

Lieb-Robinson Bounds with Exponential-in-Volume Tails

Ben T. McDonough¹, Chao Yin¹, Andrew Lucas^{1,*} and Carolyn Zhang²

¹*Department of Physics and Center for Theory of Quantum Matter, University of Colorado, Boulder, Colorado 80309, USA*

²*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*



(Received 21 February 2025; revised 30 June 2025; accepted 20 August 2025; published 3 November 2025)

Lieb-Robinson bounds demonstrate the emergence of locality in many-body quantum systems. Intuitively, Lieb-Robinson bounds state that, with local or exponentially decaying interactions, the correlation that can be built up between two sites separated by distance r after a time t decays as $\exp(vt - r)$, where v is the emergent Lieb-Robinson velocity. In many problems, it is important to also capture how much of an operator grows to act on r^d sites in d spatial dimensions. Perturbation theory and cluster expansion methods suggest that, at short times, these volume-filling operators are suppressed as $\exp(-r^d)$. We confirm this intuition, showing that, for $r > vt$, the volume-filling operator is suppressed by $\exp(-(r - vt)^d / (vt)^{d-1})$. This closes a conceptual and practical gap between the cluster expansion and the Lieb-Robinson bound. We then present two very different applications of this new bound. Firstly, we obtain improved bounds on the classical computational resources necessary to simulate many-body dynamics with error tolerance ε for any finite time t : as ε becomes sufficiently small, only $\varepsilon^{-O(d-1)}$ resources are needed. A protocol that likely saturates this bound is given. Secondly, we prove that disorder operators have volume-law suppression near the “solvable (Ising) point” in quantum phases with spontaneous symmetry breaking, which implies a new diagnostic for distinguishing many-body phases of quantum matter.

DOI: [10.1103/n242-m112](https://doi.org/10.1103/n242-m112)

I. INTRODUCTION

The Lieb-Robinson (LR) theorem [1] proves that quantum correlations and entanglement spread with at most a finite velocity in many-body quantum lattice models. While the original proof of this theorem is over 50 years old, it has recently become an extremely important technical tool in mathematical many-body physics [2]. The LR theorem underlies proofs of (1) the efficient simulatability of quantum dynamics on classical or quantum computers [3,4]; (2) lower bounds on the time needed to prepare entangled states in quantum information processors [5,6], including those with power-law interactions [7–17] and, in some cases, bosonic degrees of freedom [18–27]; (3) the prethermal robustness of gapped phases of matter [28] and the non-perturbative metastability of false vacua [29]; (4) the stability of topological order [30–32] and quantization of Hall conductance [33]; (5) the area-law of entanglement entropy in one dimension [34] (although there are

also combinatorial proofs); (6) clustering of correlations in gapped ground states [35,36]; (7) robustness of quantum metrology [37,38] (see Ref. [2] for a more complete list of applications). Besides these mathematical results, the intuition gained from the LR bound has been important in developing a new theory of many-body quantum chaos in lattice models, where the onset of chaotic behavior at early times is characterized by the growth of a Heisenberg-evolved operator from a short string of Pauli matrices to a long string [39,40]. This growth is controlled by a similar “Frobenius light cone,” which generally has a smaller velocity than the LR light cone [11]. For this reason, insight from a precise understanding of quantum operator growth may lead to improved classical algorithms to simulate hydrodynamics [41].

This diversity of applications, spanning quantum information sciences, atomic physics, condensed matter, and even high-energy physics, usually relies on a LR theorem stated as follows: given two local operators A_x and B_y , and a local many-body lattice model,

$$\| [A_x(t), B_y] \| \lesssim \exp[\mu(vt - d(x, y))], \quad (1)$$

where $d(x, y)$ denotes the distance between the degrees of freedom and μ is a constant. This bound is sufficient for many of the applications listed above. To give one example of how the simple bound (1) would be used, let us

*Contact author: andrew.j.lucas@colorado.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

briefly discuss how to bound the classical simulatability of quantum dynamics. Suppose that we wish to study the Heisenberg-evolved operator $A_x(t)$ by solving the Heisenberg equation of motion for A_x . Given finite computational resources, we simply truncate the list of Pauli strings we keep track of to those supported in a ball of radius R around site x . We refer to this truncated operator as the “fraction” of the operator acting within this ball. In d spatial dimensions, the number of such operators scales as $N \sim \exp[R^d]$. Equation (1) suggests that the error in this approximation ε will be bounded by $\varepsilon \lesssim \exp[vt - R]$. The classical resources necessary to simulate dynamics can thus be estimated as

$$N \sim \exp\left[(vt + \log \varepsilon^{-1})^d\right]. \quad (2)$$

For the rest of the introduction, $O(1)$ prefactors in scaling relations will be suppressed.

The argument above—as do many other applications of a LR bound—crucially relies on the *exponential decay in distance* in Eq. (1). Is that optimal? As phrased in Eq. (1), the bound is pretty much optimal: the tail bound can be improved to at most $\exp(-R \log R)$ with strictly local interactions [2]. This can be intuitively seen by noting that at order R , operator $A_x(t)$ can grow R sites away:

$$A_x(t) = A_x + it[H_{x+1,x}, A_x] + \cdots + \frac{(it)^R}{R!} [H_{x+R,x+R-1}, \dots, [H_{x+1,x}, A_x]] + \cdots. \quad (3)$$

On the other hand, when we bounded simulatability in $d > 1$, we used a LR bound for the spreading of an operator from x to y , and assumed that this same error controls how much of the operator might act on the whole ball of radius $d(x, y)$. This approximation might seem loose. Indeed, Eq. (3) suggests that the first terms that act on an entire ball of radius R arise at order R^d , so we might expect a tail bound that is suppressed in *volume*: $\exp(-(R - vt)^d)$. LR bounds of this kind are not known.

This understanding is physically important because it sharpens our picture of operator growth outside of the LR light cone. This has important applications to the simulatability of continuous-time dynamics and to extended operators such as disorder operators, which are used broadly in the literature. To illustrate the simulatability application, if A is a local operator and we study the Heisenberg time-evolved $A(t)$ expanded in the basis of Pauli strings, then a volume-law scaling bound suggests that the minuscule weight of volume-filling Pauli strings could overcome the large number of them— $\exp[O(t^d)]$ —to account for. We might then hope for an algorithm with an improved *polynomial* dependence in ε^{-1} :

$$N \lesssim \exp[(vt)^{d-1}(vt + \log \varepsilon^{-1})]. \quad (4)$$

Here $(vt + \log \varepsilon^{-1})$ is the radius corresponding to an error tolerance ε and the prefactor $(vt)^{d-1}$ is the surface area of the light cone, which scales with the number of paths [as in Eq. (3)] that reach the edge of the light cone.

In fact, using a very different approach—cluster expansions—it has recently been shown [42] that classical simulation algorithms do exist with polynomial error in ε^{-1} at a fixed time t :

$$N \sim \exp(e' \log \varepsilon^{-1}). \quad (5)$$

Despite scaling optimally with ε , this bound is weaker than Eq. (2) for large t and fixed ε . The mismatch in scaling with t and ε between Eqs. (2) and (5) would be closed by a bound such as in Eq. (4). To find such an optimal bound, which may also find many other applications, it is crucial to have better control over the shape of $A_x(t)$ outside of the “LR light cone”—the region within distance vt of the initial site x . This paper will address precisely this problem, and discuss the extent to which Eq. (4) might be achievable.

