



Rakovszky, T., Placke, B., Breuckmann, N. P., & Khemani, V. (2024). *Bottlenecks in quantum channels and finite temperature phases of matter*. arXiv.org. <https://doi.org/10.48550/arXiv.2412.09598>

Early version, also known as pre-print

License (if available):  
CC BY

Link to published version (if available):  
[10.48550/arXiv.2412.09598](https://doi.org/10.48550/arXiv.2412.09598)

[Link to publication record on the Bristol Research Portal](#)  
PDF-document

## University of Bristol – Bristol Research Portal

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>

# Bottlenecks in quantum channels and finite temperature phases of matter

Tibor Rakovszky,<sup>1,2,3</sup> Benedikt Placke,<sup>4</sup> Nikolas P. Breuckmann,<sup>5</sup> and Vedika Khemani<sup>1</sup>

<sup>1</sup>*Department of Physics, Stanford University, Stanford, California 94305, USA*

<sup>2</sup>*Department of Theoretical Physics, Institute of Physics,*

*Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary*

<sup>3</sup>*HUN-REN-BME Quantum Error Correcting Codes and Non-equilibrium Phases Research Group,*

*Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary*

<sup>4</sup>*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*

<sup>5</sup>*School of Mathematics, University of Bristol, Bristol BS8 1UG, United Kingdom*

We prove an analogue of the “bottleneck theorem”, well-known for classical Markov chains, for Markovian quantum channels. In particular, we show that if two regions (subspaces) of Hilbert space are separated by a region that has very low weight in the channel’s steady state, then states initialized on one side of this barrier will take a long time to relax, putting a lower bound on the mixing time in terms of an appropriately defined “quantum bottleneck ratio”. Importantly, this bottleneck ratio involves not only the probabilities of the relevant subspaces, but also the size of off-diagonal matrix elements between them. For low temperature quantum many-body systems, we use the bottleneck theorem to bound the performance of any quasi-local Gibbs sampler. This leads to a new perspective on thermally stable quantum phases in terms of a decomposition of the Gibbs state into multiple components separated by bottlenecks. As a concrete application, we show rigorously that weakly perturbed commuting projector models with extensive energy barriers (including certain classical and quantum expander codes) have exponentially large mixing times.

*Introduction.* Quantum channels are the most general kind of quantum evolutions [1–3]. Physically, they describe the dynamics of systems coupled to equilibrium and non-equilibrium environments. Computationally, they form a paradigm for quantum algorithms, such as generalizations of the Monte Carlo method [4–6]. An important context where these two perspectives intersect, is the problem of quantum Gibbs sampling [7–11]: the design of quantum channels, with favorable properties (e.g., locality) that are guaranteed to approach the Gibbs state  $\rho_G \propto e^{-H/T}$  for some chosen quantum Hamiltonian  $H$  at temperature  $T$ . From a physical perspective, a Gibbs sampler models the behavior of the system coupled to a thermal bath; while algorithmically, on a quantum computer, it can play a role analogous to that of a classical Markov-Chain Monte-Carlo algorithm. Efficient local quantum Gibbs samplers, obeying a quantum version of the detailed balance condition [12–14], have only been discovered very recently, and better understanding their properties is an important challenge [15–18].

A fundamental property of Markov chains is their *mixing time* [19], the minimal time it takes to get close to the steady state distribution starting from any initial state. The mixing time is thus closely related to the convergence of Monte Carlo algorithms and putting upper and lower bounds on it is of central importance. One standard approach that lower bounds the mixing time is the *bottleneck method* [19]: it involves finding a decomposition of the configuration space into two halves such that both have significant probabilities in the steady state, but the flow induced by the Markov chain between them is small; one can then show that a state initialized in one

half takes a long time to relax to the true steady state.

In this work, we generalize the bottleneck approach from Markov chains to Markovian quantum channels. The usual proof of the bottleneck theorem [19] involves expanding the transition probability between two regions as a sum over probabilities of paths. Thus, it is not a priori obvious whether a similar result applies to quantum systems where interference effects make probabilities non-additive and the system might tunnel through energy barriers quantum mechanically<sup>1</sup>. Here, we show that an appropriate generalization of the bottleneck theorem *does hold*: if one can find a subspace such that in order to leave it, any quantum trajectory must pass through a low-weight “bottleneck” region, then the subspace hosts a long-lived approximate steady state, lower bounding the channel’s mixing time in terms of a “quantum bottleneck ratio” (Fig. 1). The important difference is that this lower bound involves not only probabilities but also the *coherences* (off-diagonal matrix elements) present in the steady state density matrix.

To apply this result to quantum Gibbs samplers, we establish an appropriate bottleneck condition that lower bounds the mixing time of *any* Gibbs sampler with  $k$ -local Kraus operators. We argue that this should lead to superpolynomially long mixing times in many cases of interest, such as the low temperature regime of certain Ising and toric code models. We further illustrate these ideas by proving exponential mixing times in a family

<sup>1</sup> One context where analogs of the bottleneck theorem have been proven is for unitary quantum walks [20, 21].

of Hamiltonians related to good classical and quantum error correcting codes.

Beyond a bound on mixing times, the bottleneck theorem provides a novel dynamical perspective on finite-temperature phases of matter. The bottleneck ratio formalizes the idea of a *free energy barrier* separating the Gibbs state into multiple distinct components. Our theorem shows that when such barriers are present (as are expected in many non-trivial phases of matter at low temperatures), *any* local dynamics that obeys detailed balance will feature multiple near-degenerate steady states. This provides an analogy with non-trivial zero temperature phases, which are often characterized in terms of ground state degeneracies, and places thermal phases within the broader framework of phases in open quantum systems [22].

*Quantum channels and mixing times.* A quantum channel [2]  $\mathcal{M}$  is a super-operator that maps any density matrix  $\rho$  over some Hilbert space  $\mathcal{H}$  to another density matrix  $\mathcal{M}[\rho]$ . In the following we will make use of the Kraus representation of channels,  $\mathcal{M}[\rho] = \sum_i K_i \rho K_i^\dagger$  where  $\{K_i\}$  are a set of operators with the property  $\sum_i K_i^\dagger K_i = \mathbb{1}$ . We will be interested in Markovian (memoryless) evolutions, where the state at time  $t$  is obtained by repeated application of the same channel,  $\rho(t) = \mathcal{M}^t[\rho(0)]$  or, more generally, by a sequence of different channels,  $\rho(t) = (\mathcal{M}_t \dots \mathcal{M}_1)[\rho(0)]$ .

One useful representation of channels is in terms of *quantum trajectories* [23, 24]. These are an ensemble of time-evolving *pure* quantum states, where the Kraus operators correspond to random “quantum jumps”. The probabilities associated to the different trajectories are chosen such that on average, they reproduce the evolution of  $\rho(t)$  under the channel  $\mathcal{M}$ .

Every channel has at least one steady state  $\mathcal{M}[\rho_{\text{ss}}] = \rho_{\text{ss}}$  [25]. The steady state is unique, unless the Hilbert space can be decomposed into multiple orthogonal subspaces that are all individually invariant under  $\mathcal{M}$  [26]. We will be interested in the *mixing time* of  $\mathcal{M}$ , defined as the minimal time needed to reach  $\rho_{\text{ss}}$ :

$$t_{\text{mix},\varepsilon} = \min \{t \geq 0 : \|\rho(t) - \rho_{\text{ss}}\|_1 \leq \varepsilon, \forall \rho(0)\}, \quad (1)$$

where  $\|\cdot\|_1$  is the trace norm and the cutoff  $\varepsilon$  is arbitrary and often taken to be  $\varepsilon = 1/4$ . When  $\mathcal{M}$  is engineered to approach a particular steady state, as in Gibbs samplers or other Monte-Carlo-style algorithms,  $t_{\text{mix}}$  characterizes the time needed for the algorithm to converge. Next, we show how to lower bound  $t_{\text{mix}}$  entirely in terms of the steady state itself.

*Quantum bottleneck theorem.* We will now formulate a quantum bottleneck theorem. Colloquially, it says the following (see Fig. 1). Imagine that the Hilbert space can be split into three regions  $\mathcal{A}, \mathcal{B}, \mathcal{C}$  in such a way that any quantum trajectory from  $\mathcal{A}$  to  $\mathcal{C}$  has to pass through  $\mathcal{B}$ . Then, if the steady state  $\rho_{\text{ss}}$  has small weight on  $\mathcal{B}$  com-

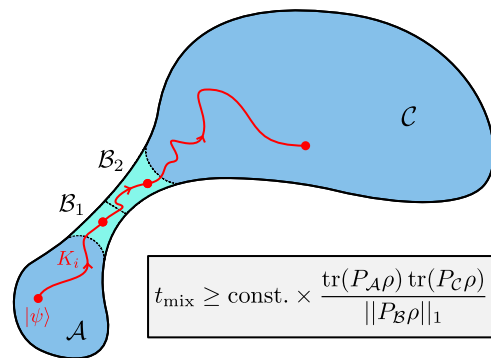


FIG. 1. **General bottleneck theorem.** The Hilbert space is split into four parts in such a way that Kraus operators only have matrix elements between neighboring subspaces. The probability of the “bottleneck region”  $\mathcal{B}$  can be used to upper bound the time it takes to cross from  $\mathcal{A}$  to  $\mathcal{C}$ .

pared to  $\mathcal{A}$ , then we can construct a density matrix supported entirely on  $\mathcal{A}$  that will remain localized in  $\mathcal{A}$  for long times. For those familiar with classical Markov chain bottlenecks, some details of our formulation might seem unusual; for this reason, we discuss the classical analogue in the supplement [27]. In particular, in order to make the proof work, we need to split the bottleneck region  $\mathcal{B}$  into two halves.

**Theorem 1** (Quantum bottleneck theorem). *Let  $\mathcal{M}$  be a quantum channel with steady state  $\rho$  and let  $\mathcal{H} = \mathcal{A} \oplus \mathcal{B}_1 \oplus \mathcal{B}_2 \oplus \mathcal{C}$  be a decomposition of the Hilbert space into orthogonal subspaces with projectors  $P_{\mathcal{A}}, P_{\mathcal{B}_1}, P_{\mathcal{B}_2}, P_{\mathcal{C}}$  and  $P_{\mathcal{B}_1} + P_{\mathcal{B}_2} \equiv P_{\mathcal{B}}$ . Assume that  $\mathcal{M}$  has a Kraus representation such that all Kraus operators  $K_i$  satisfy*

$$(P_{\mathcal{C}} + P_{\mathcal{B}_2})K_i P_{\mathcal{A}} = (P_{\mathcal{A}} + P_{\mathcal{B}_1})K_i P_{\mathcal{C}} = 0. \quad (2)$$

*Then the projected state,  $\rho_{\mathcal{A}} \equiv P_{\mathcal{A}}\rho P_{\mathcal{A}} / \text{tr}(P_{\mathcal{A}}\rho)$  satisfies*

$$\|\mathcal{M}[\rho_{\mathcal{A}}] - \rho_{\mathcal{A}}\|_1 \leq 10 \frac{\|P_{\mathcal{B}}\rho\|_1}{\text{tr}(P_{\mathcal{A}}\rho)} \equiv 10\Delta. \quad (3)$$

*Proof idea.* One first rewrites the LHS of Eq. (3) in terms of projections of  $\rho$  into the “complement”  $\mathcal{BC} \equiv \mathcal{B} \oplus \mathcal{C}$ . What happens to the state in  $\mathcal{BC}$  upon application of  $\mathcal{M}$ ? The part contained in  $\mathcal{B}_2\mathcal{C}$  cannot change, since it is far from the boundary between  $\mathcal{A}$  and  $\mathcal{B}$  and thus insensitive to the projection of  $\rho$  onto  $\mathcal{BC}$ . The part in  $\mathcal{B}_1$  can change; however, this change is bounded by the fact that it can only have weight flowing into it from  $\mathcal{B}_2$  which already has a small weight to begin with [27].

