{\mathcal{K}}(PDP) = \operatorname{supp}\operatorname{tr}_{\mathcal{K}}(PEP)$. I.e. it implies that $\|(\operatorname{tr}_{\mathcal{K}} PDP)^{\frac{1}{2}}(\operatorname{tr}_{\mathcal{K}} PEP)^{-1}(\operatorname{tr}_{\mathcal{K}} PDP)^{\frac{1}{2}} - \mathbb{1}_{\operatorname{supp}(\operatorname{tr}_{\mathcal{K}}(PDP))}\| \leq \epsilon$.
For property 4′) first note that if $D \overset{\epsilon}{\sim} E$, then $D \overset{\mu}{\sim} \lambda E$ where $\mu = \frac{\epsilon}{\lambda} + |1 - \frac{1}{\lambda}|$, since

$$\|D^{\frac{1}{2}}(\lambda E)^{-1}D^{\frac{1}{2}} - \mathbb{1}\| = |\lambda^{-1}|\|D^{\frac{1}{2}}E^{-1}D^{\frac{1}{2}} - \mathbb{1}\| + \|\lambda^{-1}\mathbb{1} - \mathbb{1}\|.$$

Now since $D \overset{\epsilon}{\sim} E \implies \operatorname{tr}_{\mathcal{K}} PDP \overset{\epsilon}{\sim} \operatorname{tr}_{\mathcal{K}} PEP \Leftrightarrow \frac{\operatorname{tr}_{\mathcal{K}} PDP}{1+\epsilon} \leq \operatorname{tr}_{\mathcal{K}} PEP \leq \frac{\operatorname{tr}_{\mathcal{K}} PDP}{1-\epsilon} \implies 1+\epsilon \geq \frac{\operatorname{Tr}[PDP]}{\operatorname{Tr}[PEP]} \geq 1-\epsilon$ by the proof of property 4. It follows that

$$\begin{aligned}\|\mathscr{X}\| &\equiv \left\| \left(\frac{\operatorname{tr}_{\mathcal{K}} PDP}{\operatorname{Tr}[PDP]}\right)^{\frac{1}{2}} \left(\frac{\operatorname{tr}_{\mathcal{K}} PEP}{\operatorname{Tr}[PEP]}\right)^{-1} \left(\frac{\operatorname{tr}_{\mathcal{K}} PDP}{\operatorname{Tr}[PDP]}\right)^{\frac{1}{2}} - \mathbb{1}\right\| \\ &= \|(\operatorname{tr}_{\mathcal{K}} PDP)^{\frac{1}{2}}(\operatorname{tr}_{\mathcal{K}} PEP)^{-1}(\operatorname{tr}_{\mathcal{K}} PDP)^{\frac{1}{2}}\lambda - \mathbb{1}\|,\end{aligned}$$

with $1-\epsilon \leq \lambda^{-1} = \frac{\operatorname{Tr}[PDP]}{\operatorname{Tr}[PEP]} \leq 1+\epsilon$. Hence $\|\mathscr{X}\| \leq \epsilon\lambda^{-1} + |1-\lambda^{-1}| \leq \epsilon(1+\epsilon)+\epsilon = \epsilon(2+\epsilon)$. □

Proof of Corollary 3.3. The corollary is easily proved by induction using the transitivity of the relation, spelled out here for convenience. Assume $A_i \overset{\epsilon}{\sim} A_{i+1}$ for all i and set $A_0 \overset{\eta_k}{\sim} A_k$, then by transitivity we have the recursion

$$\eta_{k+1} \leq \eta_k(1+\epsilon) + \epsilon.$$

For $k = 1$, noting that $\eta_1 = \epsilon$, this gives $\eta_2 = \epsilon(1+\epsilon) + \epsilon = \epsilon^2 + 2\epsilon = (1+\epsilon)^2 - 1$. For any k we have

$$\eta_{k+1} \leq \eta_k(1+\epsilon) + \epsilon = ((1+\epsilon)^k - 1)(1+\epsilon) + \epsilon = (1+\epsilon)^{k+1} - 1.$$