II. SUMMARY OF RESULTS

Our central objective is to better understand the “tail” of an operator—the fraction that acts outside the light cone. Quantum circuits with local gates have strict light cones, so time-evolved operators have no tails. This is one of the few cases where we cannot rely on intuition from quantum circuits [39,40] to inform us about the behavior of operators in continuous-time evolution. Intuitively, from Eq. (3), it appears that the exponential tail in the LR bound essentially arises due to *direct paths* between two points x and y . With a local Hamiltonian, terms arising from direct paths between points separated by a distance R grow in a sequential fashion, requiring $\Omega(R)$ steps. In contrast, if we want to fill the entire volume of sites a distance R from x , we should need $\Omega(R^d)$ terms in the series expansion (3) to ensure that all sites are touched at least once. This argument suggests that the weight of an operator that touches a finite fraction of the sites inside a ball of radius R should decay as $\exp[-R^d]$ rather than $\exp[-R]$, and that the Pauli strings grown along direct paths which dominate the behavior outside the light cone have a thin, “noodle-shaped” support [illustrated in Fig. 1(a)]. In $d > 1$, we must be careful about predicting the scaling directly from the number of terms in the perturbative expansion (3). As depicted in Fig. 1(b), the number of ways that an operator can grow to fill a large volume is exponentially large in volume itself. As we Taylor expand $A_x(t)$ at higher orders n in t , the larger support of $[H_{x_{n-1}}, \dots, [H_{x_1}, A_x]]$ means that there are yet more terms in H that may not commute with the nested commutator at the n th step. As is known [43,44], the growth in the number of possible terms is enough to break the convergence of the series. Therefore, if the weight of an operator that fills an

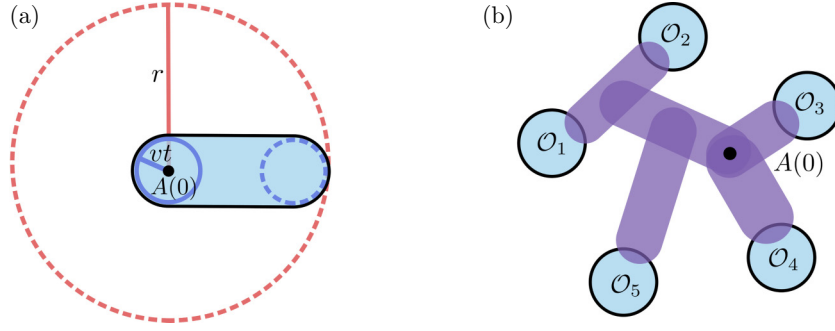


FIG. 1. (a) The dominant contribution to the growth of an operator outside of the LR light cone. Conventional LR bounds suggest that the fraction of $A(t)$ acting on the red dotted circle should be suppressed by $\exp[-O(r)]$. We show that it is suppressed by $\exp[-O(r^d)]$, which highly favors operators with “noodle-shaped” support like that shown in blue. (b) An illustration of why this problem is so difficult to approach with conventional techniques, which count the number of paths of subsets (illustrated in purple) that connect the support of $A(0)$ to the support of $O_1, \dots, O_m$. The combinatorics of counting branching paths that grow operators to have large support becomes difficult at higher orders. This makes the equivalence-class method a very valuable tool.

entire volume truly does decay as $\exp[-R^d]$, we must find a careful way to re-sum this series.

Our main technical tool for this re-summation is the equivalence-class formalism recently introduced in Ref. [45]. In Sec. III C, we use an example of a system with a weak link to illustrate the flexibility in choosing which paths to include in the bound. In Sec. IV we leverage this flexibility to introduce a reformulation of the problem, as illustrated in Fig. 2. This reformulation allows us to separate out the direct paths connecting sites within a ball of radius R to x . We phrase our theorem in terms of nested commutators as a tool for probing the support of an operator.

We use “local” to refer to Hamiltonians with finite-range interactions and “quasi-local” to refer to Hamiltonians with interactions that decay exponentially and sufficiently rapidly with the volume of their support, which may be slightly stronger than the way it has been used before in the literature. Our first result is to rigorously prove the following.

Theorem 1 (Theorems 7 and 8 (informal)). Suppose that $S_1, \dots, S_m$ are regions which are at least r apart. Let $O_1, \dots, O_m$ be operators that act nontrivially only on their respective regions, with operator norm $\|O_i\| = 1$ for simplicity. Let A be an operator with $\|A\| = 1$, and suppose that the distance between the support of A and any S_i is also at least r . If H is a local Hamiltonian then there exist constants $c_{\text{LR}}, v_{\text{LR}}$ independent of m such that

$$\|[O_m, [\dots, [O_1, A(t)]]]\| \leq c_{\text{LR}}^m \left(\frac{v_{\text{LR}} t}{r} \right)^{mr} \quad (6)$$

for any $v_{\text{LR}} t < r$. If H is quasi-local then, as long as the S_i are not “too close” together, the following slightly weaker bound holds:

$$\|[O_m, [\dots, [O_1, A(t)]]]\| \leq c_{\text{LR}}^m \exp(m\mu(v_{\text{LR}} t - r)) \quad (7)$$

with μ a constant independent of m . The value of c_{LR} is irrelevant as long as r is sufficiently large.

Intuitively, the nested commutator is a probe of the support of $A(t)$ because the operators $O_1, \dots, O_m$ supported on sets $S_1, \dots, S_m$ may be chosen to capture the contribution of Pauli strings in the expansion of $A(t)$ which act simultaneously on $S_1, \dots, S_m$. At a high level, it certainly would seem that Theorem 1 is a dramatic improvement over a conventional LR bound. Rather than scaling with the diameter of the support, the suppression depends on the volume. Just like conventional LR bounds, these results also hold for time-dependent systems; the intuition

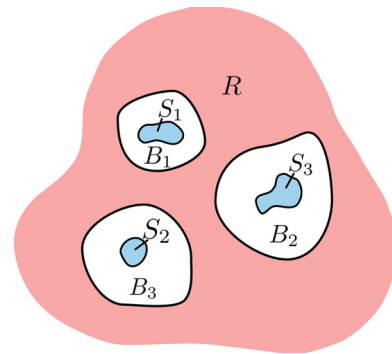


FIG. 2. Illustration of the problem setup. This illustration is schematic, and the underlying vertex set is not shown. Region R is depicted in red, and can be imagined as having infinite extent. Regions S_1, S_2, S_3 are depicted in blue, contained within the white regions B_1, B_2, B_3 which are disjoint from R . The shapes are arbitrary, but, for particular applications, we find optimal choices for the shape of these regions.

behind this generalization is that these worst-case bounds are established by counting intersections, and this combinatorial argument still works when time-dependence is introduced. We show how the proof can be generalized in Sec. IV D.

In some applications, the nested commutator itself is already an interesting object to bound. However, for applications such as bounding the simulatability of quantum dynamics on classical computers, directly characterizing the “fraction” of an operator that acts on a subset lying far outside the light cone is crucial. Specifically, an operator under time evolution becomes a complicated weighted sum of operators of different support, and such a characterization helps us design an efficient classical algorithm that keeps track of only the most important contributions to the sum.

To make the connection between the nested commutator bound and volume-law scaling precise, we observe that $m \sim (R/vt)^d$ such operators $\{O_i\}_{i=1}^m$ can fit inside of a ball of radius R , while ensuring that no light cones overlap. This leads to the following theorem.

Theorem 2 (Corollaries 5 and 7 and Lemma 8 (informal)). Given a local or quasi-local interaction and a local operator A on a cubic lattice, consider the Pauli decomposition $A(t) = \sum_{S \subseteq V} \sum_{i=1}^{3^{|S|}} \alpha_S^{(i)}(t) P_S^{(i)}$, where $P_S^{(i)}$ is a tensor product of Pauli operators supported on a subset S of vertices and $\{\alpha_S^{(i)}(t)\}$ are complex coefficients. Tile the square lattice with Euclidean boxes of side length sufficiently greater than $v_{\text{LR}}t$ such that the support of A is contained in a single box. Given a connected set of boxes $(B_1, \dots, B_m)$ called a “cluster,” where each box has side length r sufficiently greater than $v_{\text{LR}}t$, there exists μ, c_{LR} such that

$$\left\| \sum_{B_k \cap S \neq \emptyset, i} \alpha_S^{(i)}(t) P_S^{(i)} \right\| \leq c_{\text{LR}} \exp(-\mu r [3^{-d}m - 1]), \quad (8)$$

where the sum over S is taken over subsets which intersect $B_1, \dots, B_m$ non-trivially.

We expect that the restriction to a cubic lattice is just a technical convenience, and that a similar result should hold for a general graph embedded in d spatial dimensions.

The first application of this volume-law scaling bound is given in Sec. VI, where we explicitly derive an improved estimate for the simulation complexity of local and quasi-local quantum systems.

Theorem 3 (Corollary 8 (informal version)). Let H be a local or quasi-local Hamiltonian on a d -dimensional cubic lattice of qudits. If ρ is a state such that the marginal ρ_S on any connected set S can be obtained with at most exponential resources, then we provide an algorithm to compute the

expectation $\text{Tr}[A(t)\rho]$ to error ε with classical resources

$$N = \exp(O[(1 + v_{\text{LR}}t)^{d-1} (v_{\text{LR}}t + \log \varepsilon^{-1})]). \quad (9)$$

This complexity exhibits polynomial scaling in $1/\varepsilon$, and we argue that this bound has the best possible scaling in both ε and t . Note that this scaling is just slightly weaker than the naive expectation of Eq. (5). In particular, this result recovers the $\exp(O((v_{\text{LR}}t)^d))$ time-dependence that one would expect from a simulation with a strict light cone, such as a random unitary circuit [39,40], while maintaining the polynomial scaling in error tolerance ε for a fixed time with a suitably d - and t -dependent prefactor, which reduces to a constant in the case $d = 1$, as known from existing LR bounds.

