QUANTUM GIBBS SAMPLERS: THE COMMUTING CASE

Michael J. Kastoryano\textsuperscript{1,2} and Fernando G. S. L. Brandão\textsuperscript{3}

\textsuperscript{1} Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
\textsuperscript{2} Department of Computer Science, University College London, London UK
(Dated: September 12, 2014)

We analyze the problem of preparing quantum Gibbs states of lattice spin Hamiltonians with local and commuting terms on a quantum computer and in nature. Our central result is an equivalence between the behavior of correlations in the Gibbs state and the mixing time of the semigroup which drives the system to thermal equilibrium (the Gibbs sampler). We introduce a framework for analyzing the correlation and mixing characteristics of quantum Gibbs states and quantum Gibbs samplers, which is rooted in the theory of non-commutative $L_p$ spaces. We consider two distinct classes of Gibbs samplers, one of which being the well-studied Davies generators modelling the dynamics on the system due to weak-coupling with a large Markovian environment. We show that their gap is independent of system size if, and only if, a certain strong form of clustering of correlations holds in the Gibbs state. As concrete applications of our formalism, we show that for every one-dimensional lattice system, or for systems in lattices of any dimension at high enough temperatures, the Gibbs samplers of commuting Hamiltonians are always gapped, giving an efficient way of preparing these states on a quantum computer.

Contents

I. Introduction 2
   A. Summary of results 3
   B. Formal framework 4
   C. Dynamics 6

II. Conditional expectations 7
   A. Local Liouvillian projectors 8
   B. Minimal conditional expectations 9

III. Gibbs states and Gibbs samplers 10
   A. Davies generators 11
   B. Heat-Bath generators 13

IV. Decay of Correlations 15
   A. Local indistinguishability 18

V. Main Results 20
   A. Strong clustering implies gapped Gibbs sampler 21
   B. Gapped Gibbs sampler implies local clustering 26

VI. One dimensional models 29

VII. The High temperature phase 33

VIII. Outlook 34

References 37

Appendix 39
I. INTRODUCTION

Physical systems in nature very often are in thermal equilibrium. Statistical mechanics provides a microscopic theory justifying the relevance of thermal states of matter. However, fully understanding the ubiquity of this class of states from the laws of quantum theory remains an important topic in theoretical physics. Indeed the problem of thermalization in quantum systems has recently generated a lot of interest, in part because of the new set of tools available from the field of quantum information theory [4–6]. The problem can be broken up into two sets of questions: (i) under what conditions does a system thermalize in the long time limit, and (ii) assuming a system does eventually thermalize, how much time does one have to wait before this is so? Our work is concerned with the latter question in the setting of quantum lattice spin systems.

The problem of the speed of thermalization is also of practical relevance in the context of quantum simulators, where one wants to emulate the behavior of a real physical system by simulating an idealization of it on a classical or quantum computer. Given that many of the systems which one would want to simulate are thermal, it is an important task to develop simulation and sampling algorithms that can prepare large classes of thermal states of local Hamiltonians. A large body of work has already been done in the classical problem, starting with the development and analysis of Gibbs sampling algorithms of lattice systems called Glauber dynamics, which include the Metropolis and Heat-bath algorithms as special cases. Nowadays, there are dozens of variants of classical Gibbs samplers which find applications in a variety of fields of physics, computer science, and theoretical chemistry [3]. A peculiar feature of many of these algorithms is that they often provide reliable results in practice, but a systematic certification of their accuracy and efficiency is often elusive. Although a very hard problem in general, the problem of estimating the convergence time of classical Gibbs samplers has seen a number of very important breakthroughs in the past few decades. These methods, which are rooted in the analysis of Markov Chain mixing times, are also closely related to problems in combinatorial optimization, with applications in numerous fields; see [7, 8] for recent surveys.

The rigorous analysis of classical Gibbs sampling algorithms was closely tied to the development of a rigorous theory of thermodynamic quantities, which consists in analyzing the behavior of finite systems of larger and larger size in order to infer the physics in the thermodynamic limit. This is made possible because the influence of the infinite system on a subregion of the lattice can be encoded in the boundary of the region in a one-to-one fashion. This identity is often called Dobrushin-Lanford-Ruelle (DLR) theory [22, 23], and is the technical cornerstone of a lot of the early rigorous results in lattice Gibbs states. Using DLR theory, it has been possible to properly characterize phase transitions in classical lattice spin models. A series of seminal results have shown that the existence of a unique Gibbs state in the thermodynamic limit is related both to exponentially decaying correlations between local observables, and rapidly mixing dynamics to thermal equilibrium [18, 33, 36, 38, 39, 42, 44]. Furthermore, it can be shown that if a Gibbs sampler decays polynomially fast (with a polynomial degree sufficiently large, of the order of the dimension), then the convergence is in fact exponential.

The purpose of this work is to introduce a framework for analyzing Gibbs samplers in the setting of quantum systems, and to explore to what extent the classical results (equivalences) generalize to quantum lattice spin systems. We build upon a growing body of work on the theory of mixing times of quantum channels [20, 21].

Throughout this work we will restrict ourselves to commuting Hamiltonians. It is worth noting that the case of commuting Hamiltonians does not effectively reduce to the classical picture, as highly non-classical phenomenon such as topological quantum order can occur. In particular, this setting encompasses all stabilizer Hamiltonians, which have been a useful
playground for exploring unique quantum features of many-body systems.

The physical relevance of our results is twofold. First, we consider a class of Gibbs samplers (called Davies generators \[27\]) which can be derived from the weak coupling of a finite quantum system to a large thermal bath. Hence our analysis pertains to the time it takes to reach thermal equilibrium in naturally occurring systems. Secondly, all Gibbs samplers which we consider are local and bounded maps, and therefore can be prepared by dissipative engineering or digital simulation on quantum computers, or quantum simulators \[28\]. Thus, our results can also be understood as an analysis of the efficiency of quantum Gibbs sampling.

Previous efforts at proposing quantum Gibbs samplers have typically fallen into two categories: ones that have a certified runtime, that is typically exponential in the system size \[9, 10, 12\], and ones that can be implemented on a quantum computer, but where no bounds on their efficiency are currently known \[11, 29, 30\]. Our Gibbs samplers have the benefit that they are very simple, and hence amenable to analytic study. We note however they have the drawback of only being properly defined for Hamiltonians with local commuting terms.

---

A. Summary of results

In order to present our main results, we need to spend some time defining the framework and the quantities involved, which are rooted in the theory of non-commutative \(L_p\) spaces. This has value in its own right, as a systematic study of thermal states of quantum lattice systems in the spirit of DLR theory has not yet been undertaken. However we only achieve it partially, and comment along the way on the limitations of a full generalization of the classical theory in many aspects.

After a brief recollection of the formal framework and of the setting of dissipative quantum systems, we introduce a class of maps called conditional expectations which serve as local quasi-projectors onto the Gibbs state of the system. These maps play a central role in our analysis. We identify two special classes of conditional expectations: the first is purely dynamical and inherits many of the properties of the underlying dissipative generator, the second is purely static, and only depends on the reference (Gibbs) state of the system. We prove that both are local maps when the underlying Hamiltonian is commuting (Propositions 7 and 8).

We go on to define quantum lattice Gibbs states, and introduce two classes of Gibbs samplers: Davies Generators and Heat-Bath Generators. We will also call them Davies Gibbs sampler and Heat-Bath Gibbs sampler, respectively. The Davies generators are obtained from a canonical weak-coupling between a system and a large thermal bath, whereas the heat-bath generators are constructed in a manner reminiscent of the classical heat-bath Monte-Carlo algorithm. The basic properties of these maps are summarized and collected in Propositions 9 and 10.

The main purpose of the paper is to show an equivalence between the convergence time of the Gibbs sampler and the correlation behaviour of the Gibbs state. The analogous classical equivalence builds very heavily on the DLR theory of boundary conditions. As a naive extension of the DLR theory does not hold for quantum systems \[48\], we are lead to define a different notion of clustering (which we call strong clustering), that somehow incorporates the strong mixing (or complete analyticity) condition for classical systems. This condition relies on a conditional covariance, which restricts attention to a subset of the lattice. The conditional variance depends on a specific choice of conditional expectation. We show that the strong clustering condition implies the standard clustering of correlation (which we call weak clustering) condition that is usually considered in quantum lattice systems (Corollary 25). We also flesh out the connection between our notions of clustering and the local indistinguishability of Gibbs states that differ only by a distant perturbation \[17\].
Having introduced the framework of Gibbs samplers, and defined what we mean by clustering of correlations, we set out to prove the main theorem of our paper (see Theorems [21] and [24] for a precise formulation):

**Theorem 1 (informal)** Both the Davies Gibbs sampler and Heat-Bath Gibbs sampler of commuting local Hamiltonians have a gap independent of the system size if, and only if, the Gibbs state satisfies strong clustering.

The gap of a Gibbs sampler is defined in Def. [19] and is related to the rate of convergence of the Gibbs sampler to the Gibbs state.

We prove the necessity and sufficiency parts of the theorem separately, as they require quite distinct proof techniques. The only if statement is proved via methods very reminiscent of the analogous classical result [39]. The main idea of the proof is to consider the variational characterization of the spectral gap, and show, by a clever manipulation of conditional variances, that the gap of the Gibbs sampler restricted to a subsystem of minimum side length $L$ is roughly the same as the gap restricted to a subsystem of side length $2L$. Then using the same argument iteratively shows that the gap of the dynamics is approximately scale invariant. The if part of the statement, on the other hand, exploits methods from quantum information theory and quantum many-body theory. In particular, we find a mapping of our problem to properties of frustration-free gapped local Hamiltonians, and apply the so-called detectability lemma of [41].

Our main theorem becomes especially compelling for one dimensional lattice systems, where it was shown by Araki [45] that Gibbs states always satisfy weak clustering. We prove that weak clustering and strong clustering are equivalent for one dimensional systems, getting that all Gibbs samplers in this case are gapped. Exploring our mapping between Gibbs samplers and local Hamiltonians, we also prove that at high enough temperature (independent of the size of the system) the Heat-Bath Gibbs sampler is gapped. We then obtain:

**Theorem 2 (informal)** Both the Davies and the Heat-Bath Gibbs samplers give polynomial-time quantum algorithms for preparing the Gibbs state of every 1D commuting Hamiltonian at any constant temperature. The Heat-Bath Gibbs sampler gives a polynomial-time quantum algorithm for preparing every commuting Hamiltonian at temperatures above a critical temperature (that is independent of the system size).

We note that since Gibbs states of 1D commuting Hamiltonians are matrix-product operators, one can prepare them efficiently on a quantum computer using e.g. [24] (in fact this is also true for general non-commuting 1D Gibbs states [25]); here we only show another way of preparing them, which might be more resilient to noise in some circumstances.

Finally, we discuss extensions and further implications of our results. We conclude with few important questions and conjectures. We connect the conjecture with the problem of self-correcting quantum memories in two dimensional systems.

### B. Formal framework

This paper concerns quantum spin lattice systems. Although the results presented here can be extended to more general graphs, we will restrict our attention to spins living on a $d$-dimensional finite square lattice $\Lambda \subset \mathbb{Z}^d$. Lattice subsets will be denoted by upper case Roman letters; eg. $A,B \subset \Lambda$. The complement of a set $A \subset \Lambda$ will be written $A^c$. The cardinality of set $A$ will be written $|A|$. The global Hilbert space is seen to be $\mathcal{H}_\Lambda = \bigotimes_{x \in \Lambda} \mathcal{H}_x$. We assume the local Hilbert space dimension to be bounded; i.e. $\dim(\mathcal{H}_x) < \infty$. We denote the set of bounded operators on $\mathcal{H}_\Lambda$ by $\mathcal{B}_\Lambda \equiv \mathcal{B} : \mathcal{H}_\Lambda \rightarrow \mathcal{H}_\Lambda$. 

Lemma 3 \cite{1, 2} Let \( A_\Lambda \subset B_\Lambda \). The set of positive semi-definite operators is denoted \( S_\Lambda = \{ X \in A_\Lambda, X \geq 0, \text{tr} [X] = 1 \} \), and its full rank subset is analogously denoted \( S^+_\Lambda \). The elements of \( A_\Lambda \) will be called observables, and will always be denoted with lower case Latin letters \((\rho, \sigma)\), and the elements of \( S_\Lambda \) will be called states, and will be denoted with lower case Greek letters \((\rho, \sigma)\). The support of a local observable \( f \in A_\Lambda \) will be written \( \Sigma_f \). We will make use of the modified partial trace \( \text{tr}_A : B_\Lambda \to B_\Lambda \), which would read in the more traditional quantum information theory notation: \( \text{tr}_A : f \mapsto \text{tr}_A(f) \otimes \mathbb{I}_A \). For \( i, j \in \Lambda \), we denote by \( d(i, j) \) the Euclidean distance in \( \mathbb{Z}^d \). The distance between two sets \( A, B \subset \Lambda \) is \( d(A, B) = \min \{ d(i, j) \}, i \in A, j \in B \).

Many of the tools used in this work can be traced back to the theory of non-commutative \( \mathbb{L}_p \) spaces \cite{1, 2}. The central property of the non-commutative \( \mathbb{L}_p \) spaces summarized below, is that the norm as well as the scalar product is weighted with respect to some full rank reference state \( \rho \in S^+_{\Lambda} \). The non-commutative \( \mathbb{L}_p \) spaces are equipped with a weighted \( \mathbb{L}_p \)-norm which, for any \( f, g \in A_\Lambda \) and some fixed \( \rho \in S^+_{\Lambda} \), is defined as

\[
\| f \|_{p, \rho} = \text{tr} \left[ \| \rho^{\frac{1}{p}} f \rho^{\frac{1}{q}} \|_p \right]^\frac{1}{p}, \quad 1 \leq p \leq \infty.
\]

Similarly, the \( \rho \)-weighted non-commutative \( \mathbb{L}_p \) inner product is given by

\[
\langle f, g \rangle_{\rho} = \text{tr} \left[ \sqrt{\rho} f \sqrt{\rho} g \right].
\]

We summarize the basic properties of non-commutative \( \mathbb{L}_p \) spaces in the following lemma.

**Lemma 3** \cite{1, 2} Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \). The non-commutative \( \mathbb{L}_p \) spaces satisfy:

1. **(Natural ordering of the \( \mathbb{L}_p \) norms)** Let \( f \in A_\Lambda \) and \( \sigma \in S^+_\Lambda \), then for any \( p, q \in [1, \infty) \) satisfying \( p \leq q \), we get \( \| f \|_{p, \rho} \leq \| f \|_{q, \rho} \).

2. **(Hölder-type inequality)** Let \( f, g \in A_\Lambda \) and \( \rho \in S^+_\Lambda \), then for any \( p, q \in [0, \infty) \) satisfying \( 1/p + 1/q = 1 \),

\[
| \langle f, g \rangle_{\rho} | \leq \| f \|_{p, \rho} \| g \|_{q, \rho}.
\]

3. **(Duality)** Let \( f \in A_\Lambda \) and \( \rho \in S^+_\Lambda \), then for any \( p, q \in [0, \infty) \) satisfying \( 1/p + 1/q = 1 \),

\[
\| f \|_{p, \rho} = \sup \{ \langle f, g \rangle_{\rho}, g \in A_\Lambda, \| g \|_{q, \rho} \leq 1 \}.
\]

In the remainder of the paper, unless specified otherwise, we will always be working with \( \mathbb{L}_p \) norms and inner products. The reference state should always be clear from the context, and will almost always be the Gibbs state of some local commuting Hamiltonian (see Sec. III).

Finally, we will also make extensive use of the \( \mathbb{L}_p \) covariance of a state \( \rho \in S_\Lambda \), which is defined for any \( f, g \in A_\Lambda \) as

\[
\text{Cov}_\rho(f, g) = | \langle f, g \rangle_{\rho} - \text{tr} [\rho f] \text{tr} [\rho g] |.
\]

Similarly, the variance is given as \( \text{Var}_\rho(f) = \text{Cov}_\rho(f, f) \). The covariance and the variance are always positive, and that they are invariant under the transformation \( g \mapsto g + c \mathbb{I} \), for any \( c \in \mathbb{R} \).
C. Dynamics

The time evolution of an observable \((f_t \in \mathcal{A}_\Lambda)\) will be described by one-parameter semi-groups of completely positive trace preserving maps (cpt-maps), whose generator (Liouvillian) can always be written in standard Lindblad form

\[
\partial_t f = \mathcal{L}(f_t) \equiv i[H,f] + \mathcal{D}(f),
\]

with

\[
\mathcal{D}(f) = \sum_i L_i^\dagger f L_i - \frac{1}{2} \{ L_i^\dagger L_i, f \} + \mathcal{D}_Z(f),
\]

where \(\{L_i\} \in \mathcal{B}_\Lambda\) are Lindblad operators and \(H \in \mathcal{A}_\Lambda\) is a Hamiltonian operator. We will denote the semigroup generated by \(\mathcal{L}\) by \(T_t = \exp(t\mathcal{L})\). This evolution corresponds to the Heisenberg picture, which specifies the dynamics on observables rather than on states. We denote the dual generator, with respect to the Hilbert-Schmidt inner product, by \(\mathcal{L}^\ast\) which amounts to the evolution of states, i.e. the Schrödinger picture. The trace preserving condition ensures that \(\mathcal{L}(1) = 0\). If in addition \(\mathcal{L}^\ast(1) = 0\), then the dynamics are said to be unital.