Note that while the denominator of the bottleneck ratio  $\Delta$  in Eq. (3) is simply the probability associated to the subspace  $\mathcal{A}$ , the numerator does not have a similar interpretation because the projector  $P_{\mathcal{B}}$  appears only on

one side of  $\rho$ . This means that if we decompose  $\rho$  into blocks according to  $\mathcal{A}, \mathcal{B}, \mathcal{C}$ , we require not only that the diagonal block in  $\mathcal{B}$  is small, but also that the off-diagonal blocks between  $\mathcal{B}$  and either  $\mathcal{A}$  or  $\mathcal{C}$  are small also (relative to the probability of  $\mathcal{A}$  itself). However, off-diagonal and diagonal matrix elements are related by the fact that  $\rho$  has to be a positive semi-definite matrix. Using this, we can upper bound the numerator solely in terms of probabilities using the following [27]

**Lemma 1.**  $\|P_{\mathcal{B}}\rho\|_1 \leq \sqrt{\text{tr}(P_{\mathcal{B}}\rho)}$ .

Note that this bound might be very loose. For example, if  $[P_{\mathcal{B}}, \rho] = 0$  then  $\|P_{\mathcal{B}}\rho\|_1 = \text{tr}(P_{\mathcal{B}}\rho)$ , without the square root. This happens e.g. in classical systems.

The bottleneck theorem leads directly to a bound on the mixing time. In particular, let  $\{\mathcal{M}_\tau\}$  be a collection of channels that all share the same steady state  $\rho$  and all satisfy Eq. (2) with respect to the same decomposition into  $\mathcal{A}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{C}$ . A simple argument [27] shows that

$$\begin{aligned} \|\mathcal{M}_t \dots \mathcal{M}_1[\rho_{\mathcal{A}}] - \rho_{\mathcal{A}}\|_1 &\leq t \min_{\tau} \|\mathcal{M}_\tau[\rho_{\mathcal{A}}] - \rho_{\mathcal{A}}\|_1 \\ &\leq 10t\Delta, \end{aligned} \quad (4)$$

which in turn implies a lower bound on the mixing time:

**Lemma 2.**  $t_{\text{mix}, \varepsilon} \geq \frac{1 - \text{tr}(P_{\mathcal{A}}\rho)}{5\Delta} - \varepsilon \geq \frac{\text{tr}(P_{\mathcal{A}}\rho)\text{tr}(P_{\mathcal{C}}\rho)}{5\|P_{\mathcal{B}}\rho\|_1} - \varepsilon$ .

An important aspect of Theorem 1 and the resulting bound in Lemma 2 is that it depends only on the properties of the steady state  $\rho$ . The properties of the channels  $\mathcal{M}_\tau$  appeared only through condition (2). In the following, we will consider a situation where this condition follows simply from  $k$ -locality and thus applies to a large family of channels.

*Bottlenecks from locality.* We now consider systems composed of  $n$  degrees of freedom, which we take to be qubits (generalizing to qudits is straightforward), establishing a decomposition of the  $2^n$ -dimensional Hilbert space  $\mathcal{H}$  into an  $n$ -fold tensor product. We will be interested in channels that exhibit a degree of locality with respect to this decomposition. In particular, let us call an operator  $r$ -local if it can be written as a linear combination of operators acting on at most  $r$  qubits. For channels, we say that  $\mathcal{M}$  is  $r$ -local if it has a Kraus representation where all Kraus operators are  $r$ -local.

Locality also allows us to define a notion of proximity of a state in  $\mathcal{H}$  to another state, or to some subspace of  $\mathcal{H}$ . In particular, we define the neighborhood and the boundary of a subspace as follows (see also Fig. 2):

**Definition 1.** Given a subspace  $\mathcal{V} \subset \mathcal{H}$  we define its  $r$ -neighborhood as

$$\mathcal{B}_r(\mathcal{V}) := \text{Span}\{\hat{O}|\psi\rangle \mid |\psi\rangle \in \mathcal{V}, \hat{O} \text{ is } r\text{-local}\}, \quad (5)$$

and its  $r$ -boundary  $\partial_r\mathcal{V}$  as the orthogonal complement of  $\mathcal{V}$  within  $\mathcal{B}_r(\mathcal{V})$ , satisfying  $\mathcal{B}_r(\mathcal{V}) = \mathcal{V} \oplus \partial_r\mathcal{V}$ .

The reason behind these definitions is that for any  $r$ -local channel, a quantum trajectory initialized in  $\mathcal{V}$  must pass through the boundary  $\partial_r\mathcal{V}$  before it can get to the rest of the Hilbert space. This means that we can apply Theorem 1 to this situation to obtain a general bound, applicable to any local channel [27]:

**Theorem 2** (Quantum Bottleneck Theorem, local channels). Let  $\rho$  be a density matrix,  $\mathcal{V} \subset \mathcal{H}$  a subspace with orthogonal projector  $P_{\mathcal{V}}$ , and  $\rho_{\mathcal{V}} = P_{\mathcal{V}}\rho P_{\mathcal{V}} / \text{tr}(P_{\mathcal{V}}\rho)$ . We define the bottleneck ratio as

$$\Delta(r) := \frac{\|P_{\partial_r\mathcal{V}}\rho\|_1}{\text{tr}(P_{\mathcal{V}}\rho)}. \quad (6)$$

Let  $\mathcal{M}$  be a quantum channel satisfying  $\mathcal{M}[\rho] = \rho$  and assume that  $\mathcal{M}$  is  $s$ -local with  $s \leq r$ . Then

$$\|\mathcal{M}[\rho_{\mathcal{V}}] - \rho_{\mathcal{V}}\|_1 \leq 10\Delta(r). \quad (7)$$

The implications of this theorem become more transparent when we consider a family of problems for different system sizes  $n$  with some well-defined thermodynamic limit  $n \rightarrow \infty$ . Assume that we can find a family of subspaces  $\mathcal{V}(n)$  and a diverging function  $r(n)$  such that the bottleneck ratio  $\Delta(n) \equiv \Delta(r(n))$  defined in Eq. (6) goes to zero superlogarithmically in the thermodynamic limit. Then Eq. (7) implies that for any  $s$ -local channel  $\mathcal{M}$  with finite  $s$  (or, more generally,  $s = o(r(n))$ ), the projected state  $\rho_{\mathcal{V}}$  becomes an increasingly good approximate steady state of  $\mathcal{M}$ . In particular, assuming that  $\lim_{n \rightarrow \infty} \text{tr}(P_{\mathcal{V}(n)}\rho_n) < 1$ , the mixing time is lower bounded to be  $t_{\text{mix}} = \Omega(1/\Delta(n))$ .

Theorem 2 generalizes to *quasi-local* channels, i.e. those that are well-approximated by local ones. In particular, we can approximate  $\mathcal{M}$  with one that is  $r(n)$ -local. As long as the approximation error decays faster than  $\Delta(n)$ ,  $1/\Delta(n)$  still lower bounds the mixing time [27].

*Gibbs sampling and low-temperature phases.* The bottleneck theorem has important implications for the the problem of *Gibbs sampling*, i.e. the case where the steady state is  $\rho = \rho_G \propto e^{-\beta H}$  for some (usually local) Hamiltonian  $H$ . The construction of general (quasi-)local Gibbs samplers is a very recent development [15, 16]. It has been shown that for local  $H$  at sufficiently small  $\beta$  they achieve *rapid mixing*,  $t_{\text{mix}} = O(\log n)$  [18]. Our results complement this by establishing sufficient conditions for *slow mixing*. In particular, we expect that Theorem 2 implies (super)polynomial mixing times at low temperatures in many cases of interest for reasons we now explain.

For Gibbs states, the bottleneck ratio is related to the notion of a *free energy barrier*, between  $\mathcal{V}$  and the rest of

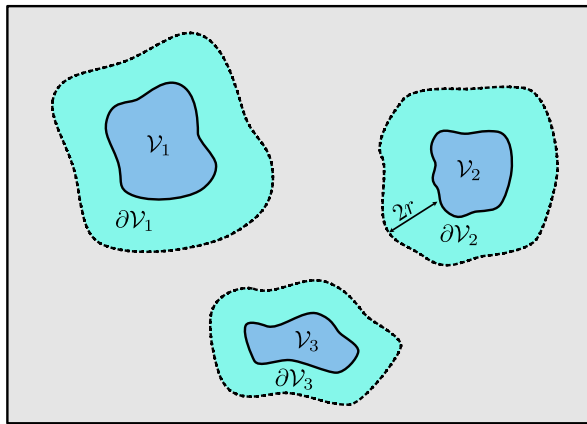


FIG. 2. **Local channels and Gibbs state structure.** For local channels, we define bottleneck regions in terms of states that can be reached by local operations. In non-trivial finite temperature phases, the Gibbs state  $\rho_G$  fragments into multiple components surrounded by large bottlenecks.

Hilbert space. Let us define the free energy of subspace  $\mathcal{V}$  as  $F(\mathcal{V}) = -\text{tr}(e^{-\beta H})/\beta$ . We can relate the bottleneck ratio Eq. (6) to the free energy in various cases<sup>2</sup>:

- (a) If  $\mathcal{V}$  has finite probability in the thermodynamic limit,  $\text{tr}(P_{\mathcal{V}}\rho_G) > c$ , then  $\Delta \leq e^{-\beta F(\partial\mathcal{V})/2}/c$ .
- (b) If  $[\rho_G, P_{\partial\mathcal{V}}] = 0$ , then  $\Delta \leq e^{-\beta(F(\partial\mathcal{V})-F(\mathcal{V}))}$ .
- (c) In the most general case, from Lemma 1,

$$\Delta \leq e^{-\frac{\beta}{2}((F(\partial\mathcal{V})-F(\mathcal{V}))+(F-E_{\min}(\mathcal{V})))}, \quad (8)$$

where  $F = F(\mathcal{H})$  is the free energy of the Gibbs state and  $E_{\min}(\mathcal{V}) \equiv \min_{|\psi\rangle \in \mathcal{V}}(\langle\psi|H|\psi\rangle)$  is the minimal energy in the subspace  $\mathcal{V}$ . As long as this minimal energy is smaller than  $F$ , this reduces to the previous bound with  $\beta \rightarrow \beta/2$ .

The upshot is that in many cases, either  $F(\partial\mathcal{V})$  or  $F(\partial\mathcal{V}) - F(\mathcal{V})$  diverging with  $n$  (a “macroscopic free energy barrier”) implies  $\Delta(n) \rightarrow 0$ . We will now discuss some cases where we expect this to happen.

One example is *spontaneous symmetry breaking* [28]. This has been established rigorously e.g. in the classical Ising model [19, 29] and is expected to hold for other Hamiltonians in the same phase, including non-commuting ones, and for other cases of discrete symmetry breaking. Analogous physical arguments suggest the existence of macroscopic free-energy barriers in the four dimensional toric code [14, 30, 31] (which spontaneously breaks so-called *higher form* symmetries [32]), and its relatives, although no proof exists to the best of our knowledge. These examples fit under case (a) above; a weaker

version (with  $c = 1/O(\text{poly}(n))$  rather than  $O(1)$ ) should apply to continuous symmetry breaking as well. Another context, fitting into case (b), where macroscopic barriers have been proven are *classical spin glasses* [33–36]. One again expects these results to generalize to related quantum models and below we use Eq. (8) to establish this for a particular family of models [27] (see also Ref. [37]).