We additionally explore the applications of bounds such as Eq. (1) in condensed matter physics. More concretely, we study spontaneous $\mathbb{Z}_2$ -symmetry-breaking phases of quantum matter, and rigorously close two gaps between known characterizations of symmetry-breaking phases and expectations from numerics and perturbation theory. First, conventional LR bounds [30] imply that the splitting $\delta = |\langle \psi_+ | H | \psi_+ \rangle - \langle \psi_- | H | \psi_- \rangle|$ between the two symmetry-broken ground states $|\psi_{\pm}\rangle$ in the ± 1 sectors of the global $\mathbb{Z}_2$ symmetry is at most exponentially small in the linear system size. However, the expectation from perturbation theory is that this splitting is exponentially small in system *volume*. Spontaneous symmetry breaking is also characterized by the rapid decay of the disorder operator $D_R := \prod_{i \in B_R(v)} X_i$ with R , which is the Ising symmetry restricted to a ball of radius R . Conventional LR bounds give a decay which is exponential in R , while the expectation from numerics and perturbation theory [46] is that this decay is exponential in the volume $\sim R^d$. We close these gaps with the following theorem.

Theorem 4 (Theorems 10 and 11 (informal)). If $|\psi_{\pm}\rangle$ are the ground states of a gapped, quasi-local Hamiltonian in d dimensions which are connected to GHZ states with eigenvalues ± 1 under the global $\mathbb{Z}_2$ symmetry by a finite-time evolution under a quasi-local Hamiltonian, then the splitting obeys the bound

$$|\langle \psi_+ | H | \psi_+ \rangle - \langle \psi_- | H | \psi_- \rangle| \leq c_{\text{LR}} L^d \exp\left(-\mu \frac{(L - v_{\text{LR}}t)^d}{(v_{\text{LR}}t)^{d-1}}\right), \quad (10)$$

where L is the linear size of the system and $c_{\text{LR}}, \mu, v_{\text{LR}}$ are constants. Furthermore, if $|\psi\rangle$ is connected to the polarized state $|\mathbf{0}\rangle$ by a finite-time evolution under a $\mathbb{Z}_2$ -symmetric,

quasi-local Hamiltonian, then the disorder parameter obeys

$$\langle \psi | D_R | \psi \rangle \leq c_{\text{LR}} \exp \left(-\mu \frac{(R - v_{\text{LR}} t)^d}{(v_{\text{LR}} t)^{d-1}} \right). \quad (11)$$

These results suggest a new diagnostic tool to distinguish quantum phases of matter, which we illustrate via a concrete example with Rokhsar-Kivelson-like states [47–49].

III. MATHEMATICAL PRELIMINARIES

The remainder of the paper will discuss the above results in a more formal way. We first introduce our notation and review some key results and ideas from previous work on LR bounds. We collect such results in this section.

A. Many-body quantum systems

To discuss locality in a many-body quantum system, it is often helpful to associate the quantum degrees of freedom (“qudits”) with the vertices of a graph. We can imbue the problem with a notion of spatial locality by adding edges between these vertices, providing a notion of distance between distinct qubits. We begin the formal discussion with problems where qudits only interact with their nearest neighbors on the resulting graph (with generalizations to exponentially decaying interactions discussed in Sec. V), and in this context it is often helpful to define a factor graph [45].

Definition 1 (Factor graph). A factor graph $G = (V, E, F)$ is a bipartite graph in which each node in the *node set* V is connected to the *factor set* F with *edges* $E \subseteq V \times F$. We assume that G is connected. If $(v, X) \in E$, we write $v \in X$: vertex $v \in V$ is connected to factor $X \in F$. The distance $d(x, y)$ between $x, y \in V$ is defined as the smallest number of factors contained in a connected path from x to y in G . For subsets $R, S \subset V$, $d(R, S) := \min_{x \in R, y \in S} d(x, y)$.

Definition 2. Given a subset $S \subset V$, define the boundary set

$$\partial S := \{x \in S : d(\{x\}, S^c) = 1\}. \quad (12)$$

The tail bounds we develop fully leverage the dimensionality of the system. To capture the dimensionality of an arbitrary graph, we use the following notion.

Definition 3 (d-dimensional system). Given a factor graph (V, E, F) , define the ball of radius r around vertex $v \in V$ as

$$B_r(v) := \{u \in V : d(u, v) \leq r\}. \quad (13)$$

We say that (V, E, F) is d dimensional if, for all $v \in V$ and sufficiently large r , there exist $O(1)$ constants $C_{1,2}$ and $C'_{1,2}$

such that

$$C'_1 r^d \leq |B_r(v)| \leq C_1 r^d, \quad (14a)$$

$$C'_2 r^{d-1} \leq |\partial B_r(v)| \leq C_2 r^{d-1}. \quad (14b)$$

Finite factor graphs represent a useful language to speak about many-body systems, as formalized in the definition below. We demonstrate all our results where V, E , and F are finite sets, but our formalism naturally extends to the thermodynamic limit.

Definition 4 (Many-body quantum system with nearest-neighbor interactions). Given a factor graph $G = (V, E, F)$, we define a many-body Hilbert space

$$\mathcal{H} = \bigotimes_{v \in V} (\mathbb{C}^q)_v =: (\mathbb{C}^q)^{\otimes V} \quad (15)$$

for some $q \in \mathbb{N}$. Here $(\mathbb{C}^q)_v$ reminds us that the q -dimensional Hilbert space is associated with a degree of freedom at vertex v . The Hamiltonian is a Hermitian operator

$$H = \sum_{X \in F} H_X, \quad (16)$$

where X , considered as a subset of V , is the support of H_X . We say that $\mathcal{H}, H$ defines a many-body system with few-body interactions if, for all $X \in F$, $|X| \leq \Delta$ for some fixed $O(1)$ constant Δ .

Definition 5 (Vector space of operators). The space of all linear operators acting on $\mathcal{H}$ forms a vector space $\mathcal{B}$. When we wish to emphasize this, we write $A \in \mathcal{B}$ as $|A\rangle$. There is a natural adjoint action of any $A \in \mathcal{B}$ on $\mathcal{B}$: $\text{ad}_A |B\rangle = |[A, B]\rangle$. We often use $A_S \in \mathcal{B}_S$ to denote that A is supported within $S \subset V$.

We can express the time evolution of an operator $\mathcal{O}$ in the Heisenberg picture using the notation above:

$$\frac{d}{dt} |\mathcal{O}(t)\rangle = i|[H, \mathcal{O}(t)]\rangle := \mathcal{L}|\mathcal{O}(t)\rangle. \quad (17)$$

The formal solution to this equation for a time-independent H , and thus a time-independent $\mathcal{L}$, is $|\mathcal{O}(t)\rangle = e^{t\mathcal{L}}|\mathcal{O}\rangle$. Since $H = \sum_{X \in F} H_X$, we can similarly associate a super-operator $\mathcal{L}_X := i \text{ad}_{H_X}$ with each H_X and, through the bilinearity of the commutator, write $\mathcal{L} = \sum_{X \in F} \mathcal{L}_X$. For most of this paper, for notational convenience, we restrict to the study of t -independent H . In Sec. IV D, we show how propagators are replaced with the appropriate time-ordered exponentials, while the remaining steps of the derivation remain largely unchanged.

B. Equivalence-class construction of LR bounds

A typical LR bound is stated as follows.

Theorem 5 (LR bound). Given a Hamiltonian H on factor graph $G = (V, F, E)$, there exist $O(1)$ constants C, μ, v_{LR} such that, for any two operators $A_{R,S} \in \mathcal{B}_{R,S}$ respectively, for subsets $R, S \subset V$,

$$\| [A_R(t), A_S] \| \leq C \min(|\partial R|, |\partial S|) e^{\mu(vt - d(R,S))}. \quad (18)$$

Here $\|\cdot\|$ denotes the operator norm (maximum singular value of its argument).

Despite its seemingly technical nature, this result has broad applications (see Ref. [2] for a recent review). Although the LR bound (18) is over 50 years old [1,50], we now focus on a much more recent [45] “equivalence-class formulation” of LR bounds, which we find will elegantly address the central question in this paper. Here, we review this equivalence-class formulation with a focus on the high-level ideas; a formal proof of the results outlined below follows from the more general method we develop in the next section.