A Liouvillian \(\mathcal{L} : \mathcal{B}_\Lambda \to \mathcal{B}_\Lambda\) is said to be primitive if it has a unique full-rank stationary state (i.e. a unique full rank state \(\rho\) s.t. \(\mathcal{L}^\ast(\rho) = 0\)). A Liouvillian is said to be reversible (or satisfy detailed balance) with respect to a state \(\rho \in \mathcal{S}_\Lambda\) if for every \(f, g \in \mathcal{A}_\Lambda\),

\[
\langle f, \mathcal{L}(g) \rangle_\rho = \langle \mathcal{L}(f), g \rangle_\rho
\]

If \(\mathcal{L}\) is reversible with respect to \(\rho\) then \(\rho\) is a stationary state of \(\mathcal{L}\) [20]. \(\mathcal{L}\) is said to be \(r\)-local if it can be written as

\[
\mathcal{L}_\Lambda(f) = \sum_{Z: Z \cap \Lambda \neq 0} i[H_Z, f] + \mathcal{D}_Z(f)
\]

where \(H_Z\) and \(\mathcal{D}_Z\) only have support on \(Z \subset \Lambda\) and \(|Z| < r\) for some constant \(r < \infty\). \(r\) will be referred to as the range of the Liouvillian. Note that in Eqn. (9), the summation is over all local terms that have non-zero intersection with \(\Lambda\). When considering restricted Liouvillians acting on subsets \(A \subset \Lambda\), we will want to label them as

\[
\mathcal{L}_A(f) = \sum_{Z: Z \cap A \neq 0} i[H_Z, f] + \mathcal{D}_Z(f)
\]

Note in particular that \(\mathcal{L}_A\) acts on \(A\) plus a “buffer” region around \(A\) whose width is determined by the range of the local terms \(|Z| < r\). We now define two auxiliary properties of local Liouvillians:

**Definition 4** Let \(\Lambda\) be a finite lattice in \(\mathbb{Z}^d\). Let \(\mathcal{L}_\Lambda : \mathcal{B}_\Lambda \to \mathcal{B}_\Lambda\) be a local Liouvillian. We say that \(\mathcal{L}_\Lambda\) is locally primitive if for any subset \(A \subset \Lambda\), \(\mathcal{L}_A(f) = 0\) implies that \(f\) has non-trivial

\[1\] The notion of locality is treated somewhat vaguely here. We will be much more specific when discussing the locality properties of Gibbs samplers.
support only on \( A^c \). Similarly, \( L_\Lambda \) is locally reversible with respect to \( \rho \) if for any \( A \subset \Lambda \) and every \( f, g \in A_\Lambda \),

\[
\langle f, L_\Lambda(g) \rangle_\rho = \langle L_\Lambda(f), g \rangle_\rho
\]

**Definition 5** Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \). Let \( L_\Lambda : B_\Lambda \rightarrow B_\Lambda \) be a local Liouvillian. We say that \( L_\Lambda \) is frustration free if for all \( A \subset \Lambda \), \( \rho \) is a stationary state of \( L_\Lambda \) whenever \( \rho \) is a stationary state of \( L_\Lambda \).

To conclude this subsection, we recall the definition of the spectral gap of a Liouvillian. Let \( L_\Lambda : B_\Lambda \rightarrow B_\Lambda \) be a primitive reversible Liouvillian with stationary state \( \rho \), then the spectral gap \( \lambda_\Lambda \) of \( L_\Lambda \) is given by

\[
\lambda_\Lambda = \inf_{f \in A_\Lambda} \frac{-\langle f, L_\Lambda(f) \rangle_\sigma}{\text{Var}_\Lambda(f)}
\]

The significance of \( \lambda_\Lambda \) follows from Theorem 2.2 of [20], where it is shown that for every state \( \sigma \),

\[
\| e^{t L_\Lambda^*} (\sigma) - \rho \|_1 \leq \lambda_{\min}(\rho)^{-1/2} e^{-\lambda_\Lambda t},
\]

with \( \rho \) the fixed point of \( L_\Lambda^* \) and \( \lambda_{\min}(\rho) \) its minimum eigenvalue. In this paper we will be interested in the case where \( \rho \) is the thermal state of a local Hamiltonian, in which case \( \lambda_{\min}(\rho) = e^{O(\Lambda)} \), with \( |\Lambda| \) the number of particles of the lattice \( \Lambda \). Thus the Liouvillian converges to a good approximation of its fixed point in time of order \( |\Lambda|/\lambda_\Lambda \).

In terms of spectral theory, the spectral gap of a primitive reversible Liouvillian is given by the smallest non-zero eigenvalue of \( -L_\Lambda \) (in the \( L^p \) space associated to \( \rho \)). In Section V we will introduce a generalization of the spectral gap to subsets of the lattice.

### II. CONDITIONAL EXPECTATIONS

In this section we introduce a set of maps called **conditional expectations**, which we denote suggestively by \( E \). These maps will later on play the role of local quasi-projectors onto the Gibbs state. In Refs. [31–33] one variant of conditional expectations was studied in very much the same context as we do here, where ergodic properties of Gibbs samplers were the main focus. Also, Petz considered a similar set of maps in the context of coarse graining operations [34].

**Definition 6 (conditional expectations)** Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \), and let \( \rho \in S^+_\Lambda \) be a full rank state. Then, we call \( E \) a conditional expectation of \( \rho \) if it satisfies the following properties:

1. (complete positivity) \( E \) is completely positive and unital.
2. (consistency) For any \( f \in A_\Lambda \), \( \text{tr} [\rho E(f)] = \text{tr} [\rho f] \)
3. (reversibility) For any \( f, g \in A_\Lambda \), \( \langle E(f), g \rangle_\rho = \langle f, E(g) \rangle_\rho \)
4. (monotonicity) For any \( f \in A_\Lambda \) and \( n \in \mathbb{N} \), \( \langle E^n(f), f \rangle_\rho \geq \langle E^{n+1}(f), f \rangle_\rho \).
The consistency condition is reminiscent of the classical conditional expectation (see e.g.), while the reversibility condition can be understood as a form of detailed balance with respect to the state $\rho$. The role of monotonicity is not a priori clear, but will turn out to be necessary for the applications which we have in mind. In practice, we will often be interested in conditional expectations that are also projectors. In that case, the monotonicity condition holds with equality.

We will describe two examples of conditional expectations which are especially useful in the context of Gibbs samplers. As we will see below and in Section III, in addition to satisfying properties $1−4$ above, they will inherit locality properties from some lattice Hamiltonian or Liouvillian.

A. Local Liouvillian projectors

Let $\Lambda \subset \mathbb{Z}^d$, and consider an local primitive Liouvillian $\mathcal{L}_\Lambda = \sum_{Z \cap \Lambda \neq 0} \mathcal{L}_Z$ with stationary state $\rho \in S^+_\Lambda$. Then the local Liouvillian projector associated with $\mathcal{L}$ on $A$ is given by

$$\mathbb{E}_\mathcal{L}^A(f) = \lim_{t \to \infty} e^{t\mathcal{L}_A}$$

(12)

If $\mathcal{L}$ is locally primitive then $\mathbb{E}_\mathcal{L}^A(f)$ has support only on $A^c$. If $\mathcal{L}$ is frustration free, then $\mathbb{E}_\mathcal{L}^A$ is a conditional expectation with respect to the stationary state of $\mathcal{L}$. Indeed

Proposition 7 Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\mathcal{L}_\Lambda : \mathcal{B}_\Lambda \to \mathcal{B}_\Lambda$ be frustration free and locally reversible with respect to $\rho \in S^+_\Lambda$. Then for any $A \subset \Lambda$, $\mathbb{E}_\mathcal{L}^A$ is a conditional expectation with respect to $\rho$.

Proof: Complete positivity follows by construction, since for any $t \geq 0$, $e^{t\mathcal{L}}$ is a completely positive unital map. Consistency follows from frustration freedom. Indeed, assume $\rho$ is a stationary state of $\mathcal{L}$, then by frustration freedom, for any $A \subset \Lambda$ and $f \in \mathcal{A}_\Lambda$,

$$\text{tr} \left[ \rho e^{t\mathcal{L}_A} (f) \right] = \text{tr} \left[ e^{t\mathcal{L}_A} (\rho f) \right] = \text{tr} [\rho f]$$

Reversibility of $\mathbb{E}_\mathcal{L}^A$ follows directly from local reversibility of $\mathcal{L}_\Lambda$. Finally, monotonicity can be seen to hold universally with equality from the projector property. For any $A \subset \Lambda$, note that

$$(\mathbb{E}_\mathcal{L}^A)^2(f) = \lim_{t \to \infty} e^{t\mathcal{L}_A} \left( \lim_{t \to \infty} e^{t\mathcal{L}_A} (f) \right)$$

$$= \lim_{t \to \infty} \lim_{t' \to \infty} e^{(t+t')} \mathcal{L}_A (f)$$

$$= \mathbb{E}_\mathcal{L}^A(f)$$

It immediately follows that $\langle (\mathbb{E}_\mathcal{L}^A)^m(f), f \rangle_\rho = \langle (\mathbb{E}_\mathcal{L}^A)^m(f), f \rangle_\rho$ for all $m, n \in \mathbb{N}$. □

It is clear from Eqn. (12) that if $\mathcal{L}$ is local, then $\mathbb{E}_\mathcal{L}^A$ acts only on $A$ plus a finite buffer region around $A$ whose width is upper bounded by the range of $\mathcal{L}$. The expectation value with respect to the full system, $\rho : f \to \text{tr} [\rho f]$ is equivalent to the local Liouvillian projector onto the whole system $\Lambda$ when $\mathcal{L}_\Lambda$ is primitive.
B. Minimal conditional expectations

The minimal conditional expectation $\mathbb{E}_A^\rho$ is as its name suggests meant to minimally affect the observables outside of $A$ while still satisfying all four conditions of Definition 6. This map has been considered previously, under the name corse graining map in Ref. [34] and block spin flip map in Ref. [31].

Let $\rho \in S_\Lambda^+$ be a full rank state on the lattice $\Lambda$, and let $A \subset \Lambda$, then the minimal conditional expectation of $\rho$ on $A$ is given by

$$\mathbb{E}_A^\rho(f) = \text{tr}_A[\eta_\Lambda^\rho f \eta_A^\rho], \quad (13)$$

where $\eta_A^\rho = (\text{tr}_A[\rho])^{-1/2}\rho^{1/2}$. Recall that $\text{tr}_A$ is not the usual partial trace, but acts as a map from $B_\Lambda \rightarrow B_A$. A moment of thought shows that $\mathbb{E}_A^\rho(f)$ is a hermitian operator on the full system, which acts as the identity on subsystem $A$, and non-trivially on the rest of the system.

It is illuminating to note that $\mathbb{E}_A^\rho$ reduces to the classical conditional expectation of $\rho$ when the input observable is taken diagonal in the eigenbasis of $\rho$. Finally,

**Proposition 8** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\rho \in S_\Lambda^+$ and let $A \subset \Lambda$, then $\mathbb{E}_A^\rho$ is a conditional expectation with respect to $\rho$.

**Proof**: Complete positivity follows directly from the explicit form of Eqn. (13), as a convolution of two completely positive maps. In order to show the other properties, we note an important property of the partial trace. Denote $\rho_{A^c} \equiv \text{tr}_A[\rho]$, then

$$\text{tr}_A[\rho_{A^c}^{-1/2}\rho^{1/2}\rho_{A^c}^{-1/2}] = \rho_{A^c}^{-1/2}\text{tr}_A[\rho^{1/2}\rho_{A^c}^{-1/2}]. \quad (14)$$

In particular, this implies that

$$\mathbb{E}_A^\rho(1) = \rho_{A^c}^{-1/2}\text{tr}_A[\rho]\rho_{A^c}^{-1/2} = 1, \quad (15)$$

showing unitality of $\mathbb{E}_A^\rho$. Consistency follows simply from Eqn. (14):

$$\text{tr}[\rho \mathbb{E}_A^\rho(f)] = \text{tr}[\rho \Gamma_{A^c}^{-1}(\text{tr}_A[\Gamma_{A^c}(f)])] = \text{tr}[\text{tr}_A[\Gamma_{A^c}^{-1}(\rho)]\text{tr}_A[\Gamma_{A^c}(f)]] = \text{tr}[\text{tr}_A[\Gamma_{A^c}(f)] = \text{tr}[f], \quad (16)$$

where for simplicity of notation we write $\Gamma_{A^c}(f) = \rho^{1/2}f\rho^{1/2}$. Reversibility follows by similar arguments.

$$\langle \mathbb{E}_A^\rho(f), g \rangle_{\rho} = \text{tr}[\text{tr}_A[\Gamma_{A^c}^{-1}(\Gamma_{A^c}(f))]\Gamma_{A^c}(g)] = \text{tr}[\text{tr}_A[\Gamma_{A^c}(f)]\text{tr}_A[\Gamma_{A^c}^{-1}(\Gamma_{A^c}(g))]] = \text{tr}[\rho^{1/2}f\rho^{1/2}\text{tr}_A[\rho_{A^c}^{-1/2}\rho^{1/2}g\rho_{A^c}^{-1/2}]] = \langle f, \mathbb{E}_A^\rho(g) \rangle_{\rho}. \quad (17)$$

We now show monotonicity of $\mathbb{E}_A^\rho$. For any $A \subset \Lambda$, $\mathbb{E}_A^\rho(\cdot)$ is a completely positive map, as can be immediately seen by inspection. Note that $\mathbb{E}_A^\rho(1) = 1$, hence it is also unital. But this in turn implies that its spectral radius is 1. Furthermore, by reversibility, $\Gamma_{A^c}^{-1/2}\mathbb{E}_A^\rho(\cdot)\Gamma_{A^c}^{-1/2}$ is...
hermitian, so its spectrum is real and its left and right eigenvectors are the same. But since the spectrum of a matrix is unchanged by a similarity transform, we can write the spectral radius of $E^\rho_A$ as

$$1 = \sup_{f = f^\dagger} \frac{\text{tr} \left[ \Gamma_{\rho}^{1/2}(f) E^\rho_A(\Gamma_{\rho}^{-1/2}(f)) \right]}{\text{tr} [f^2]} \quad (18)$$

$$= \sup_{g = g^\dagger, f = f^\dagger} \frac{\text{tr} \left[ \Gamma_{\rho}(g) E^\rho_A(g) \right]}{\text{tr} [\Gamma_{\rho}(g) g]} \quad (19)$$

$$= \|E^\rho_A\|_{2 \to 2, \rho}^2 \quad (20)$$

where the second line follows because $\Gamma_{\rho}$ is hermiticity preserving. In particular, this also implies that $\langle f - E^\rho_A(f), f \rangle_{\rho} \geq 0$.

Now let $\tilde{E}_A(f) = \Gamma_{\rho}^{1/2}(E^\rho_A(\Gamma_{\rho}^{-1/2}(f)))$ and define $\Phi(f) := \tilde{E}_A^{1/2}(\Gamma_{\rho}^{-1/2}(f))$, then

$$\sup_{f = f^\dagger} \frac{\langle E^\rho_A(f), E^\rho_A(f) \rangle_{\rho}}{\langle E^\rho_A(f), f \rangle_{\rho}} = \sup_{f = f^\dagger} \frac{\text{tr} \left[ \Gamma_{\rho}^{1/2}(f) \tilde{E}_A^{1/2}(\Gamma_{\rho}^{-1/2}(f)) \right]}{\text{tr} \left[ \Gamma_{\rho}^{1/2}(f) \tilde{E}_A(\Phi(f)) \right]} \quad (21)$$

$$= \sup_{f = f^\dagger} \frac{\text{tr} \left[ \Phi(f) \tilde{E}_A(\Phi(f)) \right]}{\text{tr} [\Phi(f) \Phi(f)]} \quad (22)$$

$$\leq \sup_{g = g^\dagger} \frac{\text{tr} \left[ g \tilde{E}_A(g) \right]}{\text{tr} [f^2]} = 1 \quad (23)$$

Thus, for all $f = f^\dagger$,

$$\langle E^\rho_A(f), E^\rho_A(f) \rangle_{\rho} \leq \langle E^\rho_A(f), f \rangle_{\rho} \quad (24)$$

By iteration, this then shows monotonicity of $E^\rho$. 

**Remark:** $E^\rho_A$ is not a projector. However, if we take the limit of infinite iterations of the minimal conditional expectation of $\rho$ on $A \subset \Lambda$: $\lim_{n \to \infty} (E^\rho_A)^n$ then we recover a local projector satisfying the monotonicity condition with equality. The minimal conditional expectation has the benefit that it is uniquely defined for any full-rank state $\rho$. In other words, it does not invoke any dynamical description of the state $\rho$ (via a Liouvillian) as is the case for the local Liouville projector. On the other hand, it has the disadvantage that the map $E^\rho_A$ can potentially not exhibit any locality properties. We will see in the next section that in the special case when $\rho$ is the Gibbs state of a commuting Hamiltonian, then $E^\rho_A$ also acts on $A$ plus a buffer region around $A$; in the same way as $E^\rho_A$.

### III. GIBBS STATES AND GIBBS SAMPLERS

The primary purpose of this paper is to analyze the efficient preparation of Gibbs states of commuting Hamiltonians on finite dimensional lattices. In this section we introduce the notion of lattice Gibbs states in the quantum setting, and we describe two classes of
Gibbs Samplers (Liouvillians) which generate the Gibbs state of a given local commuting Hamiltonian.

Given a finite lattice \( \Lambda \in \mathbb{Z}^d \), let \( \Phi_A : \Lambda \to A \) be an \( r \)-local bounded potential: i.e. for any \( j \in \Lambda \), \( \Phi_A(j) \) has support on a ball of radius \( r \) around site \( j \), and \( ||\Phi_A(j)|| < K \) for some constant \( K < \infty \). For any subset \( A \subset \Lambda \), the Hamiltonian \( H_A \) is given by

\[
H_A = \sum_{j \in A} \Phi_A(j)
\]  

(25)

We say that \( \Phi_A \) is a commuting potential if for all \( i, j \in \Lambda \), \( [\Phi_A(i), \Phi_A(j)] = 0 \). We will need to introduce some set notation at this point to characterize boundaries of sets. Let \( A \subset \Lambda \), and let \( \Phi_A \) be a bounded local potential. Then we denote the (outer) boundary of \( A \):

\[
\partial A = \{ j \in \Lambda | \text{supp}(\Phi(j)) \cap A \neq 0, j \notin A \},
\]

(26)

We will also denote the set \( A_{\partial} = A \cup \partial A \). Clearly \( H_A \) has support on \( A_{\partial} \). It is important to note at this point that the boundary conditions of \( H_A \) are encoded in the potential \( \Phi_A \).