This suggests a unifying picture of many non-trivial finite temperature phases of matter (see Fig. 2). In these phases, the Gibbs state fragments into multiple components separated by bottlenecks. Different phases are distinguished by the number and nature of these Gibbs state components. This allows to treat many different phases within the same unified framework and also connects it to the study of phases in open quantum systems [22], by suggesting that equilibrium phases can be understood in terms of the near-degenerate steady states that must be present for any local dynamics obeying detailed balance.

*Bottlenecks from extensive energy barriers.* Here, we discuss a concrete examples where Theorem 2 leads to exponential mixing times,  $t_{\text{mix}} = e^{\Omega(n)}$ .

As we discussed, the bottleneck ratio simplifies when  $\rho_G$  and  $P_{\partial\mathcal{V}}$  commute. A situation where this applies is the case of Hamiltonians  $H_0$ , which can be written as sums of commuting projectors. One often imposes a “low density” condition, meaning that each term acts on finitely many qubits and each qubit appears in finitely many terms (with “finitely” meaning  $o(1)$  in the limit  $n \rightarrow \infty$ ). While this condition implies a degree of locality, it nevertheless allows for  $H_0$  that are not geometrically local. An interesting family of such stabilizer models are those exhibiting *extensive energy barriers* [38–40]:

$$E_{\min}(\partial\mathcal{V}) - E_{\min}(\mathcal{V}) \geq \kappa n, \quad (9)$$

for some constant  $\kappa > 0$ . It is easy to see that this also leads to a bottleneck ratio vanishing as  $\Delta(n) = e^{-\Omega(n)}$ . In Ref. [37], it is shown that in some models, there exist exponentially many in  $n$  subspaces  $\mathcal{V}$  satisfying Eq. (9), including ones with  $E_{\min}(\mathcal{V}) = O(n)$ , which is interpreted in terms of spin glass order [33, 41] in these models.

A natural question to ask is about the robustness of this spin glass order to perturbations of the Hamiltonian  $H = H_0 + V$  where  $V$  is a sum of  $r$ -local terms. Here we use Theorem 2 to establish stability, proving [27]

**Theorem 3** (Stability of spin glass order). *Let  $\mathcal{V}$  be a subspace satisfying Eq. (9) for a commuting projector Hamiltonian  $H_0$ , with  $E_{\min}(\mathcal{V}) \leq \kappa n/2$ . Consider  $H = H_0 + V$  with  $\|V\| \leq gn$ . For  $\beta > \beta_*$  and  $g < g_*(\beta)$ , the bottleneck ratio of  $\Delta(\mathcal{V})$  in the state  $\rho_G \propto e^{-\beta H}$  remains exponentially small in system size,  $\Delta(\mathcal{V}) = e^{-\Omega(n)}$ .*

The key step in proving Theorem 3, beyond applying Eq. (9), is to show that eigenstates of  $H$  with energies

<sup>2</sup> In the following, we omit the  $r$ -dependence for simplicity.

close to  $E_{\min}(\mathcal{V})$  have exponentially small amplitudes in  $\partial\mathcal{V}$ . We establish this using recent results from Ref. [42].

*Conclusions.* In this paper, we proved a generalization of the bottleneck theorem of Markov chains for quantum channels, both in a general context and for the specific case of channels with local Kraus operators. We discussed the application of this result to Gibbs states of local Hamiltonians, and argued that in many physically relevant cases, it yields superpolynomial mixing times for any local Gibbs sampler, which also provides an interesting dynamical perspective on equilibrium phases of quantum matter. We further demonstrated the validity of this approach by using the bottleneck theorem to establish the stability of spin glass order in certain stabilizer models against small perturbations.

Our work opens up a number of interesting open questions. The most obvious one is to apply the bottleneck theorem to substantiate our discussion on finite temperature phases by proving large mixing times in models of interest, such as the four dimensional toric code. Another is to further explore the implications of this dynamical perspective on thermal phases, e.g. by applying the ideas developed in Ref. [22] to Gibbs samplers. Finally, we expect our bottleneck result to have applications for other Monte Carlo-style quantum algorithms beyond Gibbs sampling.

*Note added.* During the completion of this manuscript, a related work appeared, also developing quantum versions of the bottleneck theorem [43]. While their original version applies only to the case when the decomposition of the Hilbert space commutes with the steady state, and is thus less broadly applicable than ours, the updated version, appearing in the same arXiv posting as the present manuscript, involves a bottleneck theorem that should apply to similar cases as ours. The definition of the quantum bottleneck ratio is nevertheless different in the two cases and a more detailed comparison is an interesting problem for future work.

*Acknowledgements.* We thank Grace Sommers for collaboration on related work and we thank Ethan Lake, András Gilyén and especially Curt von Keyserlingk for useful discussions. T.R. was supported in part by Stanford Q-FARM Bloch Postdoctoral Fellowship in Quantum Science and Engineering and by the HUN-REN Welcome Home and Foreign Researcher Recruitment Programme 2023. B.P. acknowledges funding through a Leverhulme-Peierls Fellowship at the University of Oxford and the Alexander von Humboldt foundation through a Feodor-Lynen fellowship. This work was done in part while NPB was visiting the Simons Institute for the Theory of Computing, supported by DOE QSA grant #FP00010905. V.K. acknowledges support from the Packard Foundation through a Packard Fellowship in Science and Engineering and the US Department of Energy, Office of Science, Basic Energy Sciences, under Early Career Award Nos. DE-SC0021111.

- 
- [1] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, USA, 2002).
  - [2] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
  - [3] D. A. Lidar, Lecture notes on the theory of open quantum systems, arXiv preprint arXiv:1902.00967 (2019).
  - [4] M.-H. Yung and A. Aspuru-Guzik, A quantum–quantum metropolis algorithm, *Proceedings of the National Academy of Sciences* **109**, 754 (2012).
  - [5] D. Layden, G. Mazzola, R. V. Mishmash, M. Motta, P. Wocjan, J.-S. Kim, and S. Sheldon, Quantum-enhanced markov chain monte carlo, *Nature* **619**, 282 (2023).
  - [6] P. Wocjan and K. Temme, Szegedy walk unitaries for quantum maps, *Communications in Mathematical Physics* **402**, 3201 (2023).
  - [7] D. Poulin and P. Wocjan, Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer, *Physical review letters* **103**, 220502 (2009).
  - [8] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, Quantum metropolis sampling, *Nature* **471**, 87 (2011).
  - [9] P. Rall, C. Wang, and P. Wocjan, Thermal state preparation via rounding promises, *Quantum* **7**, 1132 (2023).
  - [10] D. Zhang, J. L. Bosse, and T. Cubitt, Dissipative quantum gibbs sampling, arXiv preprint arXiv:2304.04526 (2023).
  - [11] C.-F. Chen, M. J. Kastoryano, F. G. Brandão, and A. Gilyén, Quantum thermal state preparation, arXiv preprint arXiv:2303.18224 (2023).
  - [12] A. Kossakowski, A. Frigerio, V. Gorini, and M. Verri, Quantum detailed balance and kms condition, *Communications in Mathematical Physics* **57**, 97 (1977).
  - [13] F. Fagnola and V. Umanita, Generators of detailed balance quantum markov semigroups, *Infinite Dimensional Analysis, Quantum Probability and Related Topics* **10**, 335 (2007).
  - [14] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, On thermal stability of topological qubit in kitaev’s 4d model, *Open Systems & Information Dynamics* **17**, 1 (2010).
  - [15] C. Chen, M. Kastoryano, and A. Gilyén, An efficient and exact noncommutative quantum gibbs sampler (2023), arXiv preprint arXiv:2311.09207 (2023).
  - [16] A. Gilyén, C.-F. Chen, J. F. Doriguello, and M. J. Kastoryano, Quantum generalizations of glauber and metropolis dynamics, arXiv preprint arXiv:2405.20322 (2024).
  - [17] Z. Ding, B. Li, and L. Lin, Efficient quantum gibbs samplers with kubo–martin–schwinger detailed balance condition, arXiv preprint arXiv:2404.05998 (2024).
  - [18] C. Rouzé, D. S. França, and Á. M. Alhambra, Optimal quantum algorithm for gibbs state preparation, arXiv preprint arXiv:2411.04885 (2024).
  - [19] D. A. Levin and Y. Peres, *Markov chains and mixing times*, Vol. 107 (American Mathematical Soc., 2017).
  - [20] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani, Quantum walks on graphs, in *Proceedings of the thirty-third annual ACM symposium on Theory of computing* (2001) pp. 50–59.

- [21] M. Szegedy, Quantum speed-up of markov chain based algorithms, in *45th Annual IEEE symposium on foundations of computer science* (IEEE, 2004) pp. 32–41.
- [22] T. Rakovszky, S. Gopalakrishnan, and C. von Keyserlingk, Defining stable phases of open quantum systems, *Physical Review X* **14**, 041031 (2024).
- [23] H. Carmichael, *An open systems approach to quantum optics: lectures presented at the Université Libre de Bruxelles, October 28 to November 4, 1991*, Vol. 18 (Springer Science & Business Media, 2009).
- [24] K. Mølmer, Y. Castin, and J. Dalibard, Monte carlo wave-function method in quantum optics, *JOSA B* **10**, 524 (1993).
- [25] J. Watrous, *The theory of quantum information* (Cambridge university press, 2018).
- [26] D. Burgarth, G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa, Ergodic and mixing quantum channels in finite dimensions, *New Journal of Physics* **15**, 073045 (2013).
- [27] See Supplementary Material for proofs of Theorems 1-3 and various Lemmas and Corollaries.
- [28] A. Beekman, L. Rademaker, and J. Van Wezel, An introduction to spontaneous symmetry breaking, *SciPost Physics Lecture Notes*, 011 (2019).
- [29] L. E. Thomas, Bound on the mass gap for finite volume stochastic ising models at low temperature, *Communications in Mathematical Physics* **126**, 1 (1989).
- [30] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *Journal of Mathematical Physics* **43**, 4452 (2002).
- [31] M. B. Hastings, Topological order at nonzero temperature, *Physical review letters* **107**, 210501 (2011).
- [32] J. McGreevy, Generalized symmetries in condensed matter, *Annual Review of Condensed Matter Physics* **14**, 57 (2023).
- [33] M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, 2009).
- [34] G. B. Arous and A. Jagannath, Shattering versus metastability in spin glasses, *Communications on Pure and Applied Mathematics* **77**, 139 (2024).
- [35] A. E. Alaoui, A. Montanari, and M. Sellke, Shattering in pure spherical spin glasses, arXiv preprint arXiv:2307.04659 (2023).
- [36] D. Gamarnik, A. Jagannath, and E. C. Kızıldağ, Shattering in the ising pure  $p$ -spin model, arXiv preprint arXiv:2307.07461 (2023).
- [37] B. Placke, T. Rakovszky, N. Breuckmann, and V. Khemani, Topological quantum spin glass order and its realization in qldpc codes, (2024), in preparation.
- [38] M. Sipser and D. Spielman, Expander codes, *IEEE Transactions on Information Theory* **42**, 1710 (1996).
- [39] I. Dinur, M.-H. Hsieh, T.-C. Lin, and T. Vidick, Good quantum ldpc codes with linear time decoders, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023 (Association for Computing Machinery, New York, NY, USA, 2023) p. 905–918.
- [40] Y. Hong, J. Guo, and A. Lucas, Quantum memory at nonzero temperature in a thermodynamically trivial system, arXiv preprint arXiv:2403.10599 (2024).
- [41] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, Vol. 9 (World Scientific Publishing Company, 1987).
- [42] C. Yin, R. Nandkishore, and A. Lucas, Eigenstate localization in a many-body quantum system, arXiv preprint arXiv:2405.12279 (2024).
- [43] D. Gamarnik, B. T. Kiani, and A. Zlokapa, Slow mixing of quantum gibbs samplers, arXiv preprint arXiv:2411.04300 (2024).