Given operators A_i and $\mathcal{O}_j$ supported on sites i and j , respectively, conventional LR bounds deal with the commutator

$$C_{ij}(t) = \frac{1}{2} \| [\mathcal{O}_j, A_i(t)] \| = \frac{1}{2} \| \text{ad}_{\mathcal{O}_j} |A_i(t)\rangle \|, \quad (19)$$

analogously to Theorem 5, using single sites rather than sets for simplicity. We first express $|A_i(t)\rangle$ as

$$\text{ad}_{\mathcal{O}_j} e^{t\mathcal{L}} |A_i\rangle = \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{X_1, \dots, X_n \in F} \text{ad}_{\mathcal{O}_j} \mathcal{L}_{X_n} \cdots \mathcal{L}_{X_1} |A_i\rangle. \quad (20)$$

Each term in the sum can be uniquely labeled by a sequence of factors $M = (X_1, \dots, X_n)$. These sequences correspond to graphs whose nodes are elements of F , where X_1, X_2 are connected by an edge if they share at least one vertex, and the corresponding term is only non-vanishing if the graphs are connected. However, these graphs include redundancies, such as offshoots and cycles, which intuitively do not contribute to the growth of the operator from site i onto site j . A fundamental innovation of Ref. [45] is to formalize this intuition by bounding the commutator using only so-called *irreducible* (non-self-intersecting) paths between i and j on the factor graph. These paths correspond to sequences of factors that appear in the expansion of the time-evolution operator. Two possible connected acyclic subgraphs of the factor graph which differ only by redundant “reducible” terms but nevertheless have the same irreducible path are said to be in the same equivalence class. An example is illustrated pictorially in Fig. 3.

Grouping the terms in Eq. (20) by equivalence class, Chen and Lucas [45] obtained the following improved LR bound.

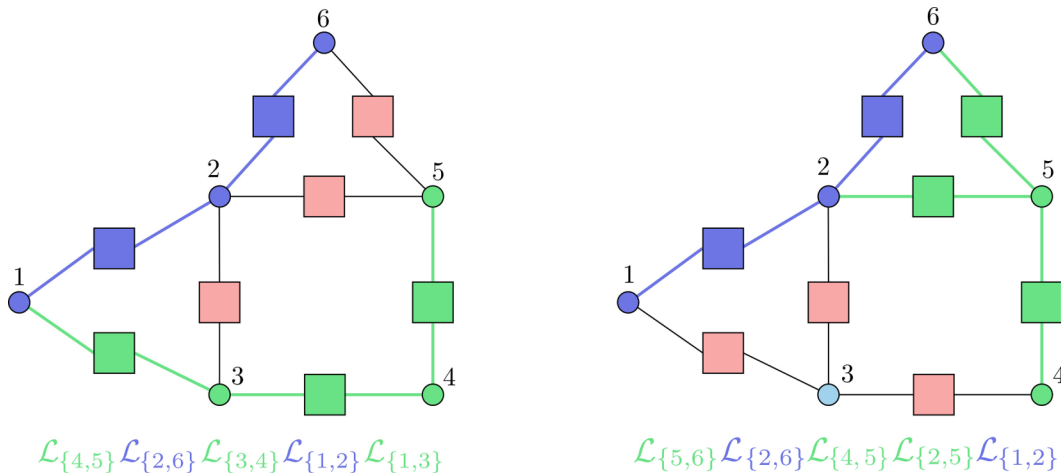


FIG. 3. Example of two sequences of factors that appear in the expansion of the Liouvillian propagator which have the same irreducible path and thus belong to the same equivalence class. The boxes are factors and the circles are vertices. The irreducible path is shown in blue, while the irrelevant terms in the sequence are shown in green. Note that the ordering of the terms in the sequence matters for the construction of the irreducible path; for instance, if $\mathcal{L}_{\{2,6\}}$ is to appear after $\mathcal{L}_{\{5,6\}}$ in the right figure, then the irreducible path would be $\{1, 2\}, \{2, 5\}, \{5, 6\}$.

Theorem 6 (Theorem 3 or Ref. [45]). If $\mathcal{O}_j \in \mathcal{B}_j$ and $A_i \in \mathcal{B}_i$ with $\|\mathcal{O}_j\| = \|A_i\| = 1$, then

$$C_{ij}(t) := \frac{1}{2} \|\text{ad}_{\mathcal{O}_j} e^{\mathcal{L}t} A_i\| \leq \sum_{[\Gamma] \in \mathcal{S}} \frac{(2|t|)^{|\Gamma|}}{|\Gamma|!} w(\Gamma), \quad (21)$$

where $w(\Gamma) = \prod_{X \in \Gamma} \|H_X\|$ is the weight of the path Γ , $[\Gamma]$ is an equivalence class of paths differing only by reducible terms, and $\mathcal{S}$ is the set of these equivalence classes.

The first main contribution of this work is to use similar techniques to prove a generalization of the previous theorem to nested commutators. As such, we will not prove Theorem 6 here, and instead defer the details of the proof to Sec. IV. We also generalize the following corollary to Theorem 6.

Corollary 1 (Corollary 6 of Ref. [45]). Define a symmetric real matrix h by

$$h_{ij} = \begin{cases} \sum_{X \in F: \{i,j\} \subseteq X} \|H_X\|, & i \neq j, \\ 0, & i = j. \end{cases} \quad (22)$$

Then

$$C_{ij}(t) \leq \exp(2h|t|)_{ij}. \quad (23)$$

Essentially, matrix h is the weighted adjacency matrix of a graph that encodes the connectivity of the Hamiltonian. It is a well-known graph-theoretic result that the matrix element $\exp(h)_{ij}$ sums the weighted paths from i to j . We likewise defer further details of the proof to Sec. IV.

C. Advantage over conventional LR bounds

There is a lot of freedom in choosing the equivalence classes in Theorem 6, which gives us the ability to obtain tighter bounds by sorting equivalence classes with a specific problem in mind. Many of the bounds to be obtained in future sections revolve around different methods of constructing equivalence classes. In this section, we provide a motivating example of such a situation.

Consider a chain consisting of four sites (see Fig. 4). The Hamiltonian is a sum of three terms $\{h_{1,2}, h_{2,3}, h_{3,4}\}$ with $\|h_{1,2}\| = \|h_{3,4}\| = h$ and a weak link $\|h_{2,3}\| = \varepsilon h$ with $\varepsilon \ll 1$. Let A be an operator supported on site 1 and B supported on site 4, both with unit norm. Using Theorem 6,

$$\|[A, B]\| \leq 2 \sum_{[\Gamma] \in \mathcal{S}_{ij}} \frac{(2|t|)^{|\Gamma|}}{|\Gamma|!} \prod_{X \in \Gamma} \|H_X\| = \frac{2\varepsilon(2h|t|)^3}{3!}, \quad (24)$$

which follows because $\Gamma = (\{1, 2\}, \{2, 3\}, \{3, 4\})$ is the only irreducible path connecting the two points.

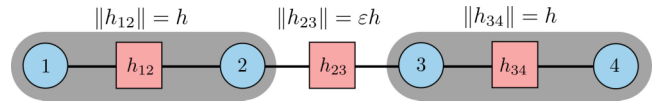


FIG. 4. Example system showing the advantage to shortening the irreducible paths.

The above does not take advantage of the fact that we know where the weak coupling is in the system. We could instead consider equivalence classes labeled by irreducible paths connecting $\{1, 2\}$ with $\{3, 4\}$, which in this simple example is trivially the factor $\mathcal{L}_{2,3}$. Similar ideas were described in Refs. [51, 52]. Applying the same theorem,

$$\|[A, B]\| \leq 4\varepsilon h|t|. \quad (25)$$

If $\varepsilon \ll 1$ then this bound is nontrivial (smaller than 1) out to later times $t \lesssim \varepsilon^{-1}$, versus $t \lesssim \varepsilon^{-1/3}$ for Eq. (24). The physical intuition for why this bound works is that we are effectively working in the interaction picture with $H_0 = h_{12} + h_{34}$, rotating out the contribution of these terms to specifically target the weak link.

IV. NESTED COMMUTATOR BOUNDS

Having reviewed the standard single-commutator LR bound, we now develop a formalism to generalize Theorem 6 to nested commutators. The proof technique is quite similar; however, there are additional details to check, including a convenient choice of equivalence classes and the need to account for the ordering of terms between the different irreducible paths, which was automatic in the case of a single commutator.

This section focuses on systems with only nearest-neighbor interactions on some interaction graph. The generalization to systems with exponential-tailed interactions (which is important for some of our applications) will be presented in Sec. V.