We assume that the systems we are working with have a natural prescription for boundary conditions; for instance periodic or closed, as is the case for example with the planer code [13]. Open boundary conditions are often problematic as they usually break primitivity at the boundary.

The Gibbs (thermal) state of the full Lattice \( \Lambda \) is

\[
\rho_{\Lambda} = \frac{e^{-\beta H_{\Lambda}}}{\text{tr}[e^{-\beta H_{\Lambda}}]} 
\]

(27)

Restricted Gibbs states will similarly be given by

\[
\rho_A = \frac{e^{-\beta H_A}}{\text{tr}[e^{-\beta H_A}]},
\]

(28)

for any \( A \subset \Lambda \), where \( \rho_A \in S^+_\beta \), when \( H_A \) is constructed from a commuting potential \( \Phi \). Unless otherwise specified, \( \rho \) will always be the Gibbs state of the full system.

For classical spin systems, Gibbs states of local hamiltonians restricted to finite subsets of the square lattice can be unambiguously related to the Gibbs state in the thermodynamic limit, by parametrizing the effect of the ambient infinite system in the form of boundary conditions on the finite system. This procedure is often referred to as the DLR theory of boundary conditions. DLR theory specifically shows that the contribution of the infinite ambient environment constitutes a convex set of (boundary) conditions [54]. Thus optimization over the set of boundary conditions can be restricted to particular (pure) configurations. This simple fact allows for remarkable simplifications when comparing properties of systems with varying lattice sizes; as is beautifully illustrated in Refs. [18, 42, 43]. It turns out that these equivalences break down in the case of quantum systems (see Ref. [38] for a detailed discussion and counter-examples). In this work, we circumvent DLR theory by working with conditional expectations of the Gibbs state. The price we pay is that our results are weaker in many instances that the analogous classical ones.

We now turn to the description of Gibbs samplers of commuting Hamiltonians.

A. Davies generators

The dissipative dynamics resulting from the weak (or singular) coupling limit of a system coupled to a large heat bath are often called Davies generators [37] or thermal Liouvillians
where the local dissipative elements are given by

\[ H_{\text{tot}} = H_{\Lambda} + H_B + \sum_{k \in \Lambda, \alpha} S_{\alpha(k)} \otimes B_{\alpha(k)} \]  

(29)

We can choose the couplings \( S_{\alpha(k)} \) and \( B_{\alpha(k)} \) to be hermitian, and in practice one usually wants the set \( \{ S_{\alpha(k)} \} \) to span the local algebra of site \( k \) (for example the Pauli operators for spin systems). Assuming the bath is in a Gibbs state, taking the coupling terms to zero, and tracing out the bath yields the Davies generators in their standard form [37]:

\[ \mathcal{L}^D_{\Lambda}(f) = i[H_{\Lambda}, f] + \sum_{k \in \Lambda} \mathcal{L}^D_k(f), \]

(30)

where the local dissipative elements are given by

\[ \mathcal{L}^D_k(f) = \sum_{\omega, \alpha(k)} \chi_{\alpha(k)}(\omega) \left( S_{\alpha(k)}^\dagger(\omega) f S_{\alpha(k)}(\omega) - \frac{1}{2} (S_{\alpha(k)}^\dagger(\omega) S_{\alpha(k)}(\omega), f) \right), \]

where \( \omega \) are the so-called Bohr frequencies, and \( \chi_{\alpha(k)}(\omega) \) are the Fourier coefficients of the two point correlation functions of the environment. The operators \( S_{\alpha(k)}(\omega) \) are the Fourier coefficients of the system couplings \( S_{\alpha(k)} \):

\[ e^{-itH} S_{\alpha(k)} e^{itH} = \sum_\omega e^{it\omega} S_{\alpha(k)}(\omega) \]

(31)

The \( S_{\alpha(k)}(\omega) \) operators can be understood as mapping eigenvectors of \( H_{\Lambda} \) with energy \( \omega \) to eigenvectors of \( H_{\Lambda} \) with energy \( E + \omega \), and hence act in the Liouvillian picture as quantum jumps which transfer energy \( \omega \) from the system to the bath and back. Reversibility of the map can be interpreted as the fact that the jumps to and from the system at a given energy are equally likely. The following useful relations hold for any \( k \in \Lambda, \alpha(k) \) and \( \omega \). Let \( \rho \) be the Gibbs state of \( H_{\Lambda} \), then for any \( s \in [0, 1] \),

\[ \chi_{\alpha(k)}(-\omega) = e^{-\beta \omega} \chi_{\alpha(k)}(\omega), \]  

(32)

\[ \rho^s S_k(\omega) = e^{s \beta \omega} S_{\alpha(k)}(\omega) \rho^s. \]  

(33)

We can naturally restrict the action of the Davies generator to the neighborhood of a subset of the lattice in the natural way: let \( A \subset \Lambda \)

\[ \mathcal{L}^D_A(f) = i[H_{\Lambda}, f] + \sum_{k \in A} \mathcal{L}^D_k \]

We collect the properties of the Davies generators in the following lemma.

**Lemma 9 (properties of Davies generators)** Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \). Let \( \Phi_{\Lambda} : \Lambda \mapsto A_\Lambda \) be an \( \tau \)-local bounded and commuting potential, then the Davies generators \( \mathcal{L}^D \), as defined in Eqs. (30-33) satisfy the following properties:

1. For any \( A \subset \Lambda \), \( \mathcal{L}^D_A \) is the generator of a completely positive unital semigroup \( e^{t \mathcal{L}^D_A} \).
2. \( \mathcal{L}^D \) is \( r \)-local, meaning that each individual Lindblad term \( \mathcal{L}^D_k = \sum_{\omega, \alpha(k)}^g \mathcal{L}^D_{\omega, \alpha(k)} \) acts non-trivially only on a constant neighborhood of size \( r \) around \( k \in \Lambda \).

3. \( \mathcal{L}^D \) is locally primitive and locally reversible with respect to the global Gibbs state \( \rho \).

4. \( \mathcal{L}^D \) is frustration free.

**Proof:**
1. Complete positivity and unitality of \( e^{it\mathcal{L}_A^D} \) follow directly from the fact that Eqn. (30) is in Lindblad form, and by plugging in \( f \propto 1 \).

2. Locality of the Liouvillian follows from locality of the Lindblad operators \( S_{\alpha(k)}(\omega) \). Given that for all \( i, j \in \Lambda \), \( [\Phi(i), \Phi(j)] = 0 \), we get for any \( \alpha(k) \),

\[
e^{-itH} S_{\alpha(k)} e^{itH} = e^{-it \sum_{j \neq d(i,k)} \Phi(j)} S_{\alpha(k)} e^{it \sum_{j \neq d(i',k)} \Phi(j')},
\]

which is manifestly local.

3. Local primitivity was shown to hold if for each \( k \in \Lambda \), \( \{ S_{\alpha(k)} \} \) generates the full matrix algebra of site \( k \). Local reversibility follows directly by exploiting the relations in Eqns. (32) and (33) to show that \( \langle f, \mathcal{L}_A^D(g) \rangle_\rho = \langle \mathcal{L}_A^D(f), g \rangle_\rho \), for any \( A \subset \Lambda \).

4. Frustration freedom of the Davies generators is also implied by the local reversibility condition. Indeed, let \( A \subset \Lambda \), then by local reversibility, for every \( f, g \in A_\Lambda \), we get

\[
\langle f, \mathcal{L}_A^D(g) \rangle_\rho = \langle \mathcal{L}_A^D(f), g \rangle_\rho
\]

In particular, frustration freedom can be made explicit by choosing \( f \propto 1 \) then

\[
\text{tr} \left[ \rho \mathcal{L}_A^D(g) \right] = \langle 1, \mathcal{L}_A^D(g) \rangle_\rho = \langle \mathcal{L}_A^D(1), g \rangle_\rho = 0,
\]

by local primitivity. This implies that if the Liouvillian \( \mathcal{L}_A^D \) satisfies detailed balance with respect to the state \( \rho \), then \( \rho \) is a stationary state of \( \mathcal{L}_A^D \).

The Davies generators are often considered a good model for exploring thermalization in quantum systems. In particular, it is the standard approach for considering environment couplings in a variety of physical scenarios (e.g. for atomic or optical systems in the quantum regime).

### B. Heat-Bath generators

We now consider a second class of Gibbs samplers which is less physically motivated, but perhaps better suited for simulations on a quantum computer because of its simple structure. Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \). Let \( \Phi_\Lambda : \Lambda \rightarrow A_\Lambda \) be a local bounded and commuting potential, and let \( \rho \in S^+_\Lambda \) be the associated Gibbs state. For some \( A \subset \Lambda \), let \( \mathbb{E}_A^\rho \) be the minimum expectation value of \( \rho \) on \( A \), then we define the **Heat bath** Liouvillian as

\[
\mathcal{L}_A^H(f) = \sum_{k \in A} (\mathbb{E}_k^\rho(f) - f).
\]

Note that the conditional expectations are taken over single sites. Given that for any set of completely positive maps \( \{ T_j \} \), \( \sum_j (T_j - id) \) is a legitimate Liouvillian, we could have defined the heat bath Liouvillian with respect to essentially any set of conditional expectations. However, this choice is easier to work with, and closely mirrors the locality properties of the Davies generators. As will be clear in Sec. VII, the Heat-Bath generators are easier to work with in some settings than the Davies generators.

The theorem below collects the relevant properties of the Heat-Bath generator:
**Theorem 10 (properties Heat-Bath generator)** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_A : \Lambda \mapsto A$ be an $r$-local bounded and commuting potential, then the heat-bath generator $\mathcal{L}_H$ satisfies the following properties:

1. For any $A \subset \Lambda$, $\mathcal{L}_A^H$ is the generator of a completely positive unital semigroup $e^{t\mathcal{L}_A^H}$.

2. $\mathcal{L}_A^H$ is $r$-local, meaning that each individual Lindblad term $\mathcal{L}_k^A$ acts non-trivially only on a constant neighborhood of sites around $k \in \Lambda$.

3. $\mathcal{L}_A^H$ is locally primitive and locally reversible with respect to the global Gibbs state $\rho$.

4. $\mathcal{L}_A^H$ is frustration free.

**Proof:**

1. Complete positivity and unitality of $e^{t\mathcal{L}_A^H}$ follows directly from Eqn. (34) and the fact that the minimal conditional expectation $E^\rho$ is completely positive and unital.

2. Locality can be seen by direct evaluation of one term in the generator. Let $k \in \Lambda$, and consider

$$E_k^\rho(f) = \text{tr}_k[\eta_k^\rho f \eta_k^\rho]$$

where $\eta_k^\rho = (\text{tr}_k[e^{-\beta H_A}])^{-1/2}e^{-\beta H_A}/2$. But given that $\Phi_A$ is a commuting potential, we have

$$(\text{tr}_k[e^{-\beta H_A}])^{-1/2} = \left(e^{-\beta H_{(k)})}}{2} \text{tr}_k[e^{-\beta H_k}] e^{-\beta H_{(k)}}{2}\right)^{-1/2}$$

$$= e^{\beta H_{(k)}{4}}(\text{tr}_k[e^{-\beta H_k}])^{-1/2} e^{\beta H_{(k)}{4}}$$

$$= (\text{tr}_k[e^{-\beta H_k}])^{-1/2} e^{\beta H_{(k)}{4}}.$$  

where $H_{(k)}(\alpha) = \sum_{j \neq k} \Phi(j)$.

Here we have used that if two invertible hermitian operators $A, B$ commute $[A, B] = 0$, then also $[A^\alpha, B^\beta] = 0$ for any scalars $\alpha, \beta \in \mathbb{R}$, since two commuting hermitian operators share the same orthonormal basis.

Then, by the same arguments $e^{-\beta H_A}/2 = e^{-\beta H_{(k)}{2}} e^{-\beta H_k}/2$. Thus we get

$$\eta_k^\rho = (\text{tr}_k[e^{-\beta H_k}])^{-1/2} e^{-\beta H_k}/2$$

Hence, if $\Phi_A$ is a commuting $r$-local potential, then $\mathcal{L}_k^H$ is at most $r$-local, for any $k \in \Lambda$.

3. Local reversibility follows directly from reversibility of $E_k^\rho$ for any $A \subset \Lambda$. Local primitivity follows from Lemma [11]. In order to prove that $\mathcal{L}_A^H$ is locally primitive, we need to show that for any $A \subset \Lambda$, $\mathcal{L}_A^H(f) = 0$ implies that $f$ has support on $A^c$. We show this by contradiction. Let $A \subset \Lambda$, and suppose that there exists an $g \in A_A$ with non-trivial support on $A$ such that $\mathcal{L}_A^H(g) = 0$. Then it follows that $\langle g, \mathcal{L}_A^H(g) \rangle_\rho = 0$. From Lemma [11], there exists a $C_A > 0$ such that

$$\langle g, -\mathcal{L}_A^H(g) \rangle_\rho = \sum_{k \in A} \langle g, g - E_k^\rho(g) \rangle_\rho$$

$$\geq \frac{1}{C_A} \langle g, g - E_k^\rho(g) \rangle_\rho > 0,$$

since $E_k^\rho(f)$ has support on $A^c$ the support of $g$ has non-zero overlap with $A$. 
4. Frustration freedom follows in much the same way as for the Davies generators from the local reversibility of $L^H\Lambda$ which is inherited from the reversibility of the minimal conditional expectation $E^\rho$.

Remark: The Heat-Bath generators have been considered previously in the context of lattice spin system in a series of papers [31–33]. There the focus was on finding general local criteria for a quantum lattice system to be well defined in the thermodynamic limit. The results in Refs. [31–33] are hence similar in spirit to ours, but quite different in scope and in terms of the methods used. Hence, the two sets of results can be seen as being complementary.

We conclude this section by pointing out that the whole framework of Heat-Bath Liouvili-ans also works if we replace the Gibbs state by some other state of the lattice $\sigma$. It can be seen that Lemma 11 will still hold. However, it will typically be very difficult to obtain a bound on the locality of the individual terms.

**Lemma 11 (equivalence of blocks [32])** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_\Lambda : \Lambda \rightarrow A_\Lambda$ be an $r$-local bounded and commuting potential, and let $\rho$ be the Gibbs state of $H_\Lambda = \sum_k \Phi_\Lambda(k)$. Let $A \subset \Lambda$, then there exist constants $c_A, C_A < \infty$ such that for any $f \in A_\Lambda$,

$$c_A \sum_{k \in A} \langle f, f - E^\rho_k(f) \rangle_\rho \leq \langle f, f - E^\rho_A(f) \rangle_\rho \leq C_A \sum_{k \in A} \langle f, f - E^\rho_k(f) \rangle_\rho. \quad (35)$$

**IV. DECAY OF CORRELATIONS**

There are a number of ways of defining the correlations between observables in a quantum system. We will be interested in describing the situation when the correlations between local observables decay rapidly (exponentially) with the distance separating their supports. This behavior typically characterizes non-critical phases of many-body systems.

Let $A \subset \Lambda$, and let $E$ be a conditional expectation of $\rho \in S^+_\Lambda$. Then we define the conditional covariance with respect to $E$ on $A$ as

$$\text{Cov}_A(f, g) = \left| \langle f - E_A(f), g - E_A(g) \rangle_\rho \right|, \quad (36)$$

for any $f, g \in A_\Lambda$ and similarly, the conditional variance is given by $\text{Var}_A(f) = \text{Cov}_A(f, f)$. We note that the conditional covariance with respect to the full lattice $\Lambda$ reduces to the usual covariance.

**Definition 12 (weak clustering)** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$, and let $\rho \in S^+_\Lambda$. We say that $\rho$ satisfies weak clustering if there exist constants $c, \xi > 0$, such that for any observables $f, g \in A_\Lambda$,

$$\text{Cov}(f, g) \leq c \|f\|_2 \|g\|_2 e^{-d(\Sigma_f, \Sigma_g)/\xi}, \quad (37)$$

where $\Sigma_f$ ($\Sigma_g$) is the support of observable $f$ ($g$).

**Definition 13 (strong clustering)** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$, and let $\rho \in S^+_\Lambda$. Let $E$ be a conditional expectation of $\rho$, then we say that $\rho$ satisfies strong clustering with respect to $E$ if for any $A, B \subset \Lambda$ with $A \cap B \neq \emptyset$, there exist constants $c, \xi > 0$ such that for any $f \in A_\Lambda$,

$$\text{Cov}_{A \cup B}(E_A(f), E_B(f)) \leq c \|f\|_2^2 e^{-d(A, B)/\xi}, \quad (38)$$

where $\bar{A} \equiv (A \cup B) / A$ and $\bar{B} \equiv (A \cup B) / B$ (see Fig. 1a).
FIG. 1: a) A subset $A \cup B \subset \Lambda$ of the full system, where $A \cap B \neq \emptyset$. The dotted lines around $A \cup B$ represent the boundary, which includes all terms of the Hamiltonian that overlap with $A \cup B$. The relevant distance is the width of the region $A \cup B$. b) $B \subset A \subset \Lambda$, where the relevant distance is between the boundary of $B$ and the boundary of $A$.