## Supplementary Material for:

# Bottlenecks in quantum channels and defining finite temperature phases of matter

### SI. A CLASSICAL BOTTLENECK THEOREM

Here we present a classical analog of our Theorem 1, applicable to classical Markov chains. While bottleneck theorems in this case are well known [S1], our formulation differs from the standard one so we include it here to showcase the logic of the more general quantum proof in a simpler setting.

Let  $\Lambda$  denote the state space. Probability distributions  $|\pi\rangle$  over  $\Lambda$  correspond to  $|\Lambda|$ -dimensional non-negative vectors with  $\| |\pi\rangle \|_1 = \langle \Lambda | \pi \rangle = 1$ , where we use  $|\Lambda\rangle$  to denote the all-1 vector and  $\| \cdot \|_1$  is the vector 1-norm (sum of absolute values of all elements). We will use  $P_\Omega$  to denote the projector onto configurations in the set  $\Omega \subseteq \Lambda$ . The total probability weight associated to this set is  $\pi(\Omega) = \langle \Lambda | P_\Omega | \pi \rangle \equiv \langle \Omega | \pi \rangle$  where  $|\Omega\rangle \equiv P_\Omega |\Lambda\rangle$ .

We are now ready to state the analog of the bottleneck theorem.

**Theorem S I.1** (Classical Bottleneck Theorem). *Let  $M$  be a Markov generator (stochastic matrix) on  $\Lambda$ , with steady state  $|\pi\rangle$ , and let  $\Lambda = A \uplus B_1 \uplus B_2 \uplus C$  be a partition of the state space into a union of four disjoint sets, such that the following condition is satisfied:*

$$P_{B_2 \uplus C} M P_A = P_{A \uplus B_1} M P_C = 0 \quad (\text{S1})$$

Define the projected state  $|\pi_A\rangle \equiv P_A |\pi\rangle / \pi(A)$ . Then

$$\| M |\pi_A\rangle - |\pi_A\rangle \|_1 \leq 2 \frac{\pi(B)}{\pi(A)}, \quad (\text{S2})$$

where  $B \equiv B_1 \uplus B_2$ .

*Proof.* First, let us introduce some notation. We will denote by  $P_{AB_1} = P_{A \uplus B_1} = P_A + P_{B_1}$  and similarly for other combinations of the four subsets. We thus have  $P_{AB_1 B_2 C} = P_{ABC} = \mathbb{1}$ .

The conditions on  $M$  mean that in order to go from a configuration in  $A$  to a configuration in  $C$ , the system has to first go through  $B_1$  and then  $B_2$ , in that order, so that the sets are naturally arranged in a sequence (see Fig. 1 of the main text).

The claim of the theorem is equivalent to

$$\| M P_A |\pi\rangle - P_A |\pi\rangle \|_1 \leq 2 \| P_B |\pi\rangle \|_1. \quad (\text{S3})$$

To show this, we first note that

$$M P_A |\pi\rangle - P_A |\pi\rangle = M(\mathbb{1} - P_{BC}) |\pi\rangle - P_A |\pi\rangle = |\pi\rangle - P_A |\pi\rangle - M P_{BC} |\pi\rangle = -(M P_{BC} |\pi\rangle - P_{BC} |\pi\rangle). \quad (\text{S4})$$

We now multiply  $M P_{BC} |\pi\rangle$  by  $\mathbb{1} = P_{AB_1} + P_{B_2 C}$  from the left and note that, due to the conditions Eq. (S1), the following two equalities hold:

$$\begin{aligned} P_{B_2 C} M P_{BC} |\pi\rangle &= P_{B_2 C} M |\pi\rangle = P_{B_2 C} |\pi\rangle, \\ P_{AB_1} M P_{BC} |\pi\rangle &= P_{AB_1} M P_B |\pi\rangle. \end{aligned} \quad (\text{S5})$$

Importantly, the first term cancels part of  $P_{BC} |\pi\rangle = (P_{B_1} + P_{B_2 C}) |\pi\rangle$ . Combining these, we thus get

$$\| M P_{BC} |\pi\rangle - P_{BC} |\pi\rangle \|_1 = \| P_{AB_1} M P_B |\pi\rangle - P_{B_1} |\pi\rangle \|_1 \leq \| P_{AB_1} M P_B |\pi\rangle \|_1 + \| P_B |\pi\rangle \|_1 \leq 2 \| P_B |\pi\rangle \|_1, \quad (\text{S6})$$

where we used that both  $M$  and multiplying by projections cannot increase the 1-norm.  $\square$



## SII. QUANTUM BOTTLENECK THEOREMS

We consider quantum channels  $\mathcal{M}$  acting on a quantum system with Hilbert space  $\mathcal{H}$ . We denote by  $\mathcal{V} \subseteq \mathcal{H}$  a linear subspace of  $\mathcal{H}$  and by  $P_{\mathcal{V}}$  the orthogonal projector onto  $\mathcal{V}$ .

### A. General theorem

Let us now prove the quantum bottleneck theorem stated in the main text.

**Theorem S II.1** (Quantum Bottleneck Theorem, general case). *Let  $\mathcal{M}$  be a quantum channel with steady state  $\rho$  and. Let  $\mathcal{H} = \mathcal{A} \oplus \mathcal{B}_1 \oplus \mathcal{B}_2 \oplus \mathcal{C}$  be a decomposition of the Hilbert space into four orthogonal subspaces and let  $P_{\mathcal{A}}, P_{\mathcal{B}_1}, P_{\mathcal{B}_2}, P_{\mathcal{C}}$  denote the corresponding projectors. Assume that  $\mathcal{M}$  has a Kraus representation such that all Kraus operators  $K_i$  satisfy*

$$(P_{\mathcal{B}_2} + P_{\mathcal{C}})K_i P_{\mathcal{A}} = (P_{\mathcal{A}} + P_{\mathcal{B}_1})K_i P_{\mathcal{C}} = 0. \quad (\text{S7})$$

Then the projected state,  $\rho_{\mathcal{A}} \equiv P_{\mathcal{A}}\rho P_{\mathcal{A}}/\text{tr}(P_{\mathcal{A}}\rho)$  satisfies

$$\|\mathcal{M}[\rho_{\mathcal{A}}] - \rho_{\mathcal{A}}\|_1 \leq 10 \frac{\|P_{\mathcal{B}}\rho\|_1}{\text{tr}(P_{\mathcal{A}}\rho)}, \quad (\text{S8})$$

where  $P_{\mathcal{B}} = P_{\mathcal{B}_1} + P_{\mathcal{B}_2}$ .

*Proof.* First, we again introduce some shorthand notation:  $P_{\mathcal{AB}_1} = P_{\mathcal{A}} + P_{\mathcal{B}_1}$ ,  $P_{\mathcal{AB}} = P_{\mathcal{A}} + P_{\mathcal{B}_1} + P_{\mathcal{B}_2}$  etc. We begin by noting that, since  $P_{\mathcal{A}} = \mathbb{1} - P_{\mathcal{BC}}$  and  $\mathcal{M}[\rho] = \rho$ ,

$$\mathcal{M}[P_{\mathcal{A}}\rho P_{\mathcal{A}}] - P_{\mathcal{A}}\rho P_{\mathcal{A}} = \mathcal{M}[P_{\mathcal{BC}}\rho P_{\mathcal{BC}} - P_{\mathcal{BC}}\rho - \rho P_{\mathcal{BC}}] - P_{\mathcal{BC}}\rho P_{\mathcal{BC}} + P_{\mathcal{BC}}\rho + \rho P_{\mathcal{BC}}. \quad (\text{S9})$$

Consider the Kraus representation,  $\mathcal{M}[\rho] = \sum_i K_i \rho K_i^\dagger$ . From the conditions imposed on  $K_i$  we have that

$$\mathcal{M}[P_{\mathcal{BC}}\rho] = (P_{\mathcal{AB}_1} + P_{\mathcal{B}_2\mathcal{C}})\mathcal{M}[P_{\mathcal{BC}}\rho] = P_{\mathcal{AB}_1}\mathcal{M}[P_{\mathcal{B}}\rho] + P_{\mathcal{B}_2\mathcal{C}}\mathcal{M}[P_{\mathcal{ABC}}\rho] = P_{\mathcal{AB}_1}\mathcal{M}[P_{\mathcal{B}}\rho] + P_{\mathcal{B}_2\mathcal{C}}\rho. \quad (\text{S10})$$

The second of these terms will cancel with the second term of  $P_{\mathcal{BC}}\rho = (P_{\mathcal{B}_1} + P_{\mathcal{B}_2\mathcal{C}})\rho$ . Similarly, we have

$$\mathcal{M}[\rho P_{\mathcal{BC}}] = \mathcal{M}[\rho P_{\mathcal{B}}]P_{\mathcal{AB}_1} + \rho P_{\mathcal{B}_2\mathcal{C}}, \quad (\text{S11})$$

where the second term will now cancel the second term of  $\rho P_{\mathcal{BC}} = \rho P_{\mathcal{B}_1} + \rho P_{\mathcal{B}_2\mathcal{C}}$ . Finally, we have that

$$\mathcal{M}[P_{\mathcal{BC}}\rho P_{\mathcal{BC}}] = P_{\mathcal{B}_2\mathcal{C}}\rho P_{\mathcal{B}_2\mathcal{C}} + P_{\mathcal{AB}_1}\mathcal{M}[P_{\mathcal{B}}\rho P_{\mathcal{B}}]P_{\mathcal{AB}_1} + P_{\mathcal{AB}_1}\mathcal{M}[P_{\mathcal{B}}\rho]P_{\mathcal{B}_2\mathcal{C}} + P_{\mathcal{B}_2\mathcal{C}}\mathcal{M}[\rho P_{\mathcal{B}}]P_{\mathcal{AB}_1}, \quad (\text{S12})$$

which should be compared with

$$P_{\mathcal{BC}}\rho P_{\mathcal{BC}} = P_{\mathcal{B}_2\mathcal{C}}\rho P_{\mathcal{B}_2\mathcal{C}} + P_{\mathcal{B}_1}\rho P_{\mathcal{B}_1} + P_{\mathcal{B}_2\mathcal{C}}\rho P_{\mathcal{B}_1} + P_{\mathcal{B}_1}\rho P_{\mathcal{B}_2\mathcal{C}}. \quad (\text{S13})$$

Thus, combining all of the above, we find that Eq. (S9) becomes a sum of 16 terms, 6 of which cancel in pairs. The remaining terms all involve the projector  $P_{\mathcal{B}}$  multiplying  $\rho$  from either the left or the right (note that  $P_{\mathcal{B}_1}\rho = P_{\mathcal{B}_1}P_{\mathcal{B}}\rho$ ) and we can therefore upper bound them with  $\|P_{\mathcal{B}}\rho\|_1$ . This follows from the fact that both the channel  $\mathcal{M}$  and multiplication by a projector have the property that they do not increase the trace norm (in the latter case, we prove this fact below in Proposition S II.2). For example we have that

$$\|P_{\mathcal{AB}_1}\mathcal{M}[P_{\mathcal{B}}\rho]P_{\mathcal{B}_2\mathcal{C}}\|_1 \leq \|\mathcal{M}[P_{\mathcal{B}}\rho]\|_1 \leq \|P_{\mathcal{B}}\rho\|_1. \quad (\text{S14})$$