A. Equivalence classes of causal trees

Let H be a Hamiltonian associated with the factor graph (F, E, V) . Let $\mathcal{O}_1, \dots, \mathcal{O}_m$ be operators with $\|\mathcal{O}_i\| = 1$ and disjoint supports $S_1, \dots, S_m$, respectively, which are contained in regions $B_1, \dots, B_m$. Consider an operator A with $\|A\| = 1$ whose support is a set R contained within the complement of $B_1, \dots, B_m$, as shown in Fig. 2. We are interested in bounding the nested commutator

$$C_S^R(t) = \sup_{\substack{\mathcal{O}_1, \dots, \mathcal{O}_m \\ A}} \frac{1}{2^m} \|[\mathcal{O}_m, [\dots, [\mathcal{O}_1, A(t)]] \dots]\|. \quad (26)$$

Note that $\prod_{i=1}^m \text{ad}_{\mathcal{O}_i}$ is a superoperator that acts trivially on any operator A such that $\text{supp}(A)$ (denoting the support)

ALGORITHM 1. An algorithm for generating a causal forest.

```

 $T_0 \leftarrow \{R\}$ 
 $M_0 \leftarrow R$ 
for  $n \in \{1, \dots, |M|\}$  do
   $T_n \leftarrow T_{n-1}$ 
  if  $\exists k < n$  s.t.  $M_n \cap M_k \neq \emptyset$  then
    if  $\neg \exists j < n$  s.t.  $M_n \cap M_j = M_n$  then
       $k \leftarrow \min(\{k \text{ s.t. } M_n \cap M_k \neq \emptyset\})$ 
       $T_n \leftarrow T_{n-1} \cup (M_n, (M_k, M_n))$ 
    end if
    for  $i$  s.t.  $S_i \cap M_n \neq \emptyset$  and  $S_i \notin T_{n-1}$  do
       $T_n \leftarrow T_n \cup (M_n, (S_i, M_n))$ 
    end for
  else
    return  $\emptyset$ 
  end if
end for
return  $T_{|M|}$ 

```

is disjoint from $\text{supp}(\mathcal{O}_i)$ for each i . This property is independent of the ordering of the nested commutator due to the Jacobi identity and the observation that $[\text{ad}_{\mathcal{O}_i}, \text{ad}_{\mathcal{O}_j}] = \text{ad}_{[\mathcal{O}_j, \mathcal{O}_i]} = 0$ by our assumption that the supports of $\mathcal{O}_i, \mathcal{O}_j$ are disjoint.

As illustrated before, we can rewrite the time-evolved operator A as

$$|A(t)\rangle = e^{\mathcal{L}t}|A\rangle = \sum_M \frac{t^{|M|}}{|M|!} \mathcal{L}_{M_{|M|}} \cdots \mathcal{L}_{M_1}|A\rangle, \quad (27)$$

where, again, $M = (M_1, M_2, \dots, M_{|M|})$ is a sequence of factors. The essence of our approach is to choose the appropriate equivalence classes to quantify the way that $A(t)$ “leaks” into the regions $B_1, \dots, B_m$ over time. To accomplish this, we construct an algorithm to map each M to a graph on subsets of V , modifying the construction in Ref. [45]. The algorithm takes as input $S_1, S_2, \dots, S_m, R$, and M , where M is a sequence of factors obtained from a term in the Liouville equation as described above, and outputs a forest (disjoint union of trees) $T(M)$, which is a graph whose nodes are subsets of V . The algorithm proceeds as follows.

The result of this algorithm is the forest $T(M)$.

Definition 6. If $S_1, \dots, S_m$ are nodes in the forest $T(M)$ then we call $T(M)$ a causal forest.

Proposition 1. If $T(M)$ is not a causal forest then $\prod_i \text{ad}_{\mathcal{O}_i} \mathcal{L}_{M_{|M|}} \mathcal{L}_{M_{|M|-1}} \cdots \mathcal{L}_{M_1}|A_R\rangle = 0$.

Proof. Let $S_i \notin T(M)$. By the design of the algorithm, this means that M does not have an ordered subsequence of factors connecting R to S_i . Since the factors correspond to the support of terms in the Hamiltonian, this means that

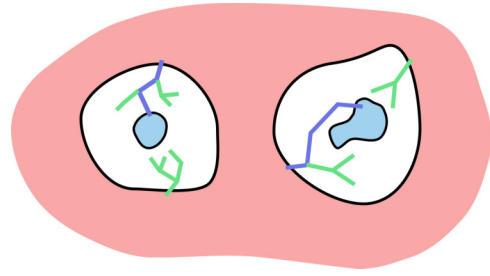


FIG. 5. Example of a causal forest constructed by the algorithm. The red region again corresponds to R and the two blue regions depict S_1, S_2 . The blue and green line segments represent factors in the sequence M . The blue segments participate in the irreducible paths, and so they form part of the corresponding irreducible skeleton, while the green ones are “reducible” terms. The ordering of the sequence is not depicted.

$|B\rangle := \mathcal{L}_{M_{|M|}} \mathcal{L}_{M_{|M|-1}} \cdots \mathcal{L}_{M_1}|A_R\rangle$ does not have support on S_i , so $\text{ad}_{\mathcal{O}_i}|B\rangle = 0$. ■

Definition 7. If $T(M)$ is a causal forest then, by construction, each S_i is contained in a tree rooted at R , so there is a unique path connecting it to R . Excluding the endpoints at R and S_i , this path corresponds to a unique sequence $\Gamma_i = (X_1, \dots, X_{l_i})$ of factors $X_i \in F$. We call $\mathcal{S}$ the set of equivalence classes of causal forests. An equivalence class of causal forests is defined by pruning all the “reducible” terms (those not involved in Γ_i) from M , arriving at a minimal ordered sequence Λ which will be called the irreducible skeleton.

An illustration of a causal forest constructed by the algorithm and the corresponding irreducible paths is depicted in Fig. 5.

B. Bounds from irreducible paths

The main technical result of this section is to show that, under the assumptions outlined above, we can derive a bound on the nested commutator which resembles a product of the single-commutator bounds presented previously. This is captured by the following result.

Theorem 7. Consider a local interaction defined by the factor graph (V, E, F) . Let $S_1, \dots, S_m$ be sets contained in disjoint regions $B_1, \dots, B_m$ separated by at least one of the subgraph of factors. Consider an operator A with $\|A\| = 1$ whose support is a set R contained within the complement of $B_1, \dots, B_m$. Let $\mathcal{O}_1, \dots, \mathcal{O}_m$ be operators also with norm 1, where each $\mathcal{O}_i$ is supported within S_i . Then on the nested

commutator we have the bound

$$\begin{aligned} & \frac{1}{2^m} \|[\mathcal{O}_m, [\dots, [\mathcal{O}_1, A(t)] \dots]]\| \\ & \leq \prod_i \sum_{\Gamma \in \Gamma_i(R \rightarrow S_i)} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} w(\Gamma), \end{aligned} \quad (28)$$

where $\Gamma_i(R \rightarrow S_i)$ denotes the set of non-self-crossing paths within B_i from R to S_i and $w(\Gamma) = \prod_{X \in \Gamma} \|H_X\|$ is the weight of the path Γ .

We begin by reorganizing the terms in Eq. (20) by equivalence class:

$$\begin{aligned} \prod_i \text{ad}_{\mathcal{O}_i} e^{\mathcal{L}_i} |A_R\rangle &= \prod_i \text{ad}_{\mathcal{O}_i} \sum_{[T] \in \mathcal{S}} \sum_{M: T(M) \in [T]} \frac{t^{|M|}}{|M|!} \mathcal{L}_{X_{|M|}} \\ &\quad \dots \mathcal{L}_{X_1} |A_R\rangle. \end{aligned} \quad (29)$$

By definition, each $T(M) \in [T]$ has the same set of irreducible paths $\{\Gamma_i\}$, so our goal is to rewrite the second sum in Eq. (29) by “factoring out” the contribution from the irreducible paths in the tree, then grouping together and safely re-summing the extraneous terms. This is accomplished in the following lemma.