It turns out that for the two conditional expectations considered in this paper it suffices to consider strong clustering for observables $f$ that act only on $A \cup B$ plus its boundary:

**Proposition 14** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_{\Lambda} : \Lambda \to A_{\Lambda}$ be an $r$-local bounded and commuting potential, and let $\rho$ be the Gibbs state of $H_\Lambda = \sum_{k} \Phi_{\Lambda}(k)$. Let $L_{\Lambda}^{\rho}$ be the Davies generator, and $\mathbb{E}_\rho$ its associated conditional expectation, and let $\mathbb{E}^{\rho}$ be the minimal expectation of $\rho$. For both $\mathbb{E} = \mathbb{E}_\rho$ and $\mathbb{E} = \mathbb{E}^{\rho}$, it follows that

$$
\sup_{f \in A_{\Lambda}} \text{Cov}_{A \cup B}(\mathbb{E}_A(f), \mathbb{E}_B(f)) \frac{||f||_2^{2,\rho}}{||f||_2^{2,\rho}} = \sup_{f \in A_{(A \cup B)_{\beta}}} \text{Cov}_{A \cup B}(\mathbb{E}_A(f), \mathbb{E}_B(f)) \frac{||f||_2^{2,\rho}}{||f||_2^{2,\rho}}.
$$

**(39)**

**Proof:** We consider the expression on the left hand side of Eqn. (39), and note that

$$
\sup_{f \in A_{\Lambda}} \text{Cov}_{A \cup B}(\mathbb{E}_A(f), \mathbb{E}_B(f)) \frac{||f||_2^{2,\rho}}{||f||_2^{2,\rho}} = \sup_{f \in A_{\Lambda}} \frac{\langle f, \mathbb{E}_A \mathbb{E}_B - \mathbb{E}_{A \cup B}(f) \rangle_\rho}{\langle f, f \rangle_\rho} = \sup_{g \in A} \frac{\text{tr}[g \mathcal{W}_{A \cup B}(g)]}{\text{tr}[g^2]}.
$$

**(40)**

where we made the replacement $g \equiv \Gamma_\rho^{1/2}(f)$, and $\Gamma_\rho(f) := \rho^{1/2} f \rho^{1/2}$. We defined the operators $\mathcal{W}_{A \cup B}(f) := \mathbb{E}_A \mathbb{E}_B(f) - \mathbb{E}_{A \cup B}$, and $\mathcal{W}_{A \cup B} = \Gamma_\rho^{1/2} \mathcal{W}_{A \cup B} \Gamma_\rho^{-1/2}$. We note that $\mathcal{W}_{A \cup B}$ is a hermitian operator, so Eqn. (40) is simply an eigenvalue equation.

Now, we will show that $\mathcal{W}_{A \cup B}$ acts non-trivially only on $(A \cup B)_{\beta}$. We write the subscript of $\rho$ explicitly so as to avoid confusion. Indeed, for any $g \in A_{\Lambda}$,

$$
\mathcal{W}_{A \cup B}(g) = \Gamma_\rho^{1/2} \circ (\mathbb{E}_A \mathbb{E}_B - \mathbb{E}_{A \cup B}) \circ \Gamma_\rho^{-1/2}(g)
$$

**(41)**

$$
= \Gamma_\rho^{1/2} \circ \rho_{(A \cup B)_{\beta}} \circ (\mathbb{E}_A \mathbb{E}_B - \mathbb{E}_{A \cup B}) \circ \Gamma_\rho^{-1/2} \rho_{(A \cup B)_{\beta}}(g)
$$

**(42)**

$$
= \Gamma_{\rho_{(A \cup B)_{\beta}}} \circ (\mathbb{E}_A \mathbb{E}_B - \mathbb{E}_{A \cup B}) \circ \Gamma_{\rho_{(A \cup B)_{\beta}}}^{-1/2}(g)
$$

**(43)**

$$
= \Phi_{(A \cup B)_{\beta}} \otimes \text{id}_{(A \cup B)_{\beta}}(g)
$$

**(44)**

for some hermitian operator $\Phi_{(A \cup B)_{\beta}}$ acting only on $(A \cup B)_{\beta}$. But it is well known that the supremum in the variational characterization of the spectral radius is obtained by 'vectors'
which have the same support as the operator. Thus, we recover the right and side of Eqn. (39).

Remarks:

i The fact that the observables can be restricted to having support only on the boundary of the region $A \cup B$ is reminiscent of the definition of clustering for classical spin systems, in which one is allowed to fix the boundary around a given region. Strong clustering goes in the direction of the Dobrushin-Schlossman uniqueness conditions [55]. However, since DLR theory of boundary conditions does not hold quantum mechanically [48], the stronger form of mixing in Eqn. (38) can not be expressed as local conditions which depend on individual boundary terms. This is because of the possibility of entangled boundary conditions, which have, as far as the authors know, not been studied much so far in the context of Gibbs states.

ii Whenever $A \cup B = \Lambda$ then weak clustering also implies strong clustering since for any conditional expectation $E$, $E_A(f)$ has support on the complement of $A$, so that for any $h \in A\Lambda$,

$$\text{Cov}_\Lambda(E_A(h), E_B(h)) \leq c |E_A(h)||E_B(h)||_2,\rho e^{-d(\overline{A}, \overline{B})}/\xi \quad (45)$$

$$\leq c |h||_2,\rho e^{-d(\overline{A}, \overline{B})}/\xi \quad (46)$$

We have defined exponential clustering in Eqns. (37) and (38) as being exponential decay of the covariance (with distance) weighted by the $L_2$ norms of the observables, instead of the operator norms, which is more common in the field. Given the ordering of $L_p$ norms (Lemma 3), the operator norm always dominates the $L_2$ norm, so that our definitions of exponential clustering are strictly stronger than the usual ones. Although classically, the $L_2$ and the $L_\infty$ clustering are equivalent [43], we do not know if that is the case for quantum systems, even for commuting Hamiltonians.

Note that the inner product used here is also unconventional. In particular, when $\rho$ is a pure state, then $\text{Cov}_\rho(f, g)$ does not reduce to the usual pure-state covariance

$$|\langle \psi | f g | \psi \rangle - \langle \psi | f | \psi \rangle \langle \psi | g | \psi \rangle|, \quad (47)$$

However, it is easy to define a modified (non-symmetric) covariance

$$\text{Cov}^{(0)}_{\Lambda}(f, g) = |\text{tr}[\rho f^\dagger g] - \text{tr}[\rho f^\dagger] \text{tr}[\rho g]|, \quad (48)$$

that reduces to Eq. (47) when $\rho$ is pure. In general, $\text{Cov}^{(0)}_{\rho}$ and $\text{Cov}_{\rho}$ are not equivalent. However, in the special case when $\rho$ is the Gibbs state of a commuting Hamiltonian, weak clustering in $\text{Cov}^{(0)}_{\rho}$ implies weak clustering in $\text{Cov}_{\rho}$. Indeed,

**Proposition 15** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_{\Lambda}: \Lambda \mapsto A_\Lambda$ be an $r$-local bounded and commuting potential, and let $\rho$ be the Gibbs state of $H_\Lambda = \sum_k \Phi_{\Lambda}(k)$. Then the following are equivalent:

- **There exist constants** $c_0, \xi_0 > 0$ **such that**

$$\text{Cov}^{(0)}_{\rho}(f, g) \leq c_0 ||f||_{2(r), \rho},||g||_{2(r), \rho} e^{-d(\Sigma_f, \Sigma_g)/\xi_0}, \quad (49)$$

- **There exist constants** $c, \xi > 0$ **such that**

$$\text{Cov}_{\rho}(f, g) \leq c ||f||_{2,\rho},||g||_{2,\rho} e^{-d(\Sigma_f, \Sigma_g)/\xi}, \quad (50)$$
where we have defined the modified $\mathbb{L}_2$ norm $||f||_{2,(0),\rho}^2 = \text{tr} \left[ \rho f^\dagger f \right]$.

**Proof:** Given an operator $f \in A_\Lambda$ with support on $A \subset \Lambda$, $\rho^s f \rho^{-s}$ has support on $A_\beta$, for any $s \in [0,1]$, since $\rho$ is the Gibbs state of a commuting Hamiltonian. Now define $\tilde{f} \equiv \rho^{-1/4} f \rho^{1/4}$ and $\tilde{g} \equiv \rho^{1/4} g \rho^{-1/4}$. Then we get

$$\text{Cov}_\rho(f, g) = \left|\text{tr} \left[ \rho^{1/2} f^\dagger \rho^{1/2} g \right] - \text{tr} \left[ \rho f^\dagger \right] \text{tr} \left[ \rho g \right] \right|$$

$$= \left|\text{tr} \left[ \rho f^\dagger \tilde{g} \right] - \text{tr} \left[ \rho \tilde{f}^\dagger \right] \text{tr} \left[ \rho \tilde{g} \right] \right|$$

$$= \text{Cov}_\rho^{(0)}(\tilde{f}, \tilde{g})$$

where we have used that

$$||\tilde{g}||^2_{2,(0),\rho} = \text{tr} \left[ \rho \tilde{g} \right]$$

$$= \text{tr} \left[ \rho^{1/2} \tilde{g} \rho \rho^{1/2} \tilde{g} \right] = ||g||^2_{2,\rho}$$

The other direction is identical, except that one has to define $\tilde{f} \equiv \rho^{1/4} f \rho^{-1/4}$ and $\tilde{g} \equiv \rho^{1/4} g \rho^{-1/4}$.

**Remark:**

- A generalized covariance which interpolates between $\text{Cov}_\rho^{(0)}$ and $\text{Cov}_\rho$ was introduced in Ref. [52], and was shown to be a necessary ingredient in a proof of the existence of a universal critical temperature above which correlations in the Gibbs state are exponentially clustering. This led to a stability theorem for locally perturbed Gibbs states at high temperatures.

### A. Local indistinguishability

One of the main contributions in this work is to introduce extended notions of clustering to characterize phases where correlations decay rapidly in a very strong sense. In this section, we consider how weak clustering relates to another important measure of correlation:

**Definition 16 (Local indistinguishability)** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_\Lambda : \Lambda \to A_\Lambda$ be an $r$-local bounded and commuting potential, and let $\rho$ be the Gibbs state of $H_\Lambda = \sum_k \Phi_\Lambda(k)$. Let $B \subset A \subset \Lambda$ (see Fig. [7]), and let $E_B$ be a conditional expectation of $\rho$. Let $\rho^A$ and $\sigma^A$ be two local Gibbs states of $A$ in the sense that they are both fixed points of the map $E_B^*$. Then, we say that the Gibbs state is **locally indistinguishable** if there exist constants $c, \xi > 0$ such that

$$||\langle \rho^A \rangle_B - \langle \sigma^A \rangle_B||_1 \leq c e^{-d(B,A)/\xi}$$

A condition similar to local indistinguishability was previously considered in Ref. [17], and called *Local Topological Quantum Order (LTQO)* because of an analogous condition for ground states of topologically ordered Hamiltonians (see also Refs. [50, 51] for a closed system analogue). However, in the Gibbs sampler setting, as far as we know this condition does not appear to be connected with topological order, which is why we give it a different name here (see however the discussion in the outlook).

We now show that a strengthening of weak clustering (changing the bound from 2-norm to the product of 1- and infinity-norms) is equivalent to local indistinguishability:
Theorem 17 Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \). Let \( \Phi_\Lambda : \Lambda \mapsto A_\Lambda \) be an \( r \)-local bounded and commuting potential, and let \( \rho \) be the Gibbs state of \( H_\Lambda = \sum_\Lambda \Phi_\Lambda(k) \), and let \( \mathbb{E} \) be a conditional expectation of \( \rho \). Then the following are equivalent:

- There exist constants \( c_0, \xi_0 > 0 \) such that for any \( f, g \in A_\Lambda \),
  \[
  \operatorname{Cov}_\rho(f, g) \leq c \|f\|_{1, \rho} \|g\|_{\infty} e^{-\ell(\Sigma_f, \Sigma_g)/\xi},
  \]  
  (59)

- \( \rho \) satisfies local indistinguishability.

PROOF: We first show that Eqn. (59) implies local indistinguishability.

Let \( \psi, \phi \in \mathcal{S}_\Lambda \) such that \( \rho^\Lambda = \mathbb{E}_\Lambda^\Lambda(\phi) \) and \( \sigma^\Lambda = \mathbb{E}_\Lambda^\Lambda(\phi) \). We now choose \( A, C \subset \Lambda \) such that \( B \cap C = \emptyset \) and \( A \cup C = \Lambda \), as illustrated in Fig. 1b), we get

\[
||(\rho^\Lambda)_B - (\sigma^\Lambda)_B||_1 = \sup_{g \in A_B, ||g|| = 1} |\text{tr} [\mathbb{E}_A(g) \phi - \psi]| \\
= \sup_{g \in A_B, ||g|| = 1} |\text{tr} [(\mathbb{E}_A(g) - \mathbb{E}_{A \cup C}(g)) \psi] - \text{tr} [(\mathbb{E}_A(g) - \mathbb{E}_{A \cup C}(g)) \phi]| \\
\leq 2 \sup_{g \in A_B, ||g|| = 1} |\text{tr} [\phi(\mathbb{E}_A(g) - \mathbb{E}_A(g))]| \\
\]

Now, defining \( y := \rho^{-1/2} \phi \rho^{-1/2} \), and noting that \( \mathbb{E}_C(g) = g \), we get

\[
||(\rho^\Lambda)_B - (\sigma^\Lambda)_B||_1 = 2 \sup_{g \in A_B, ||g|| = 1} |\text{tr} [\phi(\mathbb{E}_A\mathbb{E}_C(g) - \mathbb{E}_A(g))]| \\
= 2 \sup_{g \in A_B, ||g|| = 1} \langle y, (\mathbb{E}_A \mathbb{E}_C - \mathbb{E}_A)(g) \rangle_\rho \\
= 2 \sup_{g \in A_B, ||g|| = 1} \operatorname{Cov}_\Lambda (\mathbb{E}_A(y), \mathbb{E}_C(g)) \\
\leq 2 \sup_{g \in A_B, ||g|| = 1} ||\mathbb{E}_A(y)||_{1, \rho} ||\mathbb{E}_C(g)||_{\infty} e^{-\ell(B, A/\Lambda)/\xi} \\
= 2 \sup_{g \in A_B, ||g|| = 1} ||y||_{1, \rho} ||g||_{\infty} e^{-\ell(B, A/\Lambda)/\xi} \\
\leq 2 e^{-\ell(B, A/\Lambda)/\xi},
\]

where we have used that \( ||y||_{1, \rho} = \text{tr} [\phi] = 1 \), and that \( \mathbb{E}_C(g) = g \).

For the converse, let \( A, B \subset \Lambda \) such that \( A \cup B = \Lambda \) and let \( f, g \in A_\Lambda \), where \( f \) has support on \( B^c \) and \( g \) has support on \( A^c \). Now consider

\[
\operatorname{Cov}_\Lambda (\mathbb{E}_A(g), \mathbb{E}_B(f)) = \operatorname{Cov}_\Lambda (\mathbb{E}_A(g), \mathbb{E}_B(f)) \\
= \langle g, (\mathbb{E}_A \mathbb{E}_B - \mathbb{E}_A)(f) \rangle_\rho \\
\leq ||g||_{1, \rho} ||(\mathbb{E}_A \mathbb{E}_B - \mathbb{E}_A)(f)||_{\infty} \\
= ||g||_{1, \rho} \sup_{\varphi \in \mathcal{S}_\Lambda} |\text{tr} [\varphi(\mathbb{E}_A \mathbb{E}_B - \mathbb{E}_A)(f)]| \\
= ||g||_{1, \rho} \sup_{\varphi \in \mathcal{S}_\Lambda} |\text{tr} [\mathbb{E}_A^*(\mathbb{E}_A - \mathbb{E}_A^*)(\varphi) \mathbb{E}_B(f)]| \\
= ||g||_{1, \rho} ||\mathbb{E}_B(f)||_{\infty} \sup_{\varphi \in \mathcal{S}_\Lambda} |\text{tr} [\mathbb{E}_A^*(\mathbb{E}_A - \mathbb{E}_A^*)(\varphi)]||_1 \\
\leq c ||g||_{1, \rho} ||f||_{\infty} e^{-\ell(A, B)/\xi},
\]  
(60)
where we have used that $E_B$ is contractive in $\infty \to \infty$ norm.

Remarks:

i Eqn. (59) only differs from the definition of weak clustering in that the norms on the right hand side are different. However, it is exactly this difference that allows the connection with local indistinguishability. Combining Theorem 17 and the results of Ref. [17], we see that this form of clustering follows from having a system size-independent log-Sobolev constant, while weak clustering follows from the system having merely a constant spectral gap [19].

ii It is not known whether there is a relation between strong clustering and local indistinguishability. By analogy with classical results, one might expect that under certain conditions local indistinguishability implies strong clustering, but this is far from clear in our setting.

V. MAIN RESULTS

We are now in a position to prove the main results of the present work, namely the equivalence between strong clustering of the Gibbs state and the associated Gibbs samplers (Heat Bath or Davies) being gapped. It turns out that both directions of the proof require very different methods, hence for clarity of presentation we will separate them into two independent theorems.

To start with, we recall the definition of the conditional variance, as it will play an important role in the proof. Let $\Lambda$ be the full lattice, and let $A \subset \Lambda$, then the conditional variance of $\rho \in S_+^\Lambda$ with respect to the conditional expectation $E$ on subset $A$ is given for any $f \in A_\Lambda$ by

$$\text{Var}_A(f) = \langle f - E_A(f), f - E_A(f) \rangle_\rho = \|f - E_A(f)\|^2_{2,\rho},$$

and the conditional variance reduces to the regular variance on the full lattice: $\text{Var}_\Lambda(f) = \langle f, f \rangle_\rho - \text{tr} [\rho f]^2$.

We now give a proposition which relates the conditional variance of two subsets $A, B$ to the variance of their union $(A \cup B)$ when their overlap $A \cap B$ is non-zero (see Lemma 3.1 of [39] for a similar statement in the classical setting).