Thus, adding together all the remaining 10 terms, the theorem follows.  $\square$

**Proposition S II.2.** *Let  $O$  be an arbitrary operator and  $P$  be an orthogonal projector. Then  $\|PO\|_1 \leq \|O\|_1$  and similarly  $\|OP\|_1 \leq \|O\|_1$*

*Proof.* The statement follows from the duality property of Schatten norms [S2]:

$$\|O\|_1 = \max_{\|B\|=1} \{|\operatorname{tr}(O^\dagger B)|\}, \quad (\text{S15})$$

where  $\|\dots\|$  is the operator norm. It follows that

$$\begin{aligned} \|OP\|_1 &= \max_{\|B\|=1} \{|\operatorname{tr}(O^\dagger BP)|\} = \max_{\|B\|=1} \left\{ \|BP\| \cdot \left| \operatorname{tr} \left( O^\dagger \frac{BP}{\|BP\|} \right) \right| \right\} \leq \max_{\|B\|=1} \{\|BP\|\} \max_{\|B\|=1} \left\{ \left| \operatorname{tr} \left( O^\dagger \frac{BP}{\|BP\|} \right) \right| \right\} \\ &\leq \|O\|_1 \max_{\|B\|=1} \{\|BP\|\} \leq \|O\|_1 \cdot \|P\| \leq \|O\|_1, \end{aligned} \quad (\text{S16})$$

where we used that  $\|P\| \leq 1$  for the operator norm of a projector. The statement for  $\|PO\|_1$  follows similarly.  $\square$

If  $[P_{\mathcal{B}}, \rho] = 0$  then  $\|P_{\mathcal{B}}\rho\|_1 = \|P_{\mathcal{B}}\rho P_{\mathcal{B}}\|_1 = \operatorname{tr}(P_{\mathcal{B}}\rho)$ , so the expression in the upper bound becomes  $\frac{\operatorname{tr}(P_{\mathcal{B}}\rho)}{\operatorname{tr}(P_{\mathcal{A}}\rho)}$ , which is a ratio of probabilities associated to the two subspaces  $\mathcal{B}$  and  $\mathcal{A}$ . More generally, we can further upper bound the bottleneck ratio in a way that depends only on the probabilities of the two subspaces  $\mathcal{A}$  and  $\mathcal{B}$  evaluated in the steady state  $\rho$  using the following

**Lemma S II.3.** *For any density matrix  $\rho$  and projector  $P$ , the following inequality holds:  $\|\rho P\|_1 \leq \sqrt{\operatorname{tr}(\rho P)}$ .*

*Proof.* We first note that the positivity of  $\rho$  puts a constraint between diagonal and off-diagonal matrix elements. In particular, writing  $\rho = (\sqrt{\rho})^\dagger \sqrt{\rho}$  and using the Cauchy-Schwarz inequality yields the following

$$|\rho_{ij}| \leq \sqrt{\rho_{ii}\rho_{jj}}. \quad (\text{S17})$$

To turn this into a bound on the trace norm, we again use its dual representation. We split the Hilbert space up into  $\mathcal{H} = \mathcal{V} \oplus \mathcal{V}^\perp$  where  $\mathcal{V}$  is the subspace onto which  $P$  projects and take a basis  $\{|i\rangle\}$  that respects this decomposition (below, we write  $i \in \mathcal{V}$  to mean  $|i\rangle \in \mathcal{V}$ ). Then we have

$$\begin{aligned} \|\rho P\|_1 &= \max_{\|O\|=1} (|\operatorname{tr}(P\rho O)|) = \max_{\|O\|=1} (|\operatorname{tr}(P\rho OP)|) = \max_{\|O\|=1} \left( \left| \sum_{i \in \mathcal{V}} \sum_j \rho_{ij} O_{ji} \right| \right) \leq \max_{\|O\|=1} \left( \sum_{i \in \mathcal{V}} \sum_j |\rho_{ij}| |O_{ji}| \right) \\ &\leq \max_{\|O\|=1} \left( \sum_{i \in \mathcal{V}} \sum_j \sqrt{\rho_{ii}\rho_{jj}} |O_{ji}| \right) \equiv \max_{\|O\|=1} \left( \langle \sqrt{\rho_{\mathcal{V}}} | \tilde{O} | \sqrt{\rho} \rangle \right), \end{aligned} \quad (\text{S18})$$

where we defined  $\tilde{O}$  as the element-wise absolute value of  $O$ ,  $|\sqrt{\rho_{\mathcal{V}}}\rangle$  as a vector whose  $i$ -th element is  $\sqrt{\rho_{ii}}$  and  $|\sqrt{\rho_{\mathcal{V}}}\rangle$  as a vector whose  $j$ -th element is  $\sqrt{\rho_{jj}}$  if  $j \in \mathcal{V}$  and 0 otherwise. Then using Cauchy-Schwarz, the definition of the operator norm, and the fact that the element-wise absolute value does not decrease the operator norm, we get

$$\|\rho P\|_1 \leq \max_{\|O\|=1} \left( (\langle \sqrt{\rho_{\mathcal{V}}} | \sqrt{\rho_{\mathcal{V}}} \rangle)^{1/2} (\langle \sqrt{\rho} | \sqrt{\rho} \rangle)^{1/2} \|O\| \right) = (\langle \sqrt{\rho_{\mathcal{V}}} | \sqrt{\rho_{\mathcal{V}}} \rangle)^{1/2} = \sqrt{\sum_{j \in \mathcal{V}} \rho_{jj}}. \quad (\text{S19})$$

$\square$

We can easily generalize from a single time step to any evolution that is given by a product of channels. Let  $\{\mathcal{M}_\tau\}_{\tau=1}^t$  be an arbitrary collection of channels and  $\sigma$  an arbitrary quantum state. Define  $\Delta_t := \|\mathcal{M}_1 \dots \mathcal{M}_t[\sigma] - \sigma\|_1$  and  $\Delta_{\min} := \min_{1 \leq \tau \leq t} (\|\mathcal{M}_\tau[\sigma] - \sigma\|_1)$ . We then have that

$$\Delta_t = \|\mathcal{M}_1 \dots \mathcal{M}_{t-1}(\mathcal{M}_t[\sigma] - \sigma) + (\mathcal{M}_1 \dots \mathcal{M}_{t-1}[\sigma] - \sigma)\|_1 \leq \Delta_{\min} + \Delta_{t-1} \leq \dots \leq t\Delta_{\min}, \quad (\text{S20})$$

where we again used that channels are contracting with respect to the trace norm.

We can also use this to bound the distance from the steady state  $\rho$ :

$$\|\mathcal{M}_1 \dots \mathcal{M}_t[\sigma] - \rho\|_1 = \|(\mathcal{M}_1 \dots \mathcal{M}_t[\sigma] - \sigma) + (\sigma - \rho)\|_1 \geq \|\sigma - \rho\|_1 - \Delta_t \geq \|\sigma - \rho\|_1 - t\Delta_{\min}. \quad (\text{S21})$$

Thus the mixing time, defined as the smallest time needed for an arbitrary initial state to get within trace distance  $\varepsilon$  of  $\rho$ , obeys  $t_{\text{mix},\varepsilon} \geq \frac{\|\sigma - \rho\|_1}{\Delta_{\min}} - \varepsilon$ . Applying these considerations to Thm. SII.1 we get the following

**Corollary SII.4.** *Let  $\{\mathcal{M}_\tau\}_{\tau=1}^t$  be a collection of channels sharing the same steady state,  $\mathcal{M}_\tau[\rho] = \rho$  and each obeying the condition Eq. (S7) with respect to the same decomposition of the Hilbert space into  $\mathcal{A}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{C}$ . We then have*

$$t_{\text{mix},\varepsilon} \geq \frac{1 - \text{tr}(P_{\mathcal{A}}\rho)}{5\Delta} - \varepsilon, \quad \text{where } \Delta \equiv \frac{\|P_{\mathcal{B}}\rho\|_1}{\text{Tr}(P_{\mathcal{A}}\rho)}. \quad (\text{S22})$$

*Proof.* We take  $\sigma = \rho_{\mathcal{A}} = P_{\mathcal{A}}\rho P_{\mathcal{A}} / \text{tr}(P_{\mathcal{A}}\rho)$ . Applying Thm. SII.1, we then have that  $\Delta_{\min} = \min_{\tau} \|\mathcal{M}_\tau[\sigma] - \sigma\|_1 \leq 10\Delta$ . What remains is to bound  $\|\sigma - \rho\|_1$  which we do by using [S3]  $\|\sigma - \rho\|_1 = 2 \max_P \text{tr}(P(\sigma - \rho))$ , where the maximization goes over all projectors  $P$ . This gives

$$\|\rho_{\mathcal{A}} - \rho\|_1 = 2 \max_P \text{tr}(P(\rho_{\mathcal{A}} - \rho)) \geq 2 \left( \frac{\text{tr}(P_{\mathcal{A}}\rho P_{\mathcal{A}})}{\text{tr}(P_{\mathcal{A}}\rho)} - \text{tr}(P_{\mathcal{A}}\rho) \right) = 2(1 - \text{tr}(P_{\mathcal{A}}\rho)), \quad (\text{S23})$$

which completes the proof.  $\square$

## B. Local channels

We now specify to the case when  $\mathcal{H}$  is the Hilbert space of  $n$  qubits and consider channels that obey certain (mild) locality conditions. To define these, consider the set of all Pauli operators (products of Pauli matrices acting on different qubits), acting on at most  $r$  qubits, defined as

$$\mathcal{P}_r = \{S \mid S \in \mathcal{P}, |\text{supp}(S)| \leq r\} \subseteq \mathcal{P}, \quad (\text{S24})$$

with  $\text{supp}(S)$  being the support of  $S$ , i.e. the set of qubits it acts on non-trivially. Using this, we can now define neighborhoods and boundaries in Hilbert space as follows.

**Definition SII.5.** *The  $r$ -Hilbert space neighborhood of the subspace  $\mathcal{V}$  is defined as*

$$\mathcal{B}_r(\mathcal{V}) \equiv \text{span}\{S|\psi\rangle \mid |\psi\rangle \in \mathcal{V}, S \in \mathcal{P}_r\}. \quad (\text{S25})$$

*Note that  $\mathcal{B}_r(\mathcal{V})$  is itself a subspace containing  $\mathcal{V}$ ,  $\mathcal{V} \subseteq \mathcal{B}_r(\mathcal{V})$ , by construction. The  $r$ -boundary of  $\mathcal{V}$  is defined as the orthogonal complement of  $\mathcal{V}$  within  $\mathcal{B}_r(\mathcal{V})$ . In terms of the corresponding projectors:  $P_{\partial_r\mathcal{V}} = P_{\mathcal{B}_r(\mathcal{V})}(\mathbb{1} - P_{\mathcal{V}})$ .*

**Remark SII.6.** *Since  $\mathcal{V}$  is contained within  $\mathcal{B}_r(\mathcal{V})$ , their projectors commute and  $P_{\mathcal{B}_r(\mathcal{V})} = P_{\mathcal{V}} + P_{\partial_r\mathcal{V}}$ .*

The subspace  $\mathcal{B}_r(\mathcal{V})$  can be thought of as the states ‘‘within distance  $r$ ’’ of  $\mathcal{V}$ . This interpretation is facilitated by the following:

**Proposition SII.7.** *For any subspace  $\mathcal{V}$  and  $r, s \geq 0$ :*

$$\mathcal{B}_r(\mathcal{B}_s(\mathcal{V})) = \mathcal{B}_{r+s}(\mathcal{V}). \quad (\text{S26})$$

*Proof.* First, we will show that  $\mathcal{B}_r(\mathcal{B}_s(\mathcal{V})) \subseteq \mathcal{B}_{r+s}(\mathcal{V})$ . A generic element of  $\mathcal{B}_s(\mathcal{V})$  takes the form  $|\psi'\rangle = \sum_{S' \in \mathcal{P}_s} c'_{S'} S' |\psi_{S'}\rangle$  where  $|\psi_{S'}\rangle \in \mathcal{V}$ . We then have that a generic element of  $\mathcal{B}_r(\mathcal{B}_s(\mathcal{V}))$  can be written as

$$|\psi''\rangle = \sum_{S'' \in \mathcal{P}_r} c''_{S''} S'' |\psi'_{S''}\rangle = \sum_{S'' \in \mathcal{P}_r} \sum_{S' \in \mathcal{P}_s} c''_{S''} c'_{S'} S'' S' |\psi_{S',S''}\rangle. \quad (\text{S27})$$

Here  $|\psi_{S',S''}\rangle \in \mathcal{V}$  and since  $S'' S' \in \mathcal{P}_{r+s}$ , we have that  $|\psi''\rangle \in \mathcal{B}_{r+s}(\mathcal{V})$ .