□

Proof of Theorem 3.5: Strong local indistinguishability in 1D,. The proof follows a similar path as the proof of Proposition 8.1 in [7] and uses local indistinguishability, see e.g. Theorem 5 from [4], or Corollary 2 in [6], which follows from the assumptions above. Thus we know that $\|\operatorname{tr}_{BC}\sigma^{ABC} - \operatorname{tr}_B \sigma^{AB}\|_1 \leq K'e^{a'l}$, with $K', a' > 0$ depending only on $r, J\beta$.
Using the notation $E_{X,Y} := e^{-H_{XY}}e^{H_X+H_Y}$ for $X, Y \subset I$ disjoint, we rewrite:

$$\begin{aligned}&(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B\sigma^{AB})^{-1} \\ &= \operatorname{tr}_{BC}[e^{-H_{ABC}}]\operatorname{tr}_B[e^{-H_{AB}}]^{-1}\operatorname{Tr}[e^{-H_{ABC}}]^{-1}\operatorname{Tr}[e^{-H_{AB}}] \\ &= \operatorname{tr}_{BC}[e^{-H_{ABC}}]e^{H_A}e^{-H_A}\operatorname{tr}_B[e^{-H_{AB}}]^{-1}\operatorname{Tr}[e^{-H_{ABC}}]^{-1}\operatorname{Tr}[e^{-H_{AB}}] \\ &= \operatorname{tr}_{BC}[e^{-H_{BC}}e^{-H_{ABC}}e^{H_A}e^{H_{BC}}]\operatorname{tr}_B[e^{-H_B}e^{-H_{AB}}e^{H_A}e^{H_B}]^{-1}\operatorname{Tr}[e^{-H_{ABC}}]^{-1}\operatorname{Tr}[e^{-H_{AB}}] \\ &= \operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\operatorname{tr}_B[\sigma^B E_{A,B}]^{-1}\underbrace{\frac{\operatorname{Tr}[e^{-H_{AB}}]\operatorname{Tr}[e^{-H_{BC}}]}{\operatorname{Tr}[e^{-H_{ABC}}]\operatorname{Tr}[e^{-H_B}]}} \\ &\equiv \operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\operatorname{tr}_B[\sigma^B E_{A,B}]^{-1} \quad\cdot\quad \lambda_{ABC}^{-1}.\end{aligned}$$

In the second line we multiplied by $\mathbb{1} = e^{H_A}e^{-H_A}$, which in the third line we can separate and pull into the partial traces, since neither of them trace out the region A. Thus we may rewrite:

$$
\begin{aligned}
&\|(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B\sigma^{AB})^{-1}-\mathbb{1}\| \\
&\qquad \le \|(\operatorname{tr}_{BC}\sigma^{ABC})(\operatorname{tr}_B\sigma^{AB})^{-1} - (\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}])(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\| \\
&\qquad\quad + \|(\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}])(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1} - \mathbb{1}\| \\
&\qquad \le \|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\|\|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\||\lambda_{ABC}^{-1} - 1| \\
&\qquad\quad + \|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\|\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\|.
\end{aligned}
$$

By Corollary 4 in [7] it holds that $\|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\| \le C$ and $\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\| \le C$, for the same constant C depending only on $r, J\beta$. Furthermore in [7][Step 2] it is proven that $\lambda_{ABC} - 1$ decays exponentially, thus by the geometric series so does $\lambda_{ABC}^{-1} - 1$, i.e. there exist $K'', a'' > 0$, depending only on $r, J\beta$ s.t. $|\lambda_{ABC}^{-1} - 1| \le K''e^{a''l}$. So it remains to bound $\|(\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]) - (\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\|$ exponentially in l. To do this we adopt a similar strategy to Step 3 in [7].
Split $B = B_1B_2$ into to halves, s.t. $|B_1| = |B_2| = l$ and write

$$
\begin{aligned}
\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\| \le & \|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_{BC}[\sigma^{BC}E_{A,B_1}]\| \\
& + \|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,B_1}] - \operatorname{tr}_B[\sigma^B E_{A,B_1}]\| \\
& + \|\operatorname{tr}_B[\sigma^B E_{A,B_1}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\|.
\end{aligned} \tag{A.1}
$$