**Proposition 18** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\rho \in S_+^\Lambda$, and let $A, B \subset \Lambda$ be subsets of $\Lambda$ with non-zero overlap (i.e. $A \cap B \neq \emptyset$). Let $E$ be a conditional expectation of $\rho$, and suppose that there exists a positive constant $\epsilon > 0$ such that for any $f \in A_\Lambda$, we have

$$\text{Cov}_{A \cup B}(E_A(f), E_B(f)) \leq \epsilon \text{Var}_{A \cup B}(f),$$

then

$$\text{Var}_{A \cup B}(f) \leq (1 - 2\epsilon)^{-1}(\text{Var}_A(f) + \text{Var}_B(f)).$$

**Proof:** Consider the following identity:

$$0 \leq \|((id - E_{A \cup B}) \circ (id - E_A - E_B))(f)\|^2_{2,\rho}$$

$$= -\|((id - E_{A \cup B})(f))\|^2_{2,\rho} + \|((id - E_{A \cup B}) \circ (id - E_A)(f))\|^2_{2,\rho} + \|((id - E_{A \cup B}) \circ (id - E_B)(f))\|^2_{2,\rho} + 2 ((id - E_{A \cup B}) \circ E_A(f), (id - E_{A \cup B}) \circ E_B(f))_\rho$$

(64)
Then, recalling that $||(id - E_{A \cup B})||^2_{2,\rho} = Var_{A \cup B}(f)$, and noting that

$$||(id - E_{A \cup B}) \circ (id - E_A)(f)||^2_{2,\rho} \leq ||(f - E_A(f))||^2_{2,\rho} = Var_A(f)$$

and similarly for $E_B$, since $(id - E_{A \cup B})$ is a positive contractive map, we get from Eqn. (64) $Var_{A \cup B}(f) \leq Var_A(f) + Var_B(f) + 2Cov_{A \cup B}(E_A(f), E_B(f))$.

This leads to the desired inequality

$$Var_{A \cup B}(f) \leq (1 - 2\epsilon)^{-1}(Var_A(f) + Var_B(f)). \quad (66)$$

Note that for the proof of Proposition 18 we have only used very general properties of the conditional expectations; in particular we have not assumed that $\rho$ is a Gibbs state, or that $E$ has any local structure.

### A. Strong clustering implies gapped Gibbs sampler

We now prove the first main theorem of the paper, which states that if the Gibbs state $\rho$ of a local commuting Hamiltonian satisfies strong clustering with respect to any of the two conditional expectations defined in Sec II, then the associated Gibbs sampler (Heat Bath or Davies) has a spectral gap which is independent of the size of the lattice $|\Lambda|$

By construction, the pairs $(E^L, L^R_H)$ and $(E^\rho, L^R_A)$ have the same kernel and share the same essential properties, such as reversibility and locality. In the remainder of this section, we will explicitly consider the Davies generator and associated Liouvillian expectation $(L^D, E^L)$, but it should be clear that the proof carries through essentially unchanged for the pair $(L^H, E^\rho)$.

As the proof is based on an iterative construction comparing gaps of different sub lattices, it is necessary to define what we mean by the ‘gap of $L$’ restricted to region $A$:

**Definition 19** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_{\Lambda} : \Lambda \mapsto A_{\Lambda}$ be an $r$-local bounded and commuting potential, and let $\rho$ be the Gibbs state of $H_{\Lambda} = \sum_k \Phi_{\Lambda}(k)$. The conditional gap of $L^R_{\Lambda}$ with respect to $A \subset \Lambda$ is

$$\lambda_{\Lambda}(A) = \inf_{f \in A_{\Lambda}} \frac{\langle f, -L^R_{\Lambda}(f) \rangle_{\rho}}{Var_{\Lambda}(f)}. \quad (67)$$

Given that $E^L$ is a projector and $L^R_{\Lambda}$ is reversible, Eqn (67) is identical to the spectral definition of the gap; i.e. the smallest non-zero eigenvalue of $-L^R_{\Lambda}$. In the case of $L^H_{\Lambda}$, this will not be so, since $E^\rho$ is not a projective conditional expectation.

Note that the optimization is taken over operators with support on the full lattice $\Lambda$ with respect to the Gibbs state of the full lattice. The gap $\lambda_{\Lambda}(A)$ is non-zero for any subset $A \subset \Lambda$
because by assumption $\mathbb{E}_A$ and $L_A$ have the same kernel, and $(\mathbb{E}, L)$ are assumed to be locally primitive\(^2\).

The next lemma tells us that when the potential $\Phi_A$ is commuting, the conditional gap of $L$ with respect to $A$ in $\Lambda$ is the same as the conditional gap of $L$ with respect to $A$ in $A_\partial$.

**Lemma 20** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_A : \Lambda \rightarrow A_\Lambda$ be an $r$-local bounded and commuting potential. Let $A \subset \Lambda$, and let $\rho_A$ be the Gibbs state of $H_A = \sum_{k \in A} \Phi_A(k)$ on $S_{A_\partial}$ (see Eqn. (28)) and $\rho$ the Gibbs state of $H_\Lambda = \sum_{k \in \Lambda} \Phi_\Lambda(k)$ on $S_\Lambda$. We get that

$$
\lambda_\Lambda(A) = \lambda_{A_\partial}(A) := \inf_{f \in A_{A_\partial}} \frac{\langle f, -L_A^D(f) \rangle_{\rho_A}}{||f - \mathbb{E}_A(f)||^2_{\rho_A}}
$$

The proof is provided in the appendix, as it very closely resembles that of Prop. 14.

**Remark:**
- Lemma 20 tells us that even though the gap of $L_A$ is defined with respect to the full space $A_\Lambda$, the variational optimization reaches its maximum for observables that have support on $A_\partial$. This quasi independence on the complement of $A$ is very important in the proof of the main theorem. It should be understood as a partial quantum extension of the Dobrushin-Schlossmann conditions.

We are now in a position to state our main theorem:

**Theorem 21** Let $\Lambda$ be a finite lattice in $\mathbb{Z}^d$. Let $\Phi_A : \Lambda \rightarrow A_\Lambda$ be an $r$-local bounded and commuting potential, and let $\rho$ be the Gibbs state of $H_\Lambda = \sum_{k \in \Lambda} \Phi_\Lambda(k)$. If $\rho$ satisfies strong clustering with respect to $E_\Lambda$, then $L_D^\Lambda$ has a spectral gap which is independent of $|\Lambda|$.

The proof strategy follows closely Ref. [39], and consists in showing that for any sufficiently large subset $C \subset \Lambda$, we can choose $A, B \subset \Lambda$ such that $A \cup B = C$ and $A \cap B \neq \emptyset$ and such that the conditional gap with respect to $A$ or $B$ is approximately the same as the conditional gap with respect to $C$. Choosing $A$ and $B$ to be roughly half the size of $C$, this shows that doubling the lattice size essentially leaves the conditional gap unchanged. By applying this procedure iteratively, we can show that the gap of the full system is lower bounded by the gap of a constant size subset. Finally, invoking Lemma 20 we get that the conditional gap of a constant size subset of the lattice is lower bounded by a constant. Hence the gap of $L$ on the full system cannot depend on the system size.

**Proof:** Assume that $\Lambda \subset \mathbb{Z}^d$ is a sufficiently large rectangle.

Let $A, B \subset \Lambda$ such that $A \cup B \neq \emptyset$ and $A \cap B$ forms a rectangle of minimal side length $L$, while the overlap $A \cap B$ has minimum side length larger or equal to $\sqrt{L}$; as in Fig. 1. Recall that the strong clustering assumption implies that there exist constants $c, \xi > 0$ such that for any $f \in A_\Lambda$,

$$
\text{Cov}_{A \cup B}(\mathbb{E}_A(f), \mathbb{E}_B(f)) \leq c \text{Var}_{A \cup B}(f) e^{-\sqrt{\xi/L}}
$$

\(^2\) Recall that the entire framework only makes sense for primitive semigroups, since $L_p$ spaces rely on a full rank reference state.
Then, from Proposition (18) and the definition of the conditional gap (Eqn. (67)) we get that for any $f \in A\Lambda$:

$$\text{Var}_{A \cup B}(f) \leq (1 - 2ce^{-\sqrt{T}/\xi})^{-1}(\text{Var}_A(f) + \text{Var}_B(f))$$

$$\leq (1 - 2ce^{-\sqrt{T}/\xi})^{-1}\left(\frac{\langle f, -\mathcal{L}_A(f) \rangle_\rho}{\lambda_A(A)} + \frac{\langle f, -\mathcal{L}_B(f) \rangle_\rho}{\lambda_A(B)}\right)$$

$$\leq (1 - 2ce^{-\sqrt{T}/\xi})^{-1}\frac{1}{\lambda_A(A \land B)}\left(\langle f, -\mathcal{L}_{A \land B}(f) \rangle_\rho + \langle f, -\mathcal{L}_{A \land B}(f) \rangle_\rho\right),$$

where we have written $\lambda_A(A \land B) := \min\{\lambda_A(A), \lambda_A(B)\}$.

At this point we might be tempted to upper bound $\langle f, -\mathcal{L}_{A \land B}(f) \rangle_\rho$ by $\langle f, -\mathcal{L}_{A \cup B}(f) \rangle_\rho$. But this would provide us with a bound where the gap roughly halves in magnitude when we double the size of the system, which would lead upon iterations to a global gap decreasing exponentially with the system size. However, we can use an averaging trick originally developed for a similar purpose in classical lattice spin systems [39,42,43].

Given a rectangular subset of the lattice $C \subset \Lambda$, suppose there exists a sequence of subsets $\{A_i, B_i\}_{i=1}^s$, where $s := \lceil \sqrt{T}/3 \rceil$, with the property that for every $i = 1, ..., s$

- $A_i \cup B_i = C$,
- $A_i \cap B_i$ has minimum side length lower bounded by $\sqrt{T}$,
- $A_i \cap B_i \cap A_j \cap B_j = 0$ for all $i \neq j$

Then, by noting that

$$\sum_{i=1}^s \langle f, -\mathcal{L}_{A_i \cap B_i}(f) \rangle_\rho \leq \langle f, -\mathcal{L}_C \rangle_\rho,$$
we get
\[ \text{Var}_C(f) \leq (1 - 2ce^{\sqrt{T}/\xi})^{-1} \max_{i=1,...,s} \{ \left( \frac{1}{\lambda\Lambda(A_i)} + \frac{1}{\lambda\Lambda(B_i)} \right) \} \]
\[ \left( \langle f, -\mathcal{L}_C \rangle_{\rho} + \frac{1}{s} \sum_{i=1}^{s} \langle f, -\mathcal{L}_{A_i \cap B_i} \rangle_{\rho} \right) \]
\[ \leq (1 - 2ce^{\sqrt{T}/\xi})^{-1} \max_{i=1,...,s} \{ \left( \frac{1}{\lambda\Lambda(A_i)} + \frac{1}{\lambda\Lambda(B_i)} \right) \} \langle f, -\mathcal{L}_C \rangle_{\rho} \]
(72)

It is not difficult to see now that, for given \( c, \xi \), there exists an \( L_0 \) such that for all \( L \geq L_0 \),
\[ (1 - 2ce^{\sqrt{T}/\xi})^{-1} \left( 1 + \frac{1}{[L^{1/3}]} \right) \leq \left( 1 + \frac{2}{[L^{1/3}]} \right), \]
which leads to
\[ \lambda\Lambda(C) \geq \left( 1 + \frac{2}{[L^{1/3}]} \right)^{-1} \min_{i=1,...,s} \{ \lambda\Lambda(A_i), \lambda\Lambda(B_i) \} \]
(74)

Clearly, the specific value of \( L_0 \) depends on the constants \( c, \xi \) an not on the system size \(|\Lambda|\).

In order to complete the proof, we construct a decomposition of the full lattice \( \Lambda \) into sequential subsets in such a way that we can use the bound in Eqn. (74) iteratively, and obtain a (global) lower bound on \( \lambda\Lambda(\Lambda) \). The construction has been taken from Ref. [39] which in turn was initiated in the work of Martinelli [43]. Let \( l_k := (3/2)^k/d \), and let \( \mathcal{R}_k^d \) be the set of all rectangles in \( \mathcal{R}^d \) which, modulo translations and permutations of the coordinates, are contained in
\[ [0, l_{k+1}] \times [0, l_{k+2}] \times \ldots \times [0, l_{k+d}] \]
(75)

Assume that \( \Lambda \equiv \mathcal{R}_k^{d_{\text{max}}} \). We will later show that for any size \( \Lambda \), the gap is always lower bounded. Note that we never explicitly compare the gaps of two systems \( \Lambda_1 \) and \( \Lambda_2 \) where each has specified boundary conditions, rather some boundary conditions are fixed for \( \Lambda \equiv \mathcal{R}_k^{d_{\text{max}}} \) and left untouched thereafter. We will also define the minimum gap restricted to rectangles in \( \mathcal{R}_k^d \) as \( g_k := \inf_{V \in \mathcal{R}_k^d} \lambda\Lambda(\Lambda) \). The idea behind this construction is that each rectangle in \( \mathcal{R}_k^d/\mathcal{R}_{k-1}^d \) can be obtained as a slightly overlapping union of two rectangles in \( \mathcal{R}_{k-1}^d \). By Lemma [22] we can then get the following iterative bound:

For all \( C_k \subset \mathcal{R}_{k+1}^d/\mathcal{R}_k^d \) and \( C_{k-1} \subset \mathcal{R}_k^d/\mathcal{R}_{k-1}^d \),
\[ \lambda\Lambda(C_k) \geq (1 + 2\left( [L^{1/3}] \right))^{-1} \lambda\Lambda(C_{k-1}) \]
(76)

In particular the minimum \( L_0 \) can be associated with a minimum integer \( k_0 \), such that taking \( \Lambda \) to be the thermodynamic lattice, and taking \( k = k_0 + 1, \ldots, \infty \), we get
\[ \lim_{\Lambda \to \mathbb{Z}^d} \lambda\Lambda(\Lambda) \geq \left( \prod_{k=k_0+1}^{\infty} (1 + 2(3/2)^{-k/(3d)}) \right)^{-1} \inf_{C_{k_0} \subset \mathcal{R}_k^d} \lambda\Lambda(C_{k_0}) \]
\[ \geq \exp[-2(1 - (2/3)^{1/(3d)})] \inf_{C_{k_0} \subset \mathcal{R}_k^d} \lambda\Lambda(C_{k_0}) \]
(78)
Finally, since $k_0$ is determined only by $L_0$, it is independent of $|\Lambda|$. Hence, we get from Lemma 20 that $\lambda_\Lambda(C_{k_0})$ can be lower bounded by a constant independent of $|\Lambda|$ for any $C_{k_0} \subset \mathcal{R}_{k_0}$.

**Remarks:**

i In order for the iterative procedure to work, we don’t need strong clustering to be exponential. Any polynomial decay with sufficiently high degree (strictly larger than $1/2$) will in fact do the job. Furthermore, Theorem 21 actually shows that for every rectangle $A \subset \Lambda, \mathcal{L}_A$ is gapped.

ii If one were able to extend Prop. 18 to show that

$$\text{Ent}_{A \cup B}(f) \leq (1 - 2\epsilon)^{-1}(\text{Ent}_A(f) + \text{Ent}_B(f)),$$

where the $\text{Ent}(f)$ was defined in Ref. [20], and $\text{Ent}_A(f)$ is some appropriately chosen conditional entropy, then the same proof strategy would work to show that the Gibbs samplers satisfy a Log-Sobolev inequality. This, in turn would have strong implications for the mixing of the semigroup, and would for instance show that strong clustering implies local indistinguishability, among other things.

iii The proof of Theorem 21 requires strong clustering in order to invoke Proposition 18. If we were able to find a systematic procedure to associate boundary conditions $\zeta$ to the Hamiltonian $H_A$ for any $A \subset \Lambda$, and prove that $\lambda_\Lambda(A) \geq \lambda_\zeta(A, \partial A)$, then we would be able to prove Theorem 21 with respect to weak clustering rather than strong clustering. Indeed, whenever $A \cup B = \Lambda$ then Proposition 18 holds under weak clustering. However, there is evidence to believe that Theorem 21 does not hold in general dimensions under the weak clustering assumption. Consider the 4D toric code, whose Davies generators are known not to be gapped (in fact the gap decreases exponentially). We know that the ground state of the 4D toric code satisfies a strong form of LTQO, where the suppression at large distances is not exponentially suppressed, but exactly zero. It then is plausible (although not proven) that the 4D toric Hamiltonian Gibbs state at non-zero, but small, temperatures satisfies local indistinguishability, as in Definition 16 but with the infinity norm replaced by the $L_2$ norm. If true, this in turn would imply weak clustering at low enough temperature. Hence the equivalence between weak and strong clustering for the 4D toric Hamiltonian would lead to a contradiction in general. See the outlook for a further discussion.

iv It is worth noting that in a sequence of important papers, Majewski and Zegarlinski have considered a similar approach to generalizing Glauber dynamics analysis to the quantum setting [31–33]. There, they introduce the equivalent of our Heat Bath sampler, and show that the dynamics are well defined in the thermodynamic limit. Furthermore, they show that under some strong local ergodicity conditions reminiscent of a certain form of the Dobrushin-Schlossmann complete analytic conditions, the dynamics are rapidly mixing, and in particular are gapped. Their conditions allow to show both our strong clustering, and local indistinguishability, and we hence expect them to be overly stringent for our main theorem. Those conditions, in particular, also lead to a proof that the Heat Bath sampler is always gapped at high enough temperatures.

**Lemma 22 ([39])** For all $C \subset \mathcal{R}_d^d/\mathcal{R}_{d-1}^d$ there exists a finite sequence $\{A_i, B_i\}_{i=1}^{s_k}$, where $s_k := \lfloor l_1^{1/3} \rfloor$, such that

1. $C = A_i \cup B_i$ and $A_i, B_i \in \mathcal{R}_{d-1}^d$, for all $i = 1, ..., s_k$,
2. \( d(C/A_i, C/B_i) \geq \frac{1}{8} \sqrt{l_k}, \) for all \( i = 1, \ldots, s_k. \)

3. \( A_i \cap B_i \cap A_j \cap B_j = \emptyset \) if \( i \neq j. \)

The proof is reproduced in the Appendix for the convenience of the reader.