To see containment in the other direction,  $\mathcal{B}_{r+s}(\mathcal{V}) \subseteq \mathcal{B}_r(\mathcal{B}_s(\mathcal{V}))$ , write a generic element of  $\mathcal{B}_{r+s}(\mathcal{V})$  as  $|\psi\rangle = \sum_{S \in \mathcal{P}_{r+s}} c_S |\psi_S\rangle$ . If  $S \in \mathcal{P}_{r+s}$ , then it can be split (in a non-unique way) as  $S = S'' S'$ , where  $S' \in \mathcal{P}_s$  and  $S'' \in \mathcal{P}_r$ . Choose such a splitting for every  $S$ ; then  $S' |\psi_{S=S',S''}\rangle \in \mathcal{B}_s(\mathcal{V})$  and consequently  $S |\psi_S\rangle = S'' S' |\psi_S\rangle \in \mathcal{B}_r(\mathcal{B}_s(\mathcal{V}))$ .  $\square$

**Definition S II.8.** Let  $\mathcal{M}$  be a quantum channel. We say that  $\mathcal{M}$  is *r-local* if it has a Kraus representation,  $\mathcal{M}[\rho] = \sum_i K_i \rho K_i^\dagger$ , such that all the Kraus operators  $K_i$  act on at most  $r$  qubits.

**Remark S II.9.** If  $K$  acts on at most  $r$  qubits, it can be written as  $K = \sum_{S \in \mathcal{P}_r} a_S S$  for some set of coefficients  $a_S$ .

**Theorem S II.10** (Quantum Bottleneck Theorem, local channels). Let  $\mathcal{M}$  be an *r-local* quantum channel with steady state  $\rho$ . Let  $\mathcal{V}$  be a subspace and define the quantum bottleneck ratio as

$$\frac{\|P_{\partial_{2r}\mathcal{V}}\rho\|_1}{\text{tr}(P_{\mathcal{V}}\rho)} = \frac{\|\rho P_{\partial_{2r}\mathcal{V}}\|_1}{\text{tr}(P_{\mathcal{V}}\rho)} = \Delta. \quad (\text{S28})$$

Define the projected state  $\rho_{\mathcal{V}} = P_{\mathcal{V}}\rho P_{\mathcal{V}} / \text{tr}(P_{\mathcal{V}}\rho)$ . Then  $\rho_{\mathcal{V}}$  is an approximate steady state of  $\mathcal{M}$  in the following sense:

$$\|\mathcal{M}[\rho_{\mathcal{V}}] - \rho_{\mathcal{V}}\|_1 \leq 10\Delta. \quad (\text{S29})$$

*Proof.* The claim follows directly from Thm. S II.1 with an appropriate identification of subspaces. In particular, we split the  $2r$ -neighborhood of  $\mathcal{V}$  region into two halves by using  $\mathcal{B}_{2r}(\mathcal{V}) = \mathcal{B}_r(\mathcal{B}_r(\mathcal{V}))$  and denote by  $P_{\mathcal{A}}$ ,  $P_{\mathcal{A}\mathcal{B}_1}$ ,  $P_{\mathcal{A}\mathcal{B}_1\mathcal{B}_2} = P_{\mathcal{A}\mathcal{B}}$  the projectors onto  $\mathcal{V}$ ,  $\mathcal{B}_r(\mathcal{V})$  and  $\mathcal{B}_{2r}(\mathcal{V})$  respectively. This induces a decomposition of the Hilbert space into four orthogonal subspaces  $\mathbb{1} = P_{\mathcal{A}} + P_{\mathcal{B}_1} + P_{\mathcal{B}_2} + P_{\mathcal{C}}$ , where  $P_{\mathcal{B}_1} = P_{\mathcal{A}\mathcal{B}_1}(\mathbb{1} - P_{\mathcal{A}})$ ,  $P_{\mathcal{B}_2} = P_{\mathcal{A}\mathcal{B}_1\mathcal{B}_2}(\mathbb{1} - P_{\mathcal{A}\mathcal{B}_1})$  and  $P_{\mathcal{C}} = \mathbb{1} - P_{\mathcal{A}\mathcal{B}_1\mathcal{B}_2}$ .

Consider now the Kraus representation of  $\mathcal{M}$  with Kraus operators  $i$  that are assumed to be *r-local* operators and let  $|\psi\rangle \in \mathcal{V}$ . Then we have that  $K_i|\psi\rangle = \sum_{S \in \mathcal{P}_r} a_S^{(i)} S|\psi\rangle \in \mathcal{B}_r(\mathcal{V})$ . This implies  $P_{\mathcal{B}_2\mathcal{C}}K_i P_{\mathcal{A}} = 0$ . A similar argument also yields  $P_{\mathcal{A}\mathcal{B}_1}K_i P_{\mathcal{C}} = 0$ . Thus  $\mathcal{M}$  fulfills the conditions of Thm. S II.1 and the claim follows.  $\square$

We can generalize the theorem to a more general set of channels, which fulfill a less stringent locality condition.

**Definition S II.11.** We say that the channel  $\mathcal{M}$  is *f-quasi-local*, for a decaying function  $\lim_{r \rightarrow \infty} f(r) = 0$ , if  $\forall r \geq 1$  there exists an *r-local* channel  $\mathcal{M}^{(r)}$  such that  $\|\mathcal{M} - \mathcal{M}^{(r)}\|_{\diamond} \leq f(r)$ .

For these, we can prove a different bottleneck theorem:

**Theorem S II.12** (Quantum Bottleneck Theorem, quasi-local channels). Let  $\mathcal{M}$  be an *f-quasi-local* channel with steady state  $\rho$ . Let  $\mathcal{V}$  be a subspace and  $\rho_{\mathcal{V}} = P_{\mathcal{V}}\rho P_{\mathcal{V}} / \text{tr}(P_{\mathcal{V}}\rho)$  be the steady state projected onto this subspace. Define

$$\Delta(s) \equiv \frac{\|P_{\partial_{2s}\mathcal{V}}\rho\|_1}{\text{tr}(P_{\mathcal{V}}\rho)} \quad (\text{S30})$$

We then have

$$\|\mathcal{M}[\rho_{\mathcal{V}}] - \rho_{\mathcal{V}}\|_1 \leq \inf_{1 \leq s \leq n/2} (10\Delta(s) + f(s)). \quad (\text{S31})$$

*Proof.* Fix  $s \geq 1$  and write  $\mathcal{M} = \mathcal{M}^{(s)} + \delta\mathcal{M}^{(s)}$ , where  $\mathcal{M}^{(s)}$  is the closest *s-local* approximation of  $\mathcal{M}$ . We then have

$$\|\mathcal{M}[\rho_{\mathcal{V}}] - \rho_{\mathcal{V}}\|_1 = \|\mathcal{M}^{(s)}[\rho_{\mathcal{V}}] - \rho_{\mathcal{V}} + \delta\mathcal{M}^{(s)}[\rho_{\mathcal{V}}]\|_1 \leq \|\mathcal{M}^{(s)}[\rho_{\mathcal{V}}] - \rho_{\mathcal{V}}\|_1 + \|\delta\mathcal{M}^{(s)}[\rho_{\mathcal{V}}]\|_1 \leq 10\Delta(s) + f(s), \quad (\text{S32})$$

where we used Theorem S II.1 to bound the first term and we used that  $\|\mathcal{M}[\sigma]\|_1 \leq \|\mathcal{M}\|_{\diamond} \|\sigma\|_1$  for any super-operator  $\mathcal{M}$ . Since the upper bound applies for any choice of  $s$ , the statement follows.  $\square$

### III. STABILITY OF COMMUTING PROJECTOR HAMILTONIANS WITH EXTENSIVE ENERGY BARRIERS

We consider Hamiltonians of the form  $H = H_0 + V$  on  $n$  qubits. The unperturbed Hamiltonian,  $H_0 = \sum_k \Pi_k$ , is a sum of mutually commuting projectors  $\Pi_k$  and we assume both that each projector acts on finitely many qubits and each qubit appears in finitely many terms of  $H_0$ .  $V$  is a perturbation that is a sum of local terms and whose operator norm satisfies  $\|V\| \leq gn$  for some constant perturbation strength  $g$ .

We will assume that  $H_0$  exhibits linear energy barriers around various local energy minima. In particular, we consider the setup of Thm. SII.10 above. Let  $\mathcal{V}$  be a subspace of  $\mathcal{H}$  and  $\partial_{r(n)}\mathcal{V}$  its boundary for some function  $r(n)$  diverging super-logarithmically in  $n$ ; in the following we simplify notation and refer to this boundary subspace simply as  $\partial\mathcal{V}$ . Let  $E_{0,\min}(\mathcal{V}) \equiv \min_{|\psi\rangle \in \mathcal{V}} (\langle \psi | H_0 | \psi \rangle)$  denote the minimal energy in  $\mathcal{V}$ . We then say that  $H_0$  has an extensive energy barrier surrounding  $\mathcal{V}$  if

$$E_{0,\min}(\partial\mathcal{V}) - E_{0,\min}(\mathcal{V}) \geq \kappa n \quad (\text{S33})$$

for some constant  $\kappa > 0$ . In the following, we will write  $E_{0,\min}(\mathcal{V}) = \epsilon n$  where  $\epsilon$  defines an energy density associated to the subspace  $\mathcal{V}$  (which might be zero). Our goal is to argue that whenever Eq. (S33) holds, and  $\epsilon$  is sufficiently small, the bottleneck ratio  $\Delta$  is exponentially small in  $n$  for sufficiently large  $\beta$  and small perturbation strength  $g$ .

Before discussing the proof, let us review some examples that satisfy this property, originating from classical and quantum error correcting codes. In these cases, a property of codes called *expansion* guarantees the existence of extensive energy barriers for various choices of subspaces  $\mathcal{V}$ , including ones at finite energy density, with  $\epsilon > 0$ . Indeed, for certain codes, the number of distinct subspaces having such barriers is exponentially large in  $n$  and increases with  $\epsilon$ ; see Ref. [S4].