Lemma 1. Fix an equivalence class of causal forests $[T]$ with irreducible paths $\{\Gamma_i\}$ and irreducible skeleton Λ . Then

$$\begin{aligned} & \prod_i \text{ad}_{\mathcal{O}_i} \sum_{M: T(M) \in [T]} \frac{t^{|M|}}{|M|!} \mathcal{L}_{X_n} \dots \mathcal{L}_{X_1} |A_R\rangle \\ &= \prod_i \text{ad}_{\mathcal{O}_i} \sum_{m_1, m_2, \dots, m_l} \frac{t^n}{n!} (\mathcal{L}_l)^{m_l} \mathcal{L}_{\Lambda_l} (\mathcal{L}_{l-1})^{m_{l-1}} \\ &\quad \dots \mathcal{L}_{\Lambda_1} (\mathcal{L}_0)^{m_0} |A_R\rangle, \end{aligned} \quad (30)$$

where $l := |\Lambda|$ and $n := l + \sum_i m_i$. We have defined

$$\mathcal{L}_i := \sum_{X \cap V_i = \emptyset} \mathcal{L}_X \quad (31)$$

such that V_j is the set of “disallowed” vertices: any factor that intersects V_j would necessarily change the equivalence class of $T(M)$ if it was to appear in M in between Λ_j and Λ_{j+1} . We can break V_j up into the terms disallowed by each path, i.e.

$$V_j = \bigcup_i U_{i, \beta_i(j)}, \quad (32)$$

where $\beta_i(j) = \#\{k \leq j : i_k = i\}$ counts the number of terms in sequence Λ at or before position j that belong

to Γ_i . Let $l_i = |\Gamma_i|$. Then U_{ij} is explicitly defined as

$$U_{ij} = \begin{cases} \bigcup_{m=\beta_i(j)+2}^{l_i} (\Gamma_i)_m, & j < l_i - 1, \\ S_i, & j = l_i - 1. \end{cases} \quad (33)$$

Proof. We follow Ref. [45]. First, we show that every term on the right-hand side also appears in the summand on the left-hand side of Eq. (30). Fix an equivalence class $[T]$ with an irreducible skeleton $\Lambda = (\Lambda_1, \dots, \Lambda_l)$. Let $U_i := \{\mathcal{L}_X : X \in F, X \cap V_i = \emptyset\}$, the set of allowed terms that could appear between Λ_i and Λ_{i+1} , forming another factor sequence M such that $[T(\Lambda)] = [T(M)]$. Then we can write

$$\begin{aligned} & (\mathcal{L}_l)^{m_l} \mathcal{L}_{\Lambda_l} (\mathcal{L}_{l-1})^{m_{l-1}} \dots \mathcal{L}_{\Lambda_1} (\mathcal{L}_0)^{m_0} \\ &= \left(\sum_{X_1, \dots, X_{m_l} \in U_l} \mathcal{L}_{X_1} \dots \mathcal{L}_{X_{m_l}} \right) \mathcal{L}_{\Lambda_l} \\ &\quad \times \left(\sum_{X_1, \dots, X_{m_{l-1}} \in U_{l-1}} \mathcal{L}_{X_1} \dots \mathcal{L}_{X_{m_{l-1}}} \right) \dots \\ &\quad \times \mathcal{L}_{\Lambda_1} \left(\sum_{X_1, \dots, X_{m_0} \in U_0} \mathcal{L}_{X_1} \dots \mathcal{L}_{X_{m_0}} \right). \end{aligned} \quad (34)$$

Each term in parentheses is a sum over all possible sequences of allowed terms of length m_k , with repeated entries allowed. It is permissible for sequences which do not contribute to a causal tree to appear in the sum because the contribution of these terms must ultimately vanish. By construction, none of the terms $\mathcal{L}_{X_k}$ intersect any of the irreducible paths Γ_i past $\beta_i(k)$, and so they may not create a path connecting R to S_i . Therefore, any sequence appearing on the right-hand side which has a non-vanishing contribution still corresponds to a tree $T(M)$ with the irreducible skeleton Λ . This shows that every non-vanishing term in the summand on the right-hand side corresponds to a causal forest in the equivalence class $[T(\Lambda)]$, and so it appears on the left-hand side.

For the other direction, every $T(M) \in [T(\Lambda)]$ may be constructed as $M_l \Lambda_l M_{l-1} \dots \Lambda_1 M_0$, where $M_l, \dots, M_0$ are arbitrary sequences of terms that form part of the causal forest, but do not create additional paths between R and $\{S_1, \dots, S_j\}$; otherwise, they would change the equivalence class of $T(M)$. As explained above, these are exactly the sequences that appear in the sum on the right-hand side. Furthermore, each term on the left and each term in the summand on the right correspond to a unique M . This establishes a bijection between the non-vanishing terms in the two sums, proving the result. ■

Definition 8 (n-simplex). We define the canonical n -simplex with side length t , denoted by $\Delta^n(t)$, as an n -dimensional shape bounded by $\{x_1 < x_2 < \dots < x_n < t\}$, which generalizes a triangle in two dimensions and a tetrahedron in three dimensions to n dimensions. The volume of $\Delta^n(t)$ is given by

$$\text{vol } \Delta^n(t) = \int_0^t dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{n-1}} dt_n = \frac{t^n}{n!}. \quad (35)$$

In the next step, we can eliminate the contribution of extraneous “reducible” terms and include in the bound only contributions from the irreducible paths. The engine behind this re-exponentiation is the generalized Schwinger-Karplus identity.

Lemma 2 (Lemma 5 of Ref. [45]). The following identity holds:

$$\begin{aligned} \sum_{\vec{m}=0}^{\infty} \frac{t^n}{n!} \mathcal{F}_l^{m_l} A_l \dots \mathcal{F}_1^{m_1} A_1 \mathcal{F}_0^{m_0} \\ = \int_{\Delta^l(t)} d\vec{t} e^{\mathcal{F}_l \Delta t_l A_l} \dots e^{\mathcal{F}_1 \Delta t_1 A_1} e^{\mathcal{F}_0 \Delta t_0}. \end{aligned} \quad (36)$$

On the left-hand side we defined the term order $n := l + \sum_{k=1}^l m_k$ and on the right-hand side, $\Delta^l(t)$ is the canonical l -simplex, and $\Delta t_i = t_{i+1} - t_i$, with $\Delta t_l := t - t_l$ and $\Delta t_0 := t_1$.

Proof of Theorem 7. By applying Lemmas 1 and 2 to Eq. (30), we obtain

$$\begin{aligned} \left\| \frac{1}{2^m} \prod_i \text{ad}_{\mathcal{O}_i} e^{\mathcal{L}_i} |A_R\rangle \right\| &\leq \frac{1}{2^m} \sum_{[T] \in \mathcal{S}} \left\| \prod_i \text{ad}_{\mathcal{O}_i} \sum_{T(M) \in [T]} \frac{t^{|M|}}{|M|!} \mathcal{L}_{X_{|M|}} \dots \mathcal{L}_{X_1} |A_R\rangle \right\| \\ &\leq \sum_{\Lambda \in \mathcal{S}} \left\| \sum_{m_1, m_2, \dots, m_l} \frac{t^n}{n!} (\mathcal{L}_l)^{m_l} \mathcal{L}_{\Lambda_l} (\mathcal{L}_{l-1})^{m_{l-1}} \dots \mathcal{L}_{\Lambda_1} (\mathcal{L}_0)^{m_0} |A_R\rangle \right\| \\ &\leq \sum_{\Lambda \in \mathcal{S}} \int_{\Delta^l(t)} d\vec{t} \left\| \exp(\mathcal{L}_l \Delta t_l) \mathcal{L}_{\Lambda_l} \exp(\mathcal{L}_{l-1} \Delta t_{l-1}) \dots \mathcal{L}_{\Lambda_1} \exp(\mathcal{L}_0 \Delta t_0) |A_R\rangle \right\| \\ &\leq \sum_{\Lambda \in \mathcal{S}} \text{vol}(\Delta^l(t)) 2^l \prod_{i=1}^l \|H_{\Lambda_i}\|. \end{aligned} \quad (37)$$