Here we just used the triangle inequality of the operator norm twice. For the first and third summand in (A.1) use that the map $Q \mapsto \operatorname{tr}_X[\sigma^X Q]$ is a contraction in $B(\mathcal{H}_\Lambda) \to B(\mathcal{H}_{\Lambda\setminus X})$ by the Russo Dye theorem (see e.g. section 3.4 in [7]), and thus $\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_{BC}[\sigma^{BC}E_{A,B_1}]\| \le \|E_{A,BC} - E_{A,B_1}\|$, where the r.h.s is exponentially decaying in $|B_1| = l$ by Corollary 3.4 and Remark 3.5 in [7]. The same holds true for $\|\operatorname{tr}_B[\sigma^B E_{A,B_1}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\| \le \|E_{A,B_1} - E_{A,B}\|$.
For the second summand in equation (A.1) we use Proposition 8.5 in [7], which gives $\|\operatorname{tr}_{BC}[\sigma^{BC}E^*_{A,B_1}] - \operatorname{tr}_B[\sigma^B E^*_{A,B_1}]\| \le \tilde{K}e^{-\tilde{a}l}$ for some $\tilde{K}, \tilde{a} > 0$, depending only on $r, J\beta$. The same holds true for the adjoints, which is exactly the second summand. In total we have that there exist $K''', a''' > 0$ depending only on r,$J\beta$ s.t.

$$
\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\| \le K'''e^{a'''l}. \tag{A.2}
$$

Now putting all of the above together we have our desired result.

$$
\begin{aligned}
\|(\operatorname{tr}_{BC}[\sigma^{ABC}])(\operatorname{tr}_B[\sigma^{AB}])^{-1} - \mathbb{1}\| \le & \|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}]\|\|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\||\lambda_{ABC}^{-1} - 1| \\
& + \|(\operatorname{tr}_B[\sigma^B E_{A,B}])^{-1}\|\|\operatorname{tr}_{BC}[\sigma^{BC}E_{A,BC}] - \operatorname{tr}_B[\sigma^B E_{A,B}]\| \\
& \le C^2K''e^{a''l} + CK'''e^{a'''l}
\end{aligned}
$$

□

Proof of Lemma 4.5. If $\operatorname{dist}(A_1, A_2) \ge 2$, then by definition of the algebras $\mathcal{N}_{A_1}$ and $\mathcal{N}_{A_2}$ have only $\mathbb{1} = \mathcal{N}_\emptyset = \mathcal{N}_{A_1\cap A_2}$ in common. Since it holds, that $\partial A_1 \cup \partial A_2 = \partial(A_1 \cup A_2)$ their union is given by

$$
\begin{aligned}
\mathcal{N}_{A_1} \cup \mathcal{N}_{A_2} &= \mathcal{B}(\mathcal{H}_{A_1}) \otimes \mathcal{B}(\mathcal{H}_{A_2}) \otimes \mathbb{1}_{\mathcal{H}_{(A_1)^c}} \otimes \mathbb{1}_{\mathcal{H}_{(A_2)^c}} \bigotimes_{i\in I_{A_1}} \bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathcal{A}^j_{b_j} \otimes \bigotimes_{i\in I_{A_2}} \bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathcal{A}^j_{b_j} \\
&= \mathcal{B}(\mathcal{H}_{A_1\cup A_2}) \otimes \mathbb{1}_{\mathcal{H}_{(A_1\cup A_2)^c}} \otimes \bigotimes_{i\in I_{A_1\cup A_2}} \bigotimes_{j\in J^{(i)}\setminus\{0\}} \mathcal{A}^j_{b_j} \\
&= \mathcal{N}_{A_1\cup A_2}
\end{aligned}
$$

For the other case we can WLOG assume $A_1 \subset A_2$, hence $A_1 \cup A_2 = A_2$ and $A_1 \cap A_2 = A_1$. Then clearly $\mathcal{N}_{A_1} \cup \mathcal{N}_{A_2} = \mathcal{N}_{A_2} = \mathcal{N}_{A_1 \cup A_2}$. Similarly $\mathcal{N}_{A_1} \cap \mathcal{N}_{A_2} = \mathcal{N}_{A_1} = \mathcal{N}_{A_1 \cap A_2}$. $\square$

A.2 Finite-size-coarse-graining example

See **??** for an example of a 3-colorable, but non-2-colorable, infinite regular pentagonal graph, which after coarse graining subsets of diameter $d = 4$, becomes the 2 dimensional qubic, hence, 2-colorable lattice.