**Example SIII.1** (Classical expander codes). *We can construct Hamiltonians out of classical error correcting codes by taking  $\Pi_k = (\mathbb{1} - C_k)/2$  where  $C_k = \prod_{i \in \delta(k)} Z_i$  correspond to the parity check of a classical code (with  $Z_i$  the diagonal Pauli matrix acting on qubit  $i$ ), defined by its support  $\delta(k)$  (see e.g. [S5] for more details). Eigenstates of  $H_0$  are diagonal in the computational basis and their energy is given by the number of violated checks. We can represent each configuration by a binary vector  $\mathbf{x}$  and denote its energy by  $E_0(\mathbf{x})$ .*

*There exist examples, including various constructions of good classical codes [S6] that satisfy  $(\delta, \gamma)$ -expansion*

$$|\mathbf{x}| \leq \delta n \Rightarrow E_0(\mathbf{x}) \geq \gamma |\mathbf{x}|, \quad (\text{S34})$$

*meaning that the energy grows linearly with the Hamming weight  $|\mathbf{x}|$ . This already implies an extensive energy barrier around any of the ground states of  $H_0$  (the code's codewords). We can take  $\mathbf{x}_0$  to be such a codeword, satisfying  $E_0(\mathbf{x}_0) = 0$  and choose  $\mathcal{V}$  to be the subspace spanned by configurations that are within Hamming distance  $\leq \delta n/2$  of  $\mathbf{x}_0$ . If we then take  $r(n) \leq \delta n/2$  in the definition of  $\partial\mathcal{V}$ , it will satisfy Eq. (S33) with  $\kappa = \gamma\delta/2$ .*

*In fact, expansion also implies the existence of extensive barriers surrounding other low-energy configurations. It follows from the triangle inequality for the Hamming weight that [S4]*

$$|\mathbf{x} + \mathbf{x}_0| \leq \delta n \Rightarrow E_0(\mathbf{x}) - E_0(\mathbf{x}_0) \geq \gamma |\mathbf{x} + \mathbf{x}_0| - 2E_0(\mathbf{x}_0). \quad (\text{S35})$$

*Therefore, when  $E_0(\mathbf{x}_0)$  is sufficiently small,  $\mathbf{x}_0$  is still surrounded by an extensive energy barrier. For example, if  $E_0(\mathbf{x}_0) \leq \gamma\delta n/4$ , we can take  $\mathcal{V}$  to be a subspace defined by configurations within Hamming distance  $\frac{3}{4}\delta n$  and  $r(n) \leq \delta n/4$  to get an extensive barrier with  $\kappa = \delta\gamma/4$ .*

**Example SIII.2** (Good quantum CSS codes). *Similar expansion properties are known to hold for certain quantum error correcting codes. In these cases, there are two sets of parity checks,  $B_k = \prod_{i \in \delta_Z(k)} Z_i$  and  $A_{k'} = \prod_{i \in \delta_X(k')} X_i$ , all mutually commuting, and we can combine them into a stabilizer Hamiltonian  $H_0 = \sum_k \frac{\mathbb{1} + B_k}{2} + \sum_{k'} \frac{\mathbb{1} + A_{k'}}{2}$  whose ground states comprise the codestates of the corresponding quantum code. Other eigenstates can be written as*

$$|\mathbf{x}, \mathbf{z}\rangle \equiv \left( \prod_{i: x_i=1} X_i \right) \left( \prod_{i': z_{i'}=1} Z_{i'} \right) |\psi_0\rangle, \quad (\text{S36})$$

where  $|\psi_0\rangle$  is a groundstate of  $H_0$  and  $\mathbf{x}, \mathbf{z}$  are a pair of binary vectors. Energy with respect to  $H_0$  is given by the number of checks of each type that is violated in such a state and can be written as  $E_0(\mathbf{x}, \mathbf{z}) = E_0^Z(\mathbf{x}) + E_0^X(\mathbf{z})$ .

Importantly, the labeling of eigenstates by  $\mathbf{x}, \mathbf{z}$  is not unique: the operator  $\prod_{i:x_i=1} X_i$  can be deformed by multiplying it with the checks  $A_i$  without changing its effect on  $|\psi_0\rangle$  and similarly for the  $Z$ -type operators. One thus defines a ‘‘reduced weight’’  $|\mathbf{x}|_{\text{red}}$  as the minimum of the Hamming weight  $|\mathbf{x}|$  over all equivalent choices of  $\mathbf{x}$ . Code expansion is then defined in terms of this reduced weight as [S7, S8]

$$\begin{aligned} |\mathbf{x}|_{\text{red}} \leq \delta n &\Rightarrow E_0^Z(\mathbf{x}) \geq \gamma |\mathbf{x}|_{\text{red}}, \\ |\mathbf{z}|_{\text{red}} \leq \delta n &\Rightarrow E_0^X(\mathbf{z}) \geq \gamma |\mathbf{z}|_{\text{red}}. \end{aligned} \quad (\text{S37})$$

The discussion from classical codes then carries over using the modified definition of distance. I.e., we can fix a ground state  $|\psi_0\rangle$  and define the subspace  $\mathcal{V}$  around it to be spanned by all eigenstates  $|\mathbf{x}, \mathbf{z}\rangle$  of the form Eq. (S36) with  $|\mathbf{x}|_{\text{red}}, |\mathbf{z}|_{\text{red}} \leq \delta n/2$ . Since any operator can be written as a linear combination of products of Pauli  $X$  and  $Z$  operators, multiplying a vector in  $\mathcal{V}$  with an  $r$ -local operator results in a state that is a superposition of eigenstates with  $|\mathbf{x}|_{\text{red}}, |\mathbf{z}|_{\text{red}} \leq \delta n/2 + r$ . Therefore if we take  $r(n) \leq \delta n/2$  in the definition of  $\partial\mathcal{V}$ , we will satisfy Eq. (S33) with  $\kappa = \gamma\delta$ .

Finally, using the triangle inequality, now for  $|\bullet|_{\text{red}}$ , we get [S4]

$$\begin{aligned} |\mathbf{x} + \mathbf{x}_0|_{\text{red}} \leq \delta n &\Rightarrow E_0^Z(\mathbf{x}) - E_0^Z(\mathbf{x}_0) \geq \gamma |\mathbf{x} + \mathbf{x}_0|_{\text{red}} - 2E_0^Z(\mathbf{x}_0), \\ |\mathbf{z} + \mathbf{z}_0|_{\text{red}} \leq \delta n &\Rightarrow E_0^X(\mathbf{z}) - E_0^X(\mathbf{z}_0) \geq \gamma |\mathbf{z} + \mathbf{z}_0|_{\text{red}} - 2E_0^X(\mathbf{z}_0). \end{aligned} \quad (\text{S38})$$

We can therefore take some reference state  $|\mathbf{x}_0, \mathbf{z}_0\rangle$  with energy  $E_0(\mathbf{x}_0, \mathbf{z}_0) \leq \gamma\delta n/4$  and take  $\mathcal{V}$  to be defined by eigenstates of the form  $|\mathbf{x}_0 + \mathbf{x}, \mathbf{z}_0 + \mathbf{z}\rangle$  with  $|\mathbf{x}|_{\text{red}}, |\mathbf{z}|_{\text{red}} \leq 3\delta n/4$  and choose  $r(n) \leq \delta n/4$  to achieve Eq. (S33) with  $\kappa = \delta\gamma/2$ .

We now want to prove that extensive barriers with respect to  $H_0$  imply an exponentially small bottleneck ratio even in the perturbed model  $H_0 + V$ , assuming the perturbation is weak and the temperature is low.

To bound the bottleneck ratio, we will make use of Lemma SII.3 to write

$$\Delta \leq \frac{\sqrt{\text{tr}(\rho_G P_{\partial\mathcal{V}})}}{\text{tr}(\rho_G P_{\mathcal{V}})} = Z^{1/2} \frac{\sqrt{\text{tr}(e^{-\beta H} P_{\partial\mathcal{V}})}}{\text{tr}(e^{-\beta H} P_{\mathcal{V}})}, \quad (\text{S39})$$

where  $Z = \text{tr}(e^{-\beta H})$  is the partition function. The denominator is easy to bound:

$$\text{tr}(e^{-\beta H} P_{\mathcal{V}}) \geq \max_{|\psi\rangle \in \mathcal{V}} \langle \psi | e^{-\beta H} | \psi \rangle \geq e^{-\beta \min_{|\psi\rangle \in \mathcal{V}} (\langle \psi | H | \psi \rangle)} \geq e^{-\beta(\epsilon+g)n}, \quad (\text{S40})$$

where we used the convexity of the exponential function and the fact that the perturbation  $V$  changes the energy of any state by at most  $gn$ .

The more difficult part is upper bounding  $\text{tr}(e^{-\beta H} P_{\partial\mathcal{V}})$ . The problem here is that, even though we know that the average energy of any state in  $\partial\mathcal{V}$  is large (i.e., larger than  $(\epsilon + \kappa - g)n$ , thanks to Eq. (S33)), this in itself does not give any useful bound on the expectation value of  $e^{-\beta H}$  in this state, which could be dominated by contributions from the small components of the wavefunction on eigenstates with energies much smaller than the average. Luckily, we can adapt a proof from Ref. [S9] to bound these tails.

**Lemma SIII.3.** *Let  $|\psi\rangle$  be an eigenstate of  $H = H_0 + V$  with energy  $E < \epsilon_1 n$  and let  $Q_{>}$  be a projector onto the subspace  $\mathcal{H}_{>}$ , spanned by eigenstates of  $H_0$  with energies  $\geq \epsilon_2 n$  with  $\epsilon_2 > \epsilon_1 + 4g$ . Let each qubit be part of at most  $w_0$  terms in  $H_0$  and each term in  $V$  act on at most  $w_1$  qubits. Then*

$$\|Q_{>} |\psi\rangle\| \leq e^{-\lambda(g)n}, \quad \text{with } \lambda(g) \geq \frac{\epsilon_2 - \epsilon_1}{2w_0 w_1} \ln \left( \frac{\epsilon_2 - \epsilon_1}{2g} \right). \quad (\text{S41})$$

*Proof.* Let  $|\psi_0\rangle$  be an eigenstate of  $H_0$  with energy  $E_0$ . Such an eigenstate is characterized by some pattern of excitations, i.e. the set of projectors  $\Pi_k$  that have eigenvalue 1 in  $|\psi_0\rangle$ . Write  $V = \sum_l V_l$  where the terms  $V_l$  act on at most  $w_1$  qubits. Since each qubit is part of at most  $w_0$  terms in  $H_0$ , therefore the state  $V|\psi_0\rangle$  is contained within the subspace spanned by the eigenstates of  $H_0$  with energies in the interval  $[E_0 - w_0w_1, E_0 + w_0w_1]$ . The perturbation  $V$  thus only connects eigenstates of  $H_0$  with nearby energy.

We now divide the Hilbert space into a sum of orthogonal subspaces,

$$\mathcal{H} = \mathcal{H}_> \oplus \bigoplus_{q=1}^{q_*} \mathcal{H}_q \oplus \mathcal{H}_<, \quad (\text{S42})$$

where the different subspaces are defined as follows

- $\mathcal{H}_>$  is spanned by eigenstates of  $H_0$  with eigenvalues  $E_0 \geq \epsilon_2 n$
- $\mathcal{H}_<$  is spanned by eigenstates of  $H_0$  with eigenvalues  $E_0 < (\frac{\epsilon_2 - \epsilon_1}{2} + 2g)n$ .
- $\mathcal{H}_q$  is spanned by eigenstates of  $H_0$  with eigenvalues  $E_0 \in [E(q), E(q) + \Delta E]$ , where  $\Delta E > w_0w_1$  is an  $O(1)$  constant and  $E(q) = E(1) + (q-1)\Delta E$ , with  $E(1) = \frac{\epsilon_2 + \epsilon_1}{2}n + 2gn$  and  $E(q_*) + \Delta E = \epsilon_2 n$ . Note that  $q_*\Delta E = (\frac{\epsilon_2 - \epsilon_1}{2} - 2g)n$ .