### B. Gapped Gibbs sampler implies local clustering

We now proceed to prove the converse statement, namely that if a Gibbs state can be prepared by a gapped Gibbs sampler, then it satisfies strong clustering. The proof relies heavily on the detectability lemma developed in Ref. [40, 41]. We start by pointing out an important connection between Gibbs samplers of local commuting Hamiltonians, and general local frustration-free Hamiltonians. Indeed, let \( \mathcal{L} \) be a primitive Gibbs sampler of a local commuting Hamiltonian, and note that since it satisfy detailed balance, we get that the modified operator

\[
\hat{\mathcal{L}}(f) = \rho^{-1/4} \mathcal{L}(\rho^{1/4} f \rho^{1/4}) \rho^{-1/4}
\]

is Hermitian. In particular, if we represent \( \hat{\mathcal{L}} \) as a matrix on a doubled Hilbert space by the transformation \( |i\rangle \langle j| \mapsto |i\rangle |j\rangle \), we get that \( (-\hat{\mathcal{L}}) \) is a Hamiltonian (hermitian operator) with ground state energy 0. Throughout, we will use the same symbol for the super-operators acting on \( \mathcal{B}_\Lambda \rightarrow \mathcal{B}_\Lambda \) and their associated operator representation acting on \( \mathcal{B}_\Lambda \otimes \mathcal{B}_\Lambda \). It should be obvious from the context which representation we are working with. We furthermore have that \( (-\hat{\mathcal{L}}) \) is local and frustration free (for both the Heat Bath and Davies Liouvillians). If \( (-\hat{\mathcal{L}}) \) is gapped (in the Liouvillian sense) then \( (-\hat{\mathcal{L}}) \) is also gapped (in the Hamiltonian sense). The Gibbs state (density matrix \( \rho \)) is mapped onto a pure state \( |\sqrt{\rho}\rangle = \sqrt{\rho} \otimes |\omega\rangle \), where \( |\omega\rangle = \sum_j |jj\rangle \) is proportional to the maximally entangled state, and satisfies

\[
\hat{\mathcal{L}} |\sqrt{\rho}\rangle = 0.
\]

Similarly, if \( \mathcal{E} \) is a projective conditional expectation, then \( \mathcal{E} \) locally projects onto \( |\sqrt{\rho}\rangle \). We can summarize the correspondence as:

| Commuting Gibbs Sampler | Frustration-free Hamiltonian |
|-------------------------|----------------------------|
| State \( \rho \)        | Ground state \( |\phi\rangle \) |
| Dynamics \( \mathcal{L} \) | Hamiltonian \( H \) |
| Projectors \( \mathcal{E} \) | Ground state projectors \( P \) |
| Gap \( \mathcal{L} \)     | Spectral gap of \( H \) |
| Framework \( \mathcal{L}_p \) | Hilbert spaces \( H \) |

TABLE I: Correspondence between the Gibbs sampler framework and the Hamiltonian complexity framework.

Thus, all of the tools developed for frustration-free Hamiltonians with a unique ground state can be applied to the setting of Gibbs samplers. In particular, we have completely recovered the setting of the detectability lemma and we can invoke the results from Ref. [41], as we now explain.

Throughout the rest of this section we will be considering an \( r \)-local frustration-free Hamiltonian \( -\hat{\mathcal{L}} = \sum_{j \in \Lambda} -\hat{\mathcal{L}}_j \), which has a system size \( (\Lambda) \) independent spectral gap \( \lambda \) and ground
state energy zero. It will in fact be more convenient to work with the modified Hamiltonian

$$\hat{Q}_\Lambda = \sum_{j \in \Lambda} \hat{Q}_j$$

(82)

where each local term \(\hat{Q}_j := (1 - \hat{E}_j^2)\) is a projector. It is not difficult to show (see e.g. [41]) that if each \(\|\hat{L}_j\| \leq K\) for some constant \(K\), then the spectral gap \(\epsilon\) of \(Q\) is bounded as \(\epsilon \geq \lambda/K\). Given that \(\hat{Q}\) also has a unique ground state \(|\sqrt{\rho}\rangle\), all results will also hold for \(-\hat{L}\).

Each term \(\hat{Q}_j\) overlaps with a constant number of other local projectors \(\hat{Q}_k\), so that the terms \(\{\hat{Q}_k\}\) can be partitioned into \(g\) layers where each layer consists of non-overlapping projectors. See Fig. 3 for an illustration of the one-dimensional case when \(r = 2\), and there are only two layers. Define

$$\hat{\Pi}_j = \prod_{k \in \text{layer}(j)} (1 - \hat{Q}_k) = \prod_{k \in \text{layer}(j)} \hat{E}_k$$

(83)

and

$$\hat{\Pi} = \hat{\Pi}_g \cdots \hat{\Pi}_1.$$

(84)

Finally, define \(f(k, g)\) to be the number of sets of pyramids that are necessary to estimate the energy contribution of all the \(Q_j\) terms. In the 1D case illustrated in Fig. 3 we had \(f(k, g) = 2\). In the general case one can derive a crude upper bound: \(f(g, k) \leq (g - 1)k^9\). For more details consult Ref. [41]. Then,

**Lemma 23** [Detectability Lemma [41]] With the notation introduced above, we get

$$\|\hat{\Pi} - \hat{E}_\Lambda\| \leq \frac{1}{(\epsilon/f(k, g) + 1)^{1/3}}.$$  

(85)

Since for all \(j \in \Lambda\), \(\hat{E}_\Lambda \hat{E}_j = \hat{E}_\Lambda\), we get that

$$\lim_{n \to \infty} \hat{\Pi}^n = \hat{E}_\Lambda.$$  

(86)

Then from Lemma 23 we know that there exists constants \(C, \kappa > 0\) such that

$$\|\hat{E}_\Lambda - \hat{\Pi}^n\|_\infty \leq Ce^{-n\kappa},$$  

(87)

where \(\kappa\) is proportional to \(\epsilon\). The exact same reasoning holds true for local projectors:

$$\|\hat{E}_A - \hat{\Pi}_A^n\|_\infty \leq Ce^{-n\kappa},$$  

(88)

where \(\Pi_A\) is constructed from local projectors that intersect \(A\) (i.e. have support on \(A_{ij}\)).

Now, we can use this approximate local projection property to prove the converse of Theorem 21.

**Theorem 24** Let \(\Lambda\) be a finite lattice in \(\mathbb{Z}^d\). Let \(\Phi_\Lambda : \Lambda \mapsto A_\Lambda\) be an \(r\)-local bounded and commuting potential, and let \(\rho\) be the Gibbs state of \(H_\Lambda = \sum_k \Phi_\Lambda(k)\). Denote \((\mathcal{L}_\Lambda, \mathbb{E})\) a local Gibbs sampler of \(\rho\), and suppose that \(\mathbb{E}\) is a projective conditional expectation. If for any \(A \subset \Lambda\), \(\mathcal{L}_\Lambda\) is gapped, then \(\rho\) satisfies strong clustering with respect to \(\mathbb{E}\).
FIG. 3: The setting of the detectability lemma when the Hamiltonian consists of 2-local terms. The two non-overlapping approximate ground state projectors $E_{\text{in}}$ and $E_{\text{out}}$ are shown in red and blue.

**Proof:** Our proof resembles the proof of (weak) clustering for ground states of frustration free Hamiltonians in Ref. [41] (Sec. 6). Consider two subsets $A, B \subset \Lambda$ with $A \cap B \neq \emptyset$, and assume that the overlap has minimal side length $L$. Consider an approximate local projection $\hat{E}_{A \cup B} \approx \hat{\Pi}_{l}^{A \cup B}$, where $\hat{\Pi}_{l}^{A \cup B}$ is restricted to local projective terms that intersect with subset $A \cup B$. If $l \leq L/(gr)$, then we can write $\hat{\Pi}_{l}^{A \cup B} = \hat{E}_{\text{in}} \circ \hat{E}_{\text{out}}$, where $\hat{E}_{\text{in}}$ consists of terms intersecting $\bar{A}$ on the first level, and the light cone resulting from the iterative application of $\hat{\Pi}_{l}^{A \cup B}$. Let $f \in A$, then for $g_{\bar{A}} := E_{B}(f)$ and $g_{\bar{B}} := E_{A}(f)$, because of the frustration freedom and reversibility property of the conditional projective expectations, we get

$$E_{\text{in}}(g_{\bar{A}}) = g_{\bar{A}} \quad (90)$$
$$E_{\text{out}}(g_{\bar{B}}) = g_{\bar{B}} \quad (91)$$

Therefore, by the construction of $\hat{\Pi}_{l}^{A \cup B}$ and Eqns. (90) and (91), we get

$$\langle g_{\bar{A}}, \hat{\Pi}_{l}^{A \cup B}(g_{\bar{B}}) \rangle_{\rho} = \langle g_{\bar{A}}, E_{\text{in}} \circ E_{\text{out}}(g_{\bar{B}}) \rangle_{\rho} = \langle E_{\text{in}}(g_{\bar{A}}), E_{\text{out}}(g_{\bar{B}}) \rangle_{\rho} = \langle g_{\bar{A}}, g_{\bar{B}} \rangle_{\rho} \quad (92)$$

Then, noting that

$$||\hat{\Pi}_{l}^{A \cup B} - E_{A \cup B}||_{2-2,\rho} = ||\hat{\Pi}_{l}^{A \cup B} - \hat{E}_{A \cup B}||, \quad (93)$$

and using the $L_{p}$ Hölder’s inequality, we get

$$|\langle f, (\hat{\Pi}_{l}^{A \cup B} - E_{A \cup B})(h) \rangle_{\rho}| \leq ||f||_{2,\rho}||\hat{\Pi}_{l}^{A \cup B} - E_{A \cup B}||_{2-2,\rho} \quad (94)$$

$$\leq ||f||_{2,\rho}||\hat{\Pi}_{l}^{A \cup B} - E_{A \cup B}||_{2-2,\rho}||g||_{2,\rho}. \quad (95)$$
Eqn. (88), then leads to,
\[
\text{Cov}_{A \cup B}(E_A(f), E_B(f)) = \langle g_A, g_B \rangle - \langle g_A, E_{A \cup B}(g_B) \rangle + C \|g_A\|_2 \|g_B\|_2 e^{-\kappa l}.
\]

\[
\leq \langle g_A, g_B \rangle - \langle g_A, E_{A \cup B}(g_B) \rangle + C \|f\|_2^2 e^{-\kappa l}
\]
\[
= C \|f\|_2^2 e^{-\kappa l}.
\]

(96)

The proof of Thm. 24 can be adapted to show that a gapped Gibbs sampler also implies weak clustering, which in turn shows that strong clustering implies weak clustering.

**Corollary 25** Under the same assumptions as Theorem 24,
\[
\text{Cov}(f, g) \leq c \|f\|_2 \|g\|_2 \|E_{A \cup B}(g_B)\|_2 e^{-d(\Sigma_l, \Sigma_\rho)}/\xi
\]

for some positive $c, \xi$.

**Proof:** The proof is identical to that of Theorem 24 but setting $A \cup B = \Lambda$, and taking $f$ and $g$ instead of $g_A$ and $g_B$ in the covariance.

**Remarks:**

i We point out that a number of results have already been published which show that a gapped Liouvillian implies weak clustering in the ground state [17, 19, 53]. Those results focus on general Liouvillians and their steady states, and hence are typically weaker, but more general than Theorem 24.

ii Using the mapping described in Table 1, it can be seen that the strong clustering condition is essentially equivalent to condition $C_3$ in Ref. 50. Theorem 21 can hence also be seen as an alternative proof of Thm. 3 in Ref. 50. It is furthermore interesting to note that condition $C_3$ in Ref. 50 could in fact be related to a covariance decay condition; a connection which had thus far not be made.

iii The detectability lemma is almost sufficient to show local indistinguishability. Indeed, by Hölder duality, one gets
\[
|\langle f, (\Pi_{A \cup B} - E_{A \cup B})(g) \rangle_\rho| \leq ||f||_1 \|\Pi_{A \cup B} - E_{A \cup B}(g)\|_\infty.
\]

(98)

Thus, if one could show that $||\Pi_{A \cup B} - E_{A \cup B}||_\infty$ is exponentially decaying in $l$, then local indistinguishability would follow. In the framework of frustration-free Hamiltonians and ground states, this would connect LTQO and the detectability lemma in an intriguing way, and could potentially lead to new strategies for proving the area law conjecture (which is implied by LTQO 56).

**VI. ONE DIMENSIONAL MODELS**

In this section, we show that in the special case when the lattice is one dimensional, the system is always gapped and hence satisfies strong clustering. This can be considered as a
partial extension of the celebrated result by Araki that Gibbs states of one dimensional lattice systems always satisfy clustering of correlations \[45\]. The main technical contribution of this section is a proof that strong clustering and weak clustering are equivalent for one dimensional commuting potentials. Intuitively, this is true because strong clustering is in a sense a statement of clustering restricted to a subsystem, where the worst-case boundary conditions are taken into account. Given that in a one dimensional lattice system, the boundary has dimension zero, its contribution only provides a constant multiplicative factor in the clustering statement.

\[\text{FIG. 4: Illustration a subset } A \cup B \subset \Lambda \text{ of a one-dimensional lattice. The boundary } \partial A \cup \partial B \text{ is zero-dimensional.}\]

**Theorem 26 (1D equivalence)** Let \(\Lambda\) be a finite lattice in \(\mathbb{Z}\). Let \(\Phi : \Lambda \rightarrow A_\Lambda\) be an \(r\)-local bounded and commuting potential, and let \(\rho\) be the Gibbs state of \(H_\Lambda = \sum_k \Phi_\Lambda(k)\). Then, \(\rho\) satisfies weak clustering if, and only if, \(\rho\) satisfies strong clustering for some conditional expectation \(E\) of \(\rho\).

**Proof:** By Corollary 25, strong clustering always implies weak clustering. We show the converse below.

Given \(A, B \subset \Lambda\), with \(A \cap B \neq \emptyset\), Proposition 14 shows that the maximum in

\[
\sup_{f \in A_\Lambda} \frac{\text{Cov}_{A \cup B}(E_A(f), E_B(f))}{\text{Var}_{A \cup B}(f)} \leq ce^{-d(\bar{A}, \bar{B})/\xi},
\]  

is reached for an operator \(1 \otimes f\), with \(f \in A_{(A \cup B)\partial}\). Hence, it suffices to work with \(f \in A_{(A \cup B)\partial}\), and \(\rho \propto e^{-\beta H_{A \cup B}}\), where \(H_{A \cup B} = \sum_{k \in A \cup B} \Phi_\Lambda(k)\) (see also the comments after Lemma 20).

Denote \(h = E_A(\tilde{f})\) and \(g = E_B(\tilde{f})\) where \(\tilde{f}\) is the observable achieving the maximum of the LHS of Eqn. (99), which we know acts only on \((A \cup B)\partial = \partial A \cup A \cup B \cup \partial B\) (see Fig. 4). Throughout this proof, we will write the tensor products out explicitly so as to avoid confusion. We furthermore define the modified states \(\sigma_A \equiv \rho_A\), which for any \(A \subset \Lambda\) is the Gibbs state restricted to a subset without the Hamiltonian terms intersecting the boundary of that subset. Since the Hamiltonian of the system is commuting, we can write

\[
\rho_{(A \cup B)\partial} = P(\sigma_{\partial A} \otimes \sigma_{A \cup B} \otimes \sigma_{\partial B})Q,
\]  

where \(P\) acts only on \(\partial A\) and \(A\) (in fact only on a constant-sized region of \(A\) that touches \(\partial A\)) and \(Q\) acts only on \(\partial B\) and \(B\).

Note furthermore that \(h\) has support on the complement of \(A\) in \((A \cup B)\partial\), and similarly for \(g\) with respect to \(B\).

We will prove that local clustering is equivalent to global clustering in the covariance \(\text{Cov}_\rho^{(0)}\). This is sufficient, because as shown in Proposition 15, for commuting Hamiltonians, the two inner products are equivalent for weak clustering. Then we can write,

\[
\langle h, g \rangle^{(0)}_\rho \equiv \text{tr} [\rho hg^\dagger] = \text{tr} [\sigma_{\partial A} \otimes \sigma_{A \cup B} \otimes \sigma_{\partial B} (Q h) (g P)]
\]  

(101)
Since $Qh$ has support on the complement of $A$, we can write it in its Schmidt decomposition with respect to the Hilbert spaces $(\partial_A \partial_B, \langle \cdot, \cdot \rangle_{\sigma_{\partial A} \otimes \sigma_{\partial B}}^{(0)})$ and $(B \setminus A, \langle \cdot, \cdot \rangle_{\sigma_{\partial B \setminus A}}^{(0)})$ as
\[ Qh = \sum_k h_{\partial_A \partial_B}^k \otimes h_{B \setminus A}^k, \quad (102) \]
where $h_{\partial_A \partial_B}^k \in A_{\partial A \partial B}$ and $h_{B \setminus A}^k \in A_{B \setminus A}$ are the Schmidt coefficients satisfying
\[ \text{tr} \left[ \sigma_{\partial A} \otimes \sigma_{\partial B} (h_{\partial A \partial B}^k) (h_{\partial A \partial B}^*) \right] = \delta_{kk} \text{tr} \left[ \sigma_{\partial A} \otimes \sigma_{\partial B} (h_{\partial A \partial B}^k) (h_{\partial A \partial B}^*) \right], \]
and likewise for $\{h_{B \setminus A}^k\}$. Note that $k$ varies from 1 to $(d_{\partial A} d_{\partial B})^2$, where $d_{\partial A}$ is the dimension of the space spanned by the sites in $\partial_A$.