In the second line we use the fact that $\|\text{ad}_{\mathcal{O}_i}\|_{op} \leq 2$ for $\|\mathcal{O}_i\| = 1$, and in the fourth line we use the unitary invariance and submultiplicativity of the operator norm. Since $l = \sum_j |\Gamma_j|$ by definition, the canonical n -simplices have the property that

$$\text{vol}(\Delta^l(t)) = \frac{1}{l!} \prod_j |\Gamma_j|! \text{vol}(\Delta^{|\Gamma_j|}(t)), \quad (38)$$

so we can rewrite the right-hand side of Eq. (37) as

$$\begin{aligned} \sum_{\Lambda \in \mathcal{S}} \text{vol}(\Delta^l(t)) 2^l \prod_{i=1}^l \|H_{\Lambda_i}\| \\ = \sum_{\{\Gamma_j\}: (\Gamma_1, \dots, \Gamma_m) \in \mathcal{S}} \frac{1}{l!} \prod_{j=1}^m |\Gamma_j|! \text{vol}(\Delta^{|\Gamma_j|}(t)) 2^{|\Gamma_j|} w(\Gamma_j), \end{aligned} \quad (39)$$

where the sum on the right-hand side is taken over all possible sets of irreducible paths $\{\Gamma_j\}$ such that the irreducible skeleton $\Lambda = (\Gamma_1, \dots, \Gamma_m)$ obtained from their concatenation belongs to $\mathcal{S}$. The summand does not depend on the ordering of the individual paths, only on the set of irreducible paths $\{\Gamma_j\}$ so we can reorganize the sum over sets of paths $\{\Gamma_j\}$:

$$\begin{aligned} \sum_{\Lambda \in \mathcal{S}} \text{vol}(\Delta^l(t)) 2^l \prod_{i=1}^l \|H_{\Lambda_i}\| \\ = \sum_{\{\Gamma_i\}} \frac{l!}{|\Gamma_1|! \dots |\Gamma_m|!} \frac{1}{l!} \prod_{i=1}^m |\Gamma_i|! \text{vol}(\Delta^{|\Gamma_i|}(t)) 2^{|\Gamma_i|} w(\Gamma_i) \\ = \sum_{\{\Gamma_i\}} \prod_{i=1}^m \text{vol}(\Delta^{|\Gamma_i|}(t)) 2^{|\Gamma_i|} w(\Gamma_i) \\ = \prod_i \sum_{\Gamma \in \Gamma_i(R \rightarrow S_i)} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} w(\Gamma). \end{aligned} \quad (40)$$

Hence we obtain Eq. (28). ■

C. Combinatorial bounds

Although Eq. (28) is formally quite elegant, it is not necessarily always convenient to work with. In this subsection, we obtain more physically transparent nested commutator bounds by further simplifying Eq. (28). First, the nested commutator can be bounded using the equivalence-class bounds on single commutators from Theorem 6.

Corollary 2 (Decoupling corollary). Let $C_{ij}(t)$ be the bound on the single commutator in Theorem 6. Then

$$\frac{1}{2^m} \|[\mathcal{O}_m, [\dots, [\mathcal{O}_1, A(t)] \dots]]\| \leq \prod_i \sum_{\substack{u \in \partial B_i \\ v \in \partial S_i}} C_{uv}(t). \quad (41)$$

This bound is not quite as tight as Theorem 7 because we now include paths that enter region R . At short times, we find that the following result holds.

Corollary 3. Suppose that $\|H_X\| \leq h$ for all X that intersect with $B_1, \dots, B_m$. Let $r_i := d(R, S_i)$, where distance is defined via the factor graph as in Definition 3, and Δ be the maximum degree of the factor graph. If

$$|t| < \frac{\min_i(r_i)}{2h\Delta} \quad (42)$$

then

$$C_S^R(t) \leq \prod_i |\partial B_i| |\partial S_i| \left(\frac{2eh|t|\Delta}{r_i} \right)^{r_i}, \quad (43)$$

where $C_S^R(t)$ is defined as in Eq. (26).

Proof. The number of distinct non-self-crossing paths of length l from $u \in \partial B_i$ to v_i must be smaller than Δ^l . Given a ball B_i , the minimum length of path Γ_i in a causal tree is $r_i := d(R, v_i)$. Therefore we have

$$\sum_{\Gamma_i(R \rightarrow v_i)} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} w(\Gamma) \leq \sum_{l=r_i}^{\text{vol}(B_i)} \frac{(2th\Delta)^l}{l!}. \quad (44)$$

For any $\alpha > 1$, $\alpha^{l-r_i} > 1$ whenever $l > r_i$. We therefore find that

$$\sum_{l=r_i}^{\text{vol}(B_i)} \frac{(2th\Delta)^l}{l!} \leq \inf_{\alpha > 1} \sum_{l=0}^{\infty} \frac{(2th\Delta)^l}{l!} \alpha^{l-r_i} = \inf_{\alpha > 1} \alpha^{-r_i} e^{2th\alpha\Delta}. \quad (45)$$

If Eq. (42) holds, plugging $\alpha = (2h\Delta t)^{-1} r_i > 1$ into Eq. (45) gives

$$\inf_{\alpha > 1} \alpha^{-r_i} e^{2th\alpha\Delta} \leq \left(\frac{2eh\Delta|t|}{r_i} \right)^{r_i}. \quad (46)$$

Inserting this bound into Eq. (28), we have

$$C_S^R(t) \leq \prod_i |\partial B_i| |\partial S_i| \left(\frac{2eh\Delta|t|}{r_i} \right)^{r_i}, \quad (47)$$

as desired. ■

We can also encode the combinatorial problem of summing weighted paths between two points on the factor graph into a matrix exponential, which can be computationally useful.

Corollary 4. Define a real-symmetric matrix h elementwise by

$$h_{ij} = \begin{cases} \sum_{X \in F: i, j \in X} \|H_X\|, & i \neq j, \\ 0, & i = j. \end{cases} \quad (48)$$

Then

$$C_S^R(t) \leq \prod_i \sum_{\substack{u \in \partial B_i \\ v \in \partial S_i}} \exp(2h|t|)_{uv}. \quad (49)$$

Proof. The result follows immediately by combining Corollary 1 with Corollary 2. ■

Perhaps most importantly, we would like to have a bound that makes the *volume-law decay* of large nested commutators extremely transparent. This is accomplished with the following result.

Corollary 5. Let B_R denote a metric ball of radius R with respect to the geodesic distance on the factor graph (as defined in Definition 3). There exist constants $\gamma, c_{\text{LR}} > 0$ such that we can choose $v_1, \dots, v_m \in B_R$ (where m depends on R) for which

$$C_{v_1, \dots, v_m}^R(t) \leq c_{\text{LR}} \exp\left(-\gamma \frac{(R - v_{\text{LR}}t)^d}{(v_{\text{LR}}t)^{d-1}}\right) \quad (50)$$

for any $v_{\text{LR}}t > 1$.

Proof. We choose $S_1, \dots, S_m$ to be vertices $v_1, \dots, v_m$ and $B_1, \dots, B_m$ to be metric balls $B_i = B_\xi(v_i)$ of radius ξ centered on these vertices. We optimize ξ to obtain Eq. (5). First suppose that $d > 1$. For a lattice in d dimensions,

there exists a constant C_1 such that $|\partial B_\xi| \leq (C_1 \xi)^{d-1}$. Then directly applying Corollary 3, we have

$$C_v^R(t) \leq \left[(C_1 \xi)^{d-1} \left(\frac{v_{\text{LR}} t}{\xi} \right)^\xi \right]^m. \quad (51)$$

Let $\xi = \alpha v_{\text{LR}} t$ for some $\alpha > 1$. The requirement that ensures that $\alpha > 1$ ensures that we can apply Corollary 3. There exists a constant C_2 and, for any α , a constant K , such that, for all $v_{\text{LR}} t > 1$ and all $R > K v_{\text{LR}} t$, we can find m distinct points $v_1, \dots, v_m$ in B_R separated pairwise by more than 2ξ , where $m > C_2(R/\xi)^d$. The bound above may then be written as

$$\begin{aligned} C_v^{B_R} &\leq \left[(C_1 \alpha v_{\text{LR}} t)^{(d-1)\alpha^{-d}(v_{\text{LR}} t)^{-1}} \left(\frac{1}{\alpha} \right)^{\alpha^{1-d}} \right]^{C_2 R^d (v_{\text{LR}} t)^{1-d}} \\ &\leq \left[\left(\frac{e^{C_1(d-1)}}{\alpha} \right)^{\alpha^{1-d}} \right]^{C_2 R^d (v_{\text{LR}} t)^{1-d}}. \end{aligned} \quad (52)$$

Then, for any $\alpha > \exp(C_1(d-1))$, we have $(e^{C_1(d-1)} \alpha^{-1})^{\alpha^{d-1}} := e^{-\beta} < 1$. Plugging this in gives us

$$C_v^{D^c}(t) \leq \exp(-\beta C_2 R^d / (v_{\text{LR}} t)^{d-1}). \quad (53)$$

Additionally, β achieves a maximum at $\alpha = \exp(C_1(d-1) + 1/(d-1)) > 1$ of $\beta = \exp(-C_1(d-1)^2 - 1)/(d-1)$. This holds for $R > K v_{\text{LR}} t$, but we would like a bound that holds for all $R > v_{\text{LR}} t$. If $R/v_{\text{LR}} t < K$ then we can apply the conventional LR bound (Theorem 5). Choosing v_1 at the center of B_R , we have

$$C_v^{B_R} \leq \frac{1}{2} \|\mathcal{O}_{v_1}, A(t)\| \leq c_{\text{LR}} e^{\mu(v_{\text{LR}} t - R)}. \quad (54)$$