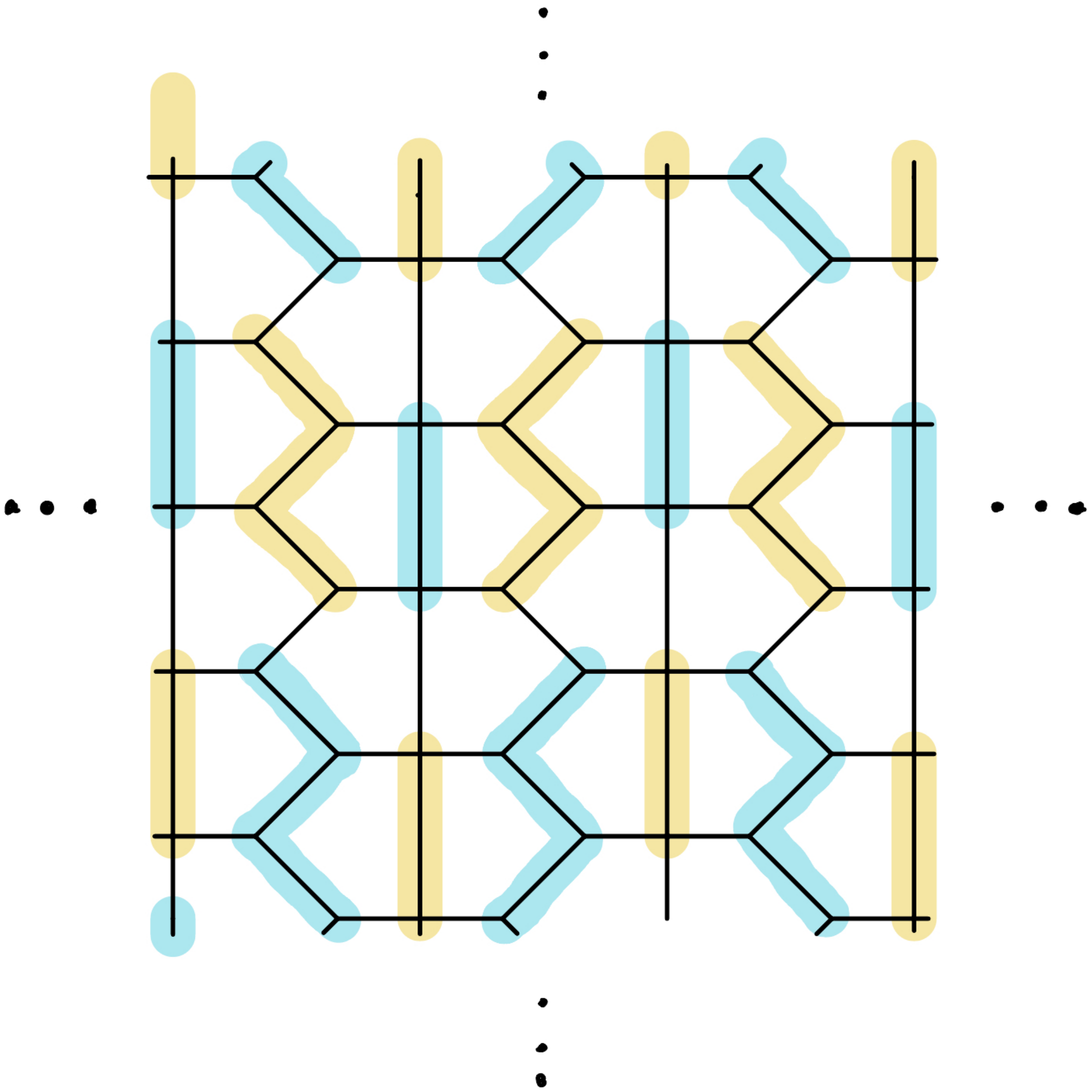

Figure A.1: Example of a section of the graph of an infinite pentagonal tiling. It is 3-colorable, but not 2-colorable, since it contains loops of odd length 5. It becomes 2-colorable, however, after finite size coarse graining of regions of size up to 4. A suitable coarse graining is given by the yellow and turquoise regions.