Similarly, since $Pg$ has support on the complement of $A$, we can write it in its Schmidt decomposition with respect to the Hilbert spaces $(\partial_A \partial_B, \langle \cdot, \cdot \rangle_{\sigma_{\partial A} \otimes \sigma_{\partial B}}^{(0)})$ and $(A \setminus B, \langle \cdot, \cdot \rangle_{\sigma_{A \setminus B}}^{(0)})$ as
\[ Pg = \sum_k g_{\partial_A \partial_B}^k \otimes g_{A \setminus B}^k, \quad (103) \]
From Eq. (101), we get
\[ \langle h, g \rangle_{\rho}^{(0)} = \sum_{k,l} \text{tr} \left[ (\sigma_{\partial A} \otimes \sigma_{\partial B} h_{\partial A \partial B}^k (g_{\partial A \partial B}^l)^\dagger \right] \text{tr} \left[ \sigma_{A \setminus B} (h_{B \setminus A}^k) (g_{A \setminus B}^l)^\dagger \right], \quad (104) \]
and thus
\[ |\langle h, g \rangle_{\rho}^{(0)}| \leq \sum_{k,l} \left| \text{tr} \left[ (\sigma_{\partial A} \otimes \sigma_{\partial B} h_{\partial A \partial B}^k (g_{\partial A \partial B}^l)^\dagger \right] \right| \left| \text{tr} \left[ \sigma_{A \setminus B} (h_{B \setminus A}^k) (g_{A \setminus B}^l)^\dagger \right] \right| \]
\[ \leq \sum_{k,l} (\text{tr} \left[ (\sigma_{\partial A} \otimes \sigma_{\partial B} h_{\partial A \partial B}^k (g_{\partial A \partial B}^l)^\dagger \right])^{1/2} (\text{tr} \left[ (\sigma_{\partial A} \otimes \sigma_{\partial B} g_{\partial A \partial B}^l (g_{\partial A \partial B}^l)^\dagger \right])^{1/2} \]
\[ \left| \text{tr} \left[ \sigma_{A \setminus B} (h_{B \setminus A}^k) (g_{A \setminus B}^l)^\dagger \right] \right| \]
Noting that
\[ |\text{tr} \left[ \sigma_{A \setminus B} g_{B \setminus A}^f \right] | \leq c |\text{tr} \left[ \rho_{A \setminus B} g_{B \setminus A}^f \right]|, \quad (105) \]
for some constant $c$ since $\rho_{A \setminus B}$ and $\sigma_{A \setminus B}$ only differ on the boundary of $A \cup B$, which is zero-dimensional.

Then, by global clustering, there exist constants $c, \xi > 0$ such that
\[ |\langle f, g \rangle_{\rho}^{(0)}| \leq c e^{-l/\xi} \sum_{k,l} \| h_{\partial A \partial B}^k \|_{2,(0),\sigma_{\partial A} \otimes \sigma_{\partial B}} \| g_{\partial A \partial B}^l \|_{2,(0),\sigma_{\partial A} \otimes \sigma_{\partial B}} \]
\[ \| h_{B \setminus A}^k \|_{2,(0),\sigma_{A \setminus B}} \| g_{A \setminus B}^l \|_{2,(0),\sigma_{A \setminus B}} \]
By concavity of $x \mapsto x^{1/2}$,
\[ |\langle f, g \rangle_{\rho}^{(0)}| \leq c e^{-l/\xi} d^2 \left( \sum_{k,l} \| h_{\partial A \partial B}^k \|_{2,(0),\sigma_{\partial A} \otimes \sigma_{\partial B}} \| g_{\partial A \partial B}^l \|_{2,(0),\sigma_{\partial A} \otimes \sigma_{\partial B}} \]
\[ \| h_{B \setminus A}^k \|_{2,(0),\sigma_{A \setminus B}} \| g_{A \setminus B}^l \|_{2,(0),\sigma_{A \setminus B}} \right)^{1/2} \]
\[ = c e^{-l/\xi} d^2 (\text{tr} \left[ \sigma' Qh(Qh)^\dagger \right] \text{tr} \left[ \sigma' Pg(Pg)^\dagger \right])^{1/2} \quad (107) \]
with \( d := d_\partial_A d_\partial_B \) and \( \sigma' := \sigma_\partial A \otimes \sigma_\partial B \otimes \sigma_{A UB} \). In the last line, we used that

\[
\text{tr} [\sigma' Q h(Qh)^\dag] = \sum_{k,k'} \text{tr} \left[ (\sigma_\partial A \otimes \sigma_\partial B \otimes \sigma_{A UB})(h_{\partial A \partial B} h_{\partial B} k_k' \otimes h_{\partial B} k_k' \otimes h_{\partial B} k_k') \right],
\]

which follows from the orthogonality of the \( h^k \)s.

The result follows from the bound

\[
\text{tr} [\sigma' Q h(Qh)^\dag] \leq C(r) \| f_{\partial A(B \setminus A) \partial B} \|_{2,(0),\sigma}^2
\]

for a function \( C(r) \) of the interaction range of the Hamiltonian.

Using Theorem [26] we can now show that commuting Gibbs samplers of one dimensional lattice systems are always gapped. At first sight the clustering proof of Araki [45], together with Thm. 26, Thm. 21, and Prop. 15, should suffice to prove that the Davies generator of a 1D commuting Hamiltonian is always gapped. However, Araki’s result has the error term expressed in infinity norm, whereas we need an \( \mathbb{L}_2 \) norm bound (see Eqn. [37]). Here we use methods from the theory of matrix product states to show that the clustering results can in fact be recast in terms of \( \mathbb{L}_2 \) norms.

**Proposition 27** Let \( \Lambda \) be a finite lattice in \( \mathbb{Z} \). Let \( \Phi_\Lambda : \Lambda \rightarrow A_\Lambda \) be an \( r \)-local bounded and commuting potential, and let \( \rho \) be the Gibbs state of \( H_\Lambda = \sum_k \Phi_\Lambda(k) \). Then, the Heat-Bath and Davies samplers are gapped.

**Proof:** Let \( H \) be a commuting Hamiltonian in one dimension with finite range \( r \), and let \( \rho \) be its Gibbs state. Group the sites into blocks of \( r/2 \) sites, so that the Hamiltonian is 2-local. Since the terms are pairwise commuting we can write the Gibbs state as

\[
\rho = \frac{1}{Z} \bigotimes_{i:\text{even}} e^{-\beta H_{i, i+1}} \bigotimes_{i:\text{odd}} e^{-\beta H_{i, i+1}}.
\]

Then it follows that

\[
|\rho^{1/2}\rangle = |\rho^{1/2} \otimes 1 \rangle |\omega\rangle
\]

is a matrix product state with bond dimension bounded by \( 2^r \), where \(|\omega\rangle = \sum_j \langle i,j | \) is proportional to the maximally entangled state. Indeed, for any bipartition \((i, \ldots, j)(j+1, \ldots, n)\), only the term \( e^{-\beta H_{i,j+1}} \) can increase the Schmidt rank. Thus \(|\rho^{1/2}\rangle\) has Schmidt rank bounded by \( 2^r \) in every bipartite cut and by Ref. [47] it is a MPS of bond dimension \( 2^r \).

Note that the MPS is also injective. This can be seen by considering a construction similar to the one in Ref. [46]. Indeed, recall that \(|\rho^{1/2}\rangle\) can be written as a local circuit, with non unitary elements \( e^{-\beta H_i} \), that commute among each other. Therefore, one can construct a transfer operator of dimension \( 2^r \). The transfer operator will be invertible as long as the Gibbs state is unique in the thermodynamic limit, which is always satisfied in 1D [45]. But this is equivalent to the MPS being injective.

Therefore by Refs. [50] (see also the argument of section 4.2 of [51]), the parent Hamiltonian of \(|\rho^{1/2}\rangle\) is gapped. Hence we can use the detectability lemma on the parent Hamiltonian of \(|\rho^{1/2}\rangle\), and Corollary 25 to get

\[
\text{Cov}^{(0)}(f, g) \leq C \| f \|_{2,(0),\rho} \| g \|_{2,(0),\rho} e^{-\kappa_d(|\Sigma f, \Sigma g|)}
\]

From Prop. [15] the same statement holds in the symmetric covariance \( \text{Cov} \) and \( \mathbb{L}_2 \) norm \( \| f \|_{2,\rho} \). Then, invoking the equivalence between strong and weak clustering for one dimensional systems of Thm. [26] we find that strong clustering holds, which in turn implies that all one-dimensional Gibbs samplers of commuting Hamiltonians are gapped via Thm. [21].

\[\square\]
VII. THE HIGH TEMPERATURE PHASE

In this section we show that for \( r \)-local commuting Hamiltonians on a \( d \)-dimensional lattice there is a temperature \( T_c(r, d) \) independent of the lattice volume such that for every \( T \geq T_c \), both the Heat Bath and the Davies generators have a constant spectral gap. Thus the Gibbs state of every commuting Hamiltonian can be created efficiently on a quantum computer and in nature at high enough temperatures. The result follows from the mapping of Section V.B between Liouvillians satisfying detailed balance and frustration-free Hamiltonians together with a technique due to Knabe \cite{63} for lower bounding the spectral gap of local Hamiltonians.

**Theorem 28** Let \( \Lambda \) be a finite lattice in \( \mathbb{Z}^d \). Let \( \Phi_\Lambda : \Lambda \mapsto \mathcal{A}_\Lambda \) be an \( r \)-local bounded and commuting potential, and let \( \rho \) be the Gibbs state of \( H_\Lambda = \sum_k \Phi_\Lambda(k) \). Then there exists a constant \( T_c(r, d) \) such that for every \( T \geq T_c \), the Heat Bath generator \( \hat{L}^H_\Lambda \) has a gap independent of \( |\Lambda| \).

**Proof:**

By Eqs. (13) and (34), for any \( f \in \mathcal{A}_\Lambda \),

\[
-\hat{L}^H_\Lambda(f) = \sum_{k \in \Lambda} \rho^{1/4}(\text{id} - E^f_k)(\rho - 1/4 f \rho - 1/4)\rho^{1/4}.
\]  

(112)

where

\[
E^f_k = \text{tr}_k[\eta^f_k \eta^f_k^\dagger] = D_k \sum_j p_j \rho^{-1/4}_k U_j \rho^{-1/4}_k \rho^{1/2} g \rho^{1/2} \rho^{-1/4}_k U_j^\dagger \rho^{-1/4}_k,
\]  

(114)

with \( \rho_y := \text{tr}_k(\rho) \), \( D_k \) the local Hilbert space dimension dimension, and \( \{p_j, U_{j,k}\} \) an ensemble of depolarizing unitaries spanning \( B_k \) such that for every \( f \in B_k \), \( \sum_j p_j U_j f U_j^\dagger = \text{tr}(f) \mathbb{1}_k / D_k \).

Using the mapping of Section V.B, the corresponding Hamiltonian \( -\hat{L}^H_\Lambda \) acting on \( \mathcal{B}_\Lambda \otimes \mathcal{B}_\Lambda \) is given by

\[
\hat{L}^H_\Lambda := \sum_{k \in \Lambda} \hat{L}^H_k
\]

\[
= \sum_{k \in \Lambda} D_k \left( \rho^{1/4} \otimes \rho^{1/4} \right) \left( \frac{1}{4} - \sum_j p_j U_{j,k} \otimes \bar{U}_{j,k} \right) \left( \rho^{-1/4} \otimes \bar{\rho}^{-1/4} \right) \left( \rho^{1/4} \otimes \rho^{1/4} \right)
\]

\[
= \sum_{k \in \Lambda} D_k \left( \rho^{1/4} \otimes \bar{\rho}^{1/4} \right) \left( \rho^{-1/4} \otimes \bar{\rho}^{-1/4} \right) \left( \mathbb{1} - w_{kk} \right) \left( \rho^{-1/4} \otimes \bar{\rho}^{-1/4} \right) \left( \rho^{1/4} \otimes \bar{\rho}^{1/4} \right),
\]  

(115)

with \( w_{kk} = \sum_{i,j} \langle jj | \langle ii | \) the maximally entangled state on \( B_k \otimes \bar{B}_k \) tensored with the identity off site \( k \). Note that each \( \hat{L}^H_k \) is local, with its locality given by the interaction range \( r \). Moreover as explained in Section V.B \( \hat{L}^H_k \) is frustration free.

Given that when \( T \to \infty \), \( \rho \to \mathbb{1}_\Lambda \) and so also \( \rho^{-1/4} \to \mathbb{1}_\Lambda \), the Hamiltonian \( -\hat{L}^H_\Lambda \) converges in the limit \( T \to \infty \) to a non-interacting Hamiltonian given by \( -\hat{L}^H_\Lambda = \sum_{k \in \Lambda} (\text{id} - w_{kk}) \), whose spectral gap is one. The statement of the theorem will follow by showing that there is a constant \( T_c(k, d) \) such that the Hamiltonian \( -\hat{L}^H_\Lambda \) has almost-commuting terms for all \( T \geq T_c(k, d) \), with commutators sufficiently small that the spectral gap is also a constant.
We proceed by employing a well-know technique due do Knabe \cite{63} for lower bounding the spectral gap of local Hamiltonians. First we consider the Hamiltonian \((-\tilde{L}_H^\Lambda) := \sum_{k \in \Lambda} P_{H,T}^k\), where \(P_{H,T}^k\) is the projector onto the non-zero eigenspace of \((-\hat{L}_H^k)\).

We have that in the limit \(T \to \infty\), \(P_{H,T}^k \to (\text{id} - w_{kk})\). Moreover, we also have that
\[
\Delta(-\tilde{L}_H^\Lambda) \geq \Omega(\Delta(-\hat{L}_H^\Lambda)). \tag{116}
\]
(see e.g. Section 2 of \cite{41}).

We now apply the Knabe bound \cite{63}. It says that given any \(k\)-local frustration-free Hamiltonian on a \(d\)-dimensional lattice formed by local projector terms, then there is an integer \(N(k,d)\) and a real number \(\lambda(k,d) < 1\) (that can be computed explicitly given the lattice and that are independent of the volume) such that
\[
\Delta(H) \geq \Omega \left( \min_{S:|S|=N} \Delta(H_S) - \lambda \right), \tag{117}
\]
where the minimum is taken over all connected sublattices of size \(N\).

For every fixed region \(S\), \(-\tilde{L}_H^S\) converges to \(\sum_{k \in S}(\text{id} - w_{kk})\) in the limit \(T \to \infty\). Since \(\sum_{k \in S}(\text{id} - w_{kk})\) one, we find that given \(N(k,d)\) and \(\lambda(k,d)\), there always exists a \(T_c\) such that for all \(T \geq T_c\),
\[
\min_{S:|S|=N} \Delta(-\tilde{L}_S^H) > \lambda, \tag{118}
\]
so indeed
\[
\Delta(-\tilde{L}_H^\Lambda) \geq \Omega(1), \tag{119}
\]
and the statement follows from Eq. \((116)\). \(\square\)

\section{VIII. OUTLOOK}

We have introduced a unified framework for analyzing quantum Gibbs samplers of Hamiltonians with commuting local terms. This includes two independent prescriptions for constructing local quantum dynamical semigroups (i.e. Gibbs samplers) that uniquely drive the system to the Gibbs state of a given commuting Hamiltonian \(H\). Associated to each Gibbs sampler, we construct local projectors onto the Gibbs state. The main result of the paper is a theorem which shows the equivalence between the rapid time convergence of the Gibbs sampler, and a new form of strong exponential clustering in the Gibbs state. We also explore how this new strong form of clustering is connected to more conversational notions of correlation decay. Finally, building upon the main theorem, we show that all Gibbs samplers of commuting Hamiltonians on a one dimensional lattice have a gap which is independent of the system size. Above a universal critical temperature, this holds true also for higher dimensional lattice models.

These results are important and useful for a number of reasons. The two Gibbs samplers that we analyze serve complementary purposes in the literature. The Davies generators are meant to model the thermal dynamics that naturally emerge for a system weakly interacting with a thermal reservoir. This situation is very generic, especially for quantum optics based experiments, hence our analysis potentially provides crucial information on time scales for optical lattice simulators, and related setups. Secondly, the heat bath generators are a simple constructive semigroup which could be useful for quantum simulations. For certain tasks, it is easier to work with than the Davies maps (ex: Ref. \cite{31}). Finally, as outlined below, in
the form of open questions, our main theorem provides a structural backbone relating several important notions, including: criticality, stability, topological order, classicality, etc.

One major drawback of our framework is that it is not very well suited for Hamiltonians with non-commuting local terms. Indeed, it is easy to see that in general $L^D$, $L^H$, $E^p$, and $E^L$ all become non-local when $H$ is non-commuting, and very little of the framework can be recovered. It would be very interesting to explore extensions of our results to non-commuting Hamiltonians, as it would incorporate many of the more interesting models in quantum statistical mechanics. Still in the setting of commuting Gibbs samplers, it would be very important to figure out whether the spectral gap of the Gibbs sampler is equivalent to the Log-Sobolev inequality. This equivalence holds for classical Gibbs samplers, and allows a tremendous strengthening of Theorems 21 and 24. If one were able to extend the theory to Log-Sobolev inequalities, then it would be able to show that a Gibbs samplers have a relaxation time which is either exponential in the number of sites or logarithmic; i.e. there is no intermediate mixing regime [36].

Another very interesting direction to be explored in more detail is the connection between Gibbs samplers and frustration-free Hamiltonians outlined in Sec. V B. Many relevant problems in Quantum Hamiltonian complexity [58, 59] involve frustration-free Hamiltonians, and it is conceivable that by exploiting this new connection the fields of quantum Gibbs samplers and Quantum Hamiltonian complexity can mutually benefit from their respective methods. In particular, it would be very interesting to understand to what extent the theory of Hypercontractive semigroups [20, 60, 61] can be applied to problems of Hamiltonian complexity.

We conclude with a list of questions and conjectures together with some compelling implications:

1. **The equivalence of weak and strong clustering in higher dimensions**

   Theorem [26] shows that for 1D Lattice systems, the strong and weak clustering conditions are equivalent, up to a multiplicative constant. Is this also true in higher dimensions? Although the proof of theorem [26] clearly does not carry through to higher dimension because it relies heavily on a Schmidt decomposition of the boundary terms, there are reasons to believe that the equivalence could extend to two dimensional lattice systems. Indeed, the conditional covariance, and the strong clustering condition are an attempt to recover the situation when a state is clustering on a subset of the full lattice independently of the “boundary conditions” that are chosen around the lattice restriction. In classical lattice systems, exotic phase transitions can be driven along the boundary of a material whose bulk is in a thermally non-critical phase [43]. Such a phenomenon has been coined a boundary phase transition, and appears not to be such an exotic phenomenon in three and higher spacial dimensions [62]. However, in two dimensional classical spin systems, this phenomenon cannot occur [44]. The heuristic reason for this is that the boundary of a two dimensional lattice model is effectively a one dimensional lattice spin system, for which we know that no critical behavior can be found.