In this decomposition,  $H_0$  is block-diagonal and  $V$  is block-tri-diagonal, only connecting neighboring subspaces, so we can write

$$H_0 = \begin{pmatrix} h_> & 0 & 0 & 0 & \dots \\ 0 & h_{q_*} & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \dots & 0 & 0 & h_1 & 0 \\ \dots & 0 & 0 & 0 & h_< \end{pmatrix}; \quad V = \begin{pmatrix} v_> & v_{>,q_*} & 0 & 0 & \dots \\ v_{>,q_*}^\dagger & v_{q_*} & v_{q_*,q_*-1} & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \dots & 0 & v_{2,1}^\dagger & v_1 & v_{1,<} \\ \dots & 0 & 0 & v_{1,<}^\dagger & v_< \end{pmatrix}. \quad (\text{S43})$$

Now consider an eigenstate  $|\psi\rangle$  of  $H = H_0 + V$  with energy  $E$ , which we decompose according to Eq. (S42) as

$$|\psi\rangle = c_> |\psi_>\rangle + \sum_q c_q |\psi_q\rangle + c_< |\psi_<\rangle, \quad (\text{S44})$$

where the states  $|\psi_q\rangle$ ,  $|\psi_>\rangle$  and  $|\psi_<\rangle$  are normalized and the coefficients  $c_q$ ,  $c_>$  and  $c_<$  can be chosen to be real and non-negative. Using Eq. (S43), the eigenvalue equation implies

$$Ec_> |\psi_>\rangle = c_> (h_> + v_>) |\psi_>\rangle + c_{q_*} v_{>,q_*} |\psi_{q_*}\rangle. \quad (\text{S45})$$

Multiplying by  $\langle\psi_>|$  and using  $\|V\| \leq gn$ , along with the definition of  $\mathcal{H}_>$ , we find

$$c_{q_*} gn \geq -c_{q_*} \langle\psi_>|v_{>,q_*}|\psi_>\rangle = c_> (\langle\psi_>|h_>|\psi_>\rangle + \langle\psi_>|v_>|\psi_>\rangle - E) \geq c_> (\epsilon_2 n - gn - E). \quad (\text{S46})$$

Rearranging this we get that  $c_>$  is upper bounded by  $c_{q_*}$  as

$$c_> \leq \frac{gn}{\epsilon_2 n - gn - E} c_{q_*} \leq c_{q_*}. \quad (\text{S47})$$

We can now iterate this bound to get an increasingly strong upper bound on  $c_>$ . In particular, for any  $1 \leq q \leq q_*$  the eigenvalue equation yields (identifying  $\mathcal{H}_>$  with  $\mathcal{H}_{q_*+1}$  and  $\mathcal{H}_<$  with  $\mathcal{H}_0$ ):

$$Ec_q |\psi_q\rangle = c_q (h_q + v_q) |\psi_q\rangle + c_{q+1} v_{q+1,q}^\dagger |\psi_{q+1}\rangle + c_{q-1} v_{q,q-1}^\dagger |\psi_{q-1}\rangle \quad (\text{S48})$$

Let us now assume that  $c_{q+1} \leq c_q$ . Then multiplying Eq. (S48) by  $\langle\psi_q|$  and rearranging we get the following inequality:

$$c_q \leq \frac{gn}{(E(q) - gn - E) - c_{q+1}gn} c_{q-1} \leq \frac{gn}{E(q) - 2gn - E} c_{q-1} \leq c_{q-1}. \quad (\text{S49})$$

We can thus iterate this bound all the way to  $q = 1$  to get:

$$c_{>} \leq \prod_{q=1}^{q_*} \frac{gn}{E(q) - 2gn - E} c_{<} \leq \left( \frac{2gn}{E(1) - 2gn - E} \right)^{q_*} = \left( \frac{2g}{\epsilon_2 - \epsilon_1} \right)^{q_*} \equiv e^{-\lambda(g)n}, \quad (\text{S50})$$

where we used  $E = \epsilon_1 n$  and defined

$$\lambda(g) = \frac{\epsilon_2 - \epsilon_1 - 4g}{2\Delta E} \ln \left( \frac{\epsilon_2 - \epsilon_1}{2g} \right). \quad (\text{S51})$$

Finally, we can choose  $\Delta E \leq w_0 w_1 \frac{\epsilon_2 - \epsilon_1}{\epsilon_2 - \epsilon_1 - 4g}$  which finishes the proof.  $\square$

With Lemma SIII.3 in hand, we can now bound the numerator in Eq. (S39) as follows.

**Theorem SIII.4** (Stability of states with extensive energy barriers). *Let  $\mathcal{V}$  satisfying Eq. (S33) with  $\epsilon = E_{0,\min}(\mathcal{V})/n < \kappa/2$ . Then there exists  $\beta_* < \frac{\kappa - 2\epsilon}{\ln 2}$  and  $g_*(\beta) > 0$  such that for  $\beta > \beta_*$  and  $g < g_*(\beta)$ , the perturbed Hamiltonian  $H_0 + V$  has bottleneck ratio*

$$\Delta = \frac{\|P_{\partial\mathcal{V}}\rho_G\|_1}{\text{tr}(P_{\mathcal{V}}\rho)} = e^{-\Omega(n)}. \quad (\text{S52})$$

*Proof.* We have already bounded the denominator of the bottleneck ratio in Eq. (S40). We will now use Lemma SIII.3 to bound the numerator, assuming  $g \leq \kappa/8$ . To do so, consider a state  $|\varphi\rangle \in \partial\mathcal{V}$  which is guaranteed to be in the subspace  $\mathcal{H}_{>}$  relative to eigenstates  $|i\rangle$  of  $H$  with energy  $E_i \leq (\epsilon + \kappa/2)n$ . We can thus bound the expectation value of  $e^{-\beta H}$  in this case by expanding it in the eigenbasis and splitting the sum into its low- and high-energy parts:

$$\begin{aligned} \langle \varphi | e^{-\beta H} | \varphi \rangle &= \sum_{i: E_i \leq (\epsilon + \kappa/2)n} e^{-\beta E_i} |\langle i | \varphi \rangle|^2 + \sum_{i: E_i > (\epsilon + \kappa/2)n} e^{-\beta E_i} |\langle i | \varphi \rangle|^2 \\ &\leq \sum_{i: E_i \leq (\epsilon + \kappa/2)n} e^{-\beta E_i} e^{-\lambda_\kappa(g)n} + \sum_{i: E_i > (\epsilon + \kappa/2)n} e^{-\beta(\epsilon + \kappa/2)n} |\langle i | \varphi \rangle|^2 \\ &\leq 2^n e^{-\lambda_\kappa(g)n} + e^{-\beta(\epsilon + \kappa/2)n}, \end{aligned} \quad (\text{S53})$$

where  $\lambda_\kappa(g) \geq \frac{\kappa}{4w_0w_1} \ln \left( \frac{\kappa}{4g} \right)$ , with  $w_0, w_1$  being the constants defined in Lemma SIII.3.

To bound  $\text{tr}(P_{\partial\mathcal{V}}e^{-\beta H})$  we can apply Eq. (S53) for each term, which yields an additional factor of  $2^n$ . The partition function is also upper bounded by  $Z \leq 2^n$ . Combining all these and using Eq. (S39) we get an upper bound on the bottleneck ratio

$$\Delta^2 \leq (8^n e^{-\lambda_\kappa(g)n} + 4^n e^{-\beta(\epsilon + \kappa/2)n}) e^{2\beta(\epsilon + g)n} = e^{(3 \ln 2 - \lambda_\kappa(g) + 2\beta(\epsilon + g))n} + e^{(2 \ln 2 + \beta(\epsilon + 2g - \kappa/2))n}. \quad (\text{S54})$$

We can ensure that the right hand side is exponentially small in  $n$  provided that the following two conditions are met:

$$\beta \left( \frac{\kappa}{2} - \epsilon - 2g \right) > 2 \ln 2, \quad (\text{S55})$$

$$\lambda_\kappa(g) - 2\beta(\epsilon + g) > 3 \ln 2. \quad (\text{S56})$$

For  $\frac{\kappa - 2\epsilon}{\ln 2} < \beta < \infty$  we can always choose  $g$  small enough to ensure that both of these conditions are satisfied. For the first, we need  $4g < \kappa - 2\epsilon - \frac{1}{\beta} \ln 2$ . For the second, it is sufficient to satisfy

$$\frac{\kappa}{4w_0w_1} \ln \left( \frac{\kappa}{4g} \right) > 3 \ln 2 + \beta(\kappa/2 + \epsilon), \quad (\text{S57})$$

which is satisfied by taking

$$g < \frac{\kappa}{4} e^{-4w_0w_1(\beta(\kappa/2 + \epsilon) + 3 \ln 2)/\kappa}. \quad (\text{S58})$$

$\square$



**Remark S III.5.** A few closing remarks are in order about Thm. S III.4. First, note that the restriction on  $\epsilon < \kappa/2$  could easily be relaxed to  $\epsilon < \kappa$  by appropriately adjusting the various constant appearing in the proof. Thus the condition is that the minimal energy density in  $\partial\mathcal{V}$  is at least twice as that of  $\mathcal{V}$ . This condition was brought about because of the appearance of the square root in the numerator of Eq. (S39), which in turn corresponds to the worst case scenario, where the matrix elements of  $\rho_G$  that are off-block-diagonal with respect to the decomposition  $\mathcal{H} = \partial\mathcal{V} \oplus \partial\mathcal{V}^\perp$  are the same size as the matrix elements within the block  $\partial\mathcal{V}$ . In fact,  $H_0$  is block-diagonal with respect to this decomposition so the off-diagonal contributions go to zero in the limit  $g \rightarrow 0$ . We thus expect that the proof can be strengthened and possibly the restriction on  $\epsilon$  removed completely.

Similarly, the bound Eq. (S58) on  $g$  gets weaker with decreasing temperature, and goes to zero in the zero-temperature limit  $\beta \rightarrow \infty$ . This contrast with physical intuition that suggests that stability should be enhanced at lower temperatures, and we indeed expect that it should be possible to keep  $g$  finite in the zero temperature limit without invalidating the conclusions of the theorem.

- 
- [S1] David A Levin and Yuval Peres, *Markov chains and mixing times*, Vol. 107 (American Mathematical Soc., 2017).
- [S2] John Watrous, *The theory of quantum information* (Cambridge university press, 2018).
- [S3] Michael A Nielsen and Isaac L Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
- [S4] Benedikt Placke, Tibor Rakovszky, Nikolas Breuckmann, and Vedika Khemani, “Topological quantum spin glass order and its realization in qldpc codes,” (2024), in preparation.
- [S5] Tibor Rakovszky and Vedika Khemani, “The physics of (good) ldpc codes i. gauging and dualities,” arXiv preprint arXiv:2310.16032 (2023).
- [S6] M. Sipser and D.A. Spielman, “Expander codes,” *IEEE Transactions on Information Theory* **42**, 1710–1722 (1996).
- [S7] Anthony Leverrier, Jean-Pierre Tillich, and Gilles Zémor, “Quantum expander codes,” in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (2015) pp. 810–824.
- [S8] Irit Dinur, Min-Hsiu Hsieh, Ting-Chun Lin, and Thomas Vidick, “Good quantum ldpc codes with linear time decoders,” in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023 (Association for Computing Machinery, New York, NY, USA, 2023) p. 905–918.
- [S9] Chao Yin, Rahul Nandkishore, and Andrew Lucas, “Eigenstate localization in a many-body quantum system,” arXiv preprint arXiv:2405.12279 (2024).