   Hence it is tempting to conjecture that: the Gibbs state of a commuting Hamiltonian on a two dimensional lattice satisfies weak clustering if and only if it satisfies strong clustering.

2. **The behavior of correlations as the temperature goes to zero**

   Physicists study the ground states of Hamiltonians, because in many situations it is believed that the actual state of the experiment is a Gibbs state at very low temperature, and the essential physics is governed by the properties of the ground state. The framework of Gibbs samplers provides a good setting for testing or confirming this intuition. In particular, if the ground state of a commuting Hamiltonian satisfies certain constraints on spacial correlations then one might expect that this still holds true at small non-zero temperature.
We therefore raise the following questions: (i) if the ground state of a local commuting Hamiltonian satisfies clustering of correlations, does the same hold true for the Gibbs state at some/all non-zero temperature? (ii) if a local commuting Hamiltonian satisfies LTQO, does the Gibbs sampler satisfy local indistinguishability?

Partial answers can be given to these questions (in the operator norm), by using the approximation results in Ref. [25], however a full answer is still elusive. One of the bottlenecks is that the in the Gibbs sampler setting, we have access to the machinery of $\mathbb{L}_p$ norms whereas for (ground)-states the only natural norms are the operator norm for observables, and the inner product for states. Our theorems all depend on $\mathbb{L}_p$ bounds, so a proper interpolation between zero temperature and finite temperature results is not obvious. A candidate for an $\mathbb{L}_p$ quasi-norm on states is: for $f \in A_A$ and a pure state $\varphi \in S_A$, define $\| f \|_{p,\varphi} := \langle \varphi | f^p | \varphi \rangle$ with $1 \leq p \leq \infty$. However, most interesting commuting Hamiltonian models, such as those exhibiting topological order, have a degenerate ground subspace, and with no preferential state it is hard to work with the semi-norms $\| f \|_{p,\varphi}$. Hence, important obstacles still remain.

3. Absence of self-correction for 2D commuting Hamiltonians

If the above two questions turn out to be true, then it would likely lead to a very strong result in the theory of self-correcting quantum memories: topological order of 2D commuting Hamiltonians on a Lattice is unstable under thermal noise.

By self-correcting memory, we mean a Hamiltonian with a topologically stable ground subspace that remains a metastable subspace under thermal noise for a time that grows exponentially in the linear system size $L$. The thermal noise is usually modeled by Davies generators \([14] [15]\), or by a diagonal variant of it \([64]\). The prototypical example of a genuinely self-correcting memory is the 4D toric code \([15]\). There have been a number of no-go theorems for self-correcting memories for 2 and 3 dimensional stabilizer codes \([65] [66] [68] [69]\). All of the existing no-go theorems prove that under certain assumptions on the Hamiltonian, it only takes a constant amount of energy to flip from one logical eigenstate to another. According to the heuristic Arrhenius Law, which states that the survival time scales as an exponential of the free-energy barrier, this would prevent the system from being self-correcting.

Arrhenius’ law is known to be neither necessary nor sufficient in general for the existence of metastable states, so it would be desirable to have direct proofs that certain classes of Hamiltonians are not good quantum memories. Showing that the Davies generators of all 2D commuting Hamiltonians on a lattice are gapped would provide a definitive blow to self-correction in 2D, and would nicely complement the results in \([68] [69]\).

Assuming questions 1) and 2) are shown to be true, the argument for a no-go theorem would go as follows:

If $H_A$ satisfies a specific form of Local Topological Quantum Order (LTQO), similar to the one defined in Ref. \([17]\), then the Davies generators $L^D$ are gapped, and hence, no state (quantum or classical) can survive for a time longer than polynomial in the system size.

We first define $LTQO_p$. Let $H_A$ be a Hamiltonian restricted to subset $A \subseteq \Lambda$, and let $B \subseteq A$. If for any two ground states $\phi, \varphi$ of $H_A$, we have

\[
| \langle \phi | f | \phi \rangle - \langle \varphi | f | \varphi \rangle | \leq c \langle \phi | f | \phi \rangle^{1/p} e^{-d(B,\partial A)/\xi},
\]

(120)

for any local observable $f \in A_B$, and some constants $c, \xi > 0$, then we say that $H_A$ satisfies $LTQO_p$. In particular all stabilizer hamiltonians satisfy $LTQO_p$ for all $p \geq 1$ since the RHS of Eqn. (120) is strictly zero beyond some constant distance.

We therefore raise the following questions: (i) if a Hamiltonian satisfies $LTQO_p$, then its ground state is clustering in the following sense: there exist constants $c, \xi > 0$ such that

\[
| \langle \phi | f g | \phi \rangle - \langle \phi | f | \phi \rangle \langle \phi | g | \phi \rangle | \leq c \langle \phi | f | \phi \rangle \langle \phi | g | \phi \rangle^{1/p} e^{-d(S_f, S_g)/\xi}
\]

(121)
If one is then able to show that clustering in the form of Eqn. (121) also holds for non-zero temperature (i.e. question 2. above), then one would recover weak clustering in the Gibbs state for \( p = 2 \). If in turn, weak and strong clustering are equivalent for 2D Gibbs samplers (question 1.), then one gets that topological order \((LTQO_2)\) implies that the Gibbs sampler is gapped for all finite temperatures.

This type of reasoning, unrigorous at this point, shows the power of our main theorems (Thms. [21] and [24]) in terms of relating static and dynamical properties of spin systems in thermal equilibrium.

Acknowledgements: We gratefully acknowledge fruitful discussions with T. Cubitt, J. Eisert, M. Friesdorf, A. Lucia, F. Pastawski, K. Temme, and R. F. Werner. MJK was supported by the Alexander von Humboldt foundation and by the EU (SIQS, RAQUEL). FB was supported by an EPSRC Early Career fellowship.

[1] U. Haagerup. Lp-spaces associated with an arbitrary von neumann algebra. *Algebres d'opérateurs et leurs applications en physique mathématique*, CNRS, 15:175–184, 1979.
[2] M. Terp. Lp spaces associated with von neumann algebras. *Notes, Math. Institute, Copenhagen Univ.*, 3, 1981.
[3] K. Binder and D. Heermann, Monte Carlo simulation in statistical physics: an introduction, Springer 2010.
[4] A. Riera, C. Gogolin, and J. Eisert, Thermalization in nature and on a quantum computer, Phys. Rev. Lett. 108, 080402 (2012).
[5] A. J. Short and T. C. Farrelly, Quantum equilibration in finite time New J. of Phys., 14, 013063 (2012).
[6] M. P. Mueller, E. Adlam, L. Masanes, and N. Wiebe, Thermalization and canonical typicality in translation-invariant quantum lattice systems, arXiv preprint [arXiv:1312.7420]
[7] D. A. Levin, Y. Peres, E. L. Wilmer, Markov chains and mixing times, American Mathematical Soc. (2009)
[8] P. Diaconis, The markov chain monte carlo revolution, Bulletin of the American Mathematical Society, 46, 179 (2009).
[9] D. Poulin, P. Wocjan, Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer, Phys. Rev. Lett. 103, 220502 (2009).
[10] B. M. Terhal, D. P. DiVincenzo, Problem of equilibration and the computation of correlation functions on a quantum computer, Phys. Rev. A 61 (2), 022301 (2000).
[11] K. Temme, T.J. Osborne, K.G. Vollbrecht, D. Poulin, F. Verstraete, Quantum Metropolis Sampling, Nature 471:87,(2011)
[12] A. Riera, C. Gogolin, J. Eisert, Thermalization in Nature and on a Quantum Computer, Phys. Rev. Lett. 108, 080402 (2012).
[13] E. Dennis, A. Kitaev, A. Landahl and J. Preskill, Topological quantum memory, J. Math. Phys. 43, 4452 (2002).
[14] R. Alicki, M. Horodecki, P. Horodecki, R. Horodecki, On thermal stability of topological qubit in Kitaev’s 4D model, Open Syst. Inf. Dyn. 17 (2010).
[15] R. Alicki, M. Fannes, M. Horodecki, On thermalization in Kitaev’s 2D model, J. Phys. A: Math. Theor. 42 (2009) 065303.
[16] R. Alicki and K. Lendi, Quantum Dynamical Semigroups and Applications, Lecture Notes in Physics, 286, Springer (1987).
[17] T. Cubitt, A. Lucia, S. Michalakis, D. Perez-Garcia, Stability of local quantum dissipative systems, arXiv:1303.4744
[18] A. Guionnet, B. Zegarlański, Lectures on logarithmic Sobolev inequalities, Springer, (2003).
[19] M. J. Kastoryano and J. Eisert, Rapid mixing implies exponential decay of correlations, J. Math. Phys. 54, 102201 (2013)
[20] M. J. Kastoryano, K. Temme, Quantum logarithmic Sobolev inequalities and rapid mixing, J.
[21] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, F. Verstraete, The $\chi^2$ - divergence and Mixing times of quantum Markov processes. J. Math. Phys. 51, 122201 (2010).
[22] R.L. Dobrushin. Description of a random field by means of conditional probabilities and the conditions governing its regularity. Theor. Prob. Appl. 13:19717224, 1968.
[23] O.E. Lanford. Observables at infinity and states with short range correlations in statistical mechanics. Comm. Math. Phys. 13:19417215, 1969.
[24] C. Schoen, E. Solano, F. Verstraete, J. I. Cirac, M. M. Wolf. Sequential generation of entangled multi-qubit states. Phys. Rev. Lett. 95, 110503 (2005).
[25] M.B. Hastings. Solving Gapped Hamiltonians Locally. Phys. Rev. B 73, 085115 (2006).
[26] Kristan Temme, Lower bounds to the spectral gap of Davies generators. Journal of Mathematical Physics, 54(12):122110, 2013.
[27] E.B. Davies. Generators of dynamical semigroups. Journal of Functional Analysis 34, 421 (1979).
[28] M. Kliesch, T. Barthel, C. Gogolin, M. Kastoryano, J. Eisert. A dissipative quantum Church-Turing theorem. Phys. Rev. Lett. 107, 120501 (2011)
[29] Man-Hong Yung and Aln Aspuru-Guzik. A Quantum-Quantum Metropolis Algorithm. Proc. Natl. Acad. Sci. USA 109, 754 (2012)
[30] Maris Ozols, Martin Roetteler, Jrmie Roland. Quantum rejection sampling. Proceedings of the 3rd Conference on Innovations in Theoretical Computer Science (ITCS’12), ACM Press, 2012, pages 290-308
[31] A. W. Majewski, B. Zegarlinski, quantum stochastic dynamics I: spin systems on a lattice, MPEJ, (1995).
[32] A. W. Majewski, R. Olkiewicz, B. Zegarlinski, Dissipative dynamics for quantum spin systems on a lattice, J. Phys. A: Math. Gen. 31 (1998) 2045
[33] A. W. Majewski, B. Zegarlinski, Quantum stochastic dynamics II, Reviews in Mathematical Physics, 8, 689 (1996).
[34] D. Petz, Quantum Information Theory and Quantum Statistics, Theoretical and Mathematical Physics, Springer, Berlin Heidelberg (2008).
[35] N. Yoshida, The equivalence of the Log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice, Ann. Inst. H. Poincare, Prob. et Stat. 37, 223 (2001)
[36] N. Yoshida, Relaxed Criteria of the Dobrushin-Shlosman Mixing Condition, J. Stat. Phys. 87, 1, (1997)
[37] E. B. Davies, One-parameter semigroups Academic press London (1980).
[38] F. Cesi, Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields, Probab. Theory Relat. Fields 120, 569-584 (2001)
[39] L. Bertini, N. Cancrini, F. Cesi, The spectral gap for a Glauber-type dynamics in a continuous gas, Annales de l’Institut Henri Poincare, 38, (2002) 9117108
[40] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, The detectability lemma and quantum gap amplification, Proceedings of the 41st annual ACM symposium on Theory of computing, 417 (2009).
[41] D. Aharonov, I. Arad, Z. Landau, U. Vazirani, The detectability lemma and quantum gap amplification, Communications of the ACM, 417 (2011).
[42] F. Martinelli, Relaxation Times of Markov Chains in Stastical Mechanics and Combinatorial Structures, Probability on Discrete Structures, Springer (2000)
[43] F. Martinelli, Lectures on Glauber dynamics for discrete spin systems, Springer, Lecture Notes in Mathematics Volume 1717, 93, (1999).
[44] F. Martinelli, E. Olivieri, R. H. Schonmann, For 2D lattice spin systems weak mixing implies strong mixing, Comm. Math. Phys. 165, 33 (1994).
[45] H. Araki, Gibbs states of a one dimensional quantum lattice, Comm. Math. Phys. 14, 120 (1969).
[46] D. Pérez-García, F. Verstraete, J. I. Cirac and M. M. Wolf, PEPS as unique ground states of local Hamiltonians, Quant. Inf. Comp., 8(6), 650-663 (2008).
[47] G. Vidal. Efficient simulation of one-dimensional quantum many-body systems. Phys. Rev. Lett. 93, 040502 (2004).
[48] M. Fannes, R. F. Werner, Boundary Conditions for Quantum Lattice Systems, Helvetica physica acta, 68, 635, (1995).
[49] M. B. Hastings, Entropy and Entanglement in Quantum Ground States, Phys. Rev. B 76, 035114
Proof of Lemma 20. We consider the expression for $\lambda_\Lambda(A)$, and note that

$$\lambda_\Lambda(A) = \inf_{f \in A_\Lambda} \frac{-\text{tr} [\Gamma_\rho(f) L^D_f]}{\text{tr} [Q_A(f) \Gamma_\rho(Q_A(f))]}$$

(122)

$$= \inf_{g \in A_\Lambda} \frac{-\text{tr} [g L^D_g]}{\text{tr} [\hat{Q}_A(g) \hat{Q}_A(g)]},$$

(123)

where we made the replacement $g \equiv \Gamma_\rho^{1/2}(f)$, and $\Gamma_\rho(f) = \rho^{1/2} f \rho^{1/2}$. We defined the operators $Q_A(f) := f - \mathbb{E}_A(f)$, $L^D_f = \Gamma_\rho^{1/2} L^D_A \Gamma_\rho^{-1/2}$, and likewise $\hat{Q}_A = \Gamma_\rho^{1/2} Q_A \Gamma_\rho^{-1/2}$. We
Note that \( \hat{L}_A^D \) and \( \hat{Q}_A \) are both hermitian operators, so Eqn. (123) is a generalized eigenvalue equation which can be recasted as

\[ \lambda_A(A) = \inf \{ \lambda \mid \det(\hat{L}_A^D + \lambda \hat{Q}_A^2) = 0 \} \] (124)

Now, we will show that \( \hat{Q}_A \) acts non-trivially only on \( A \partial \). Similarly, we can show that \( \hat{L}_A^D \) can be written as \( \hat{K}_{A \partial} \otimes \mathbb{1}_{A \partial} \) for some hermitian operator \( \hat{K}_{A \partial} \). Then, Eqn. (124) can be rewritten as

\[ \lambda_A(A) = \inf \{ \lambda \mid \det((\hat{K}_{A \partial} + \lambda \hat{Q}_{A \partial}^2) \otimes \text{id}_{(A \partial)^r}) = 0 \} \] (130)

Recalling now that \( \det(A \otimes 1) = \det(A)^n \), where \( n \) is the dimension of matrix \( A \), we get that

\[ \lambda_A(A) = \inf \{ \lambda \mid \det((\hat{K}_{A \partial} + \lambda \hat{Q}_{A \partial}^2) \otimes \text{id}_{(A \partial)^r}) = 0 \} = \inf \{ \lambda \mid \det(\hat{K}_{A \partial} + \lambda \hat{Q}_{A \partial}^2) = 0 \} = \lambda_A(A), \] (132)

completing the proof.

**Proof of Lemma 22.** Let \( C := [a_1, b_1] \times \ldots \times [a_d, b_d] \in \mathcal{R}_k^d/\mathcal{R}_{k-1}^d \). We can assume that \( a_n = 0 \) and \( b_n \leq l_{k+n} \) for \( n = 1, \ldots, d \). Then necessarily \( b_d > l_k \), since otherwise \( C \in \mathcal{R}_{k-1}^d \). Define

\[ A_i := [0, b_1] \times \ldots \times [0, b_{d-1}] \times [0, \frac{b_d}{2} + \frac{2i}{8} \sqrt{l_k}], \] (133)

\[ B_i := [0, b_1] \times \ldots \times [0, b_{d-1}] \times [\frac{b_d}{2} + \frac{1}{8} \sqrt{l_k}, b_d] \] (134)

We have \( d(C/A - i, C/B - 1) = \frac{1}{2} \sqrt{l_k} \). Furthermore,

\[ \frac{b_d}{2} + \frac{2s_k}{8} \sqrt{l_k} \leq \frac{d+1}{2} + \frac{1}{4} s_k/6 \leq 3 \left( \frac{1}{4} k \right)^{5/6} \leq l_{k+1+d} \] (135)

which together with the fact that \( l_k \leq b_d \), implies that \( A_i \) and \( B_i \) are both subsets of \( C \). Moreover, since for all \( i = 1, \ldots, s_k \)

\[ \frac{b_d}{2} + \frac{2i}{8} \sqrt{l_k} \leq l_k, \quad b \leq l_{k+1}, \ldots, b_{d-1} \leq l_{k+1+d} \] (136)

we find that \( A_i \) belongs to \( \mathcal{R}_{k-1}^d \). The sets \( B_i \)'s also belong to \( \mathcal{R}_{k-1}^d \), since they are smaller than the \( A_i \)'s.