Let $C'_2 := \beta C_2$, and, for simplicity, assume that $c_{\text{LR}} \geq 1$. For any $\lambda \leq 1$, we clearly have

$$\begin{aligned} \exp(-C'_2 R^d / (v_{\text{LR}} t)^{d-1}) &\leq c_{\text{LR}} \\ \exp(-C'_2 \lambda (R - v_{\text{LR}} t)^d / (v_{\text{LR}} t)^{d-1}). \end{aligned} \quad (55)$$

Then we find that

$$e^{\mu(v_{\text{LR}} t - R)} \leq \exp(-C'_2 \lambda (R - v_{\text{LR}} t)^d / (v_{\text{LR}} t)^{d-1}) \quad (56)$$

when $\lambda \leq \mu/[C'_2(R/v_{\text{LR}} t - 1)^{d-1}]$. If $1 < R/v_{\text{LR}} t < K$ then $\lambda = \min(\mu/[C'_2(K-1)^{d-1}], 1) < \mu/[C'_2(R/v_{\text{LR}} t - 1)^{d-1}]$, so, for all $R/v_{\text{LR}} t > 1$, we have

$$C_v^{B_R}(t) \leq \exp(-\lambda C'_2 (R - v_{\text{LR}} t)^d / (v_{\text{LR}} t)^{d-1}). \quad (57)$$

This completes the proof. $\blacksquare$

D. Generalizing to time-dependent interactions

LR bounds, at least those developed here, are “worst-case-scenario” bounds. This is because the operator norm measures the *most* that an operator can increase the norm of a vector, and then we repeatedly apply the triangle inequality. Thus, the argument that leads to the bounds is fundamentally combinatorial in nature, and it should be unsurprising that these arguments generalize to time-dependent systems. In this subsection, we explain why the argument in Sec. IV generalizes to time-dependent systems.

In the time-dependent case, we can instead expand the propagator as a Dyson series. Let $\mathcal{L}(t) := -i \text{ad}_{H(t)}$ be the system Liouvillian. Furthermore, we write $\mathcal{L}_X(t) := -i \text{ad}_{H_X(t)}$ for $X \subset V$ when the support of $H_X(t)$ is contained within X at all times. The Dyson series expansion is

$$\begin{aligned} e^{\mathcal{L}t} &= \sum_{X_1, \dots, X_n \subset V} \int_0^t ds_n \int_0^{s_n} ds_{n-1} \cdots \int_0^{s_2} ds_1 \mathcal{L}_{X_n}(s_n) \\ &\quad \cdots \mathcal{L}_{X_1}(s_1) \\ &:= \sum_{X_1, \dots, X_n \subset V} \mathcal{I}_t^{(n)}(\mathcal{L}_{X_1}, \dots, \mathcal{L}_{X_n}), \end{aligned} \quad (58)$$

where n is summed over and we have defined $\mathcal{I}_t^{(n)}$ as the convolution operator above labeling sequences of operators appearing in the expansion. This is the appropriate generalization of the Taylor-series expansion in the time-independent case (20). Despite the time-dependence, the terms in Hamiltonian $H_X(t)$ have supports which are bounded by X , so the algorithm for constructing the equivalence classes is unaffected. Therefore, we simply need to develop a suitable generalization of the Schwinger-Karplus identity (Lemma 2).

Lemma 3. Given time-dependent operators $\mathcal{F}_l, \mathcal{A}_l, \dots, \mathcal{F}_1, \mathcal{A}_1, \mathcal{F}_0$, the following identity holds:

$$\begin{aligned} &\sum_{m_0, \dots, m_l=0}^{\infty} \mathcal{I}_t^{(l+\sum m_l)}(\mathcal{F}_l \times^{m_l}, \mathcal{A}_l, \dots, \mathcal{F}_1 \times^{m_1}, \mathcal{A}_1, \mathcal{F}_0 \times^{m_0}) \\ &= \int_0^t ds_l \int_0^{s_l} ds_{l-1} \cdots \int_0^{s_2} ds_1 \mathcal{U}^{(l)}(t, s_l) \\ &\quad \times \mathcal{A}_l(s_l) \mathcal{U}^{(l-1)}(s_l, s_{l-1}) \mathcal{A}_{l-1}(s_{l-1}) \cdots \\ &\quad \times \mathcal{U}^{(1)}(s_2, s_1) \mathcal{A}_1(s_1) \mathcal{U}^{(0)}(s_1, 0). \end{aligned} \quad (59)$$

Here

$$\mathcal{U}^{(k)}(t, t_0) := \mathcal{T} \exp \left(\int_{t_0}^t ds \mathcal{F}_k(s) \right) \quad (60)$$

with $\mathcal{T}$ standing for time-ordering.

In the time-dependent form, the expansion is arguably more transparent than the time-independent form. For convenience, given any operator $\mathcal{F}$, define

$$\mathcal{F}^{(n)}(t, t_0) := \int_{t_0}^t ds_n \int_{t_0}^{s_n} ds_{n-1} \cdots \int_{t_0}^{s_2} ds_1 \mathcal{F}(s_n) \cdots \mathcal{F}(s_1), \quad (61)$$

where $t_0 = 0$ is implied if the second argument is omitted.

Proof of Lemma 3. Consider the Dyson series expansion

$$\mathcal{U}^{(k)}(t, t_0) = \mathcal{T} \exp\left(\int_{t_0}^t ds \mathcal{F}_k(s)\right) = \sum_{n=0}^{\infty} \mathcal{F}_k^{(n)}(t, t_0). \quad (62)$$

Examining the left-hand side of the equation above, we can play with the time-ordering to put it in a form resembling this:

$$\begin{aligned} & \sum_{m_0, \dots, m_l=0}^{\infty} \mathcal{I}_l^{(l+\sum m_l)}(\mathcal{F}_l^{\times m_l}, \mathcal{A}_l, \dots, \mathcal{F}_1^{\times m_1}, \mathcal{A}_1, \mathcal{F}_0^{\times m_0}) \\ &= \int_0^t ds_l \cdots \int_0^{s_2} ds_1 \left(\sum_{m_l=0}^{\infty} \mathcal{F}_l^{(m_l)}(t, s_l) \right) \mathcal{A}_l(s_l) \cdots \left(\sum_{m_1=0}^{\infty} \mathcal{F}_1^{(m_1)}(s_2, s_1) \right) \mathcal{A}_1(s_1) \left(\sum_{m_0=0}^{\infty} \mathcal{F}_0^{(m_0)}(s_1, 0) \right) \\ &= \int_0^t ds_l \cdots \int_0^{s_2} ds_1 \mathcal{U}^{(l)}(t, s_l) \mathcal{A}_l(s_l) \cdots \mathcal{U}^{(1)}(s_2, s_1) \mathcal{A}_1(s_1) \mathcal{U}^{(0)}(s_1, 0). \end{aligned} \quad (63)$$

The second line follows by repeatedly applying

$$\begin{aligned} & \int_0^{s_k} ds'_k \int_0^{s'_k} ds_{k-1} \int_0^{s_{k-1}} ds'_{k-1} [\dots] \\ &= \int_0^{s_k} ds'_k \int_0^{s'_k} ds'_{k-1} \int_{s'_{k-1}}^{s'_k} ds_k [\dots] \end{aligned} \quad (64)$$

to interchange the order of the integrals. $\blacksquare$

From here, one recognizes that the manipulations proceed exactly as in the time-independent case. Fix an equivalence class of causal forests $[T]$ with an irreducible skeleton Λ . Then

$$\begin{aligned} & \prod_i \text{ad}_{\mathcal{O}_i} \sum_{M: T(M) \in [T]} \int_0^t ds_n \cdots \\ & \times \int_0^{s_2} ds_1 \mathcal{L}_{X_n}(s_l) \cdots \mathcal{L}_{X_1}(s_1) |A_R\rangle \\ &= \prod_i \text{ad}_{\mathcal{O}_i} \sum_{\vec{m}} \int_0^t ds_l \cdots \int_0^{s_2} ds_1 \mathcal{L}_l^{(m_l)}(t, s_l) \mathcal{L}_{\Lambda_l}(s_l) \\ & \cdots \mathcal{L}_1^{(m_1)}(s_2, s_1) \mathcal{L}_{\Lambda_1}(s_1) \mathcal{L}_0^{(m_0)}(s_1) |A_R\rangle \\ &= \prod_i \text{ad}_{\mathcal{O}_i} \int_0^t ds_l \cdots \int_0^{s_2} ds_1 \mathcal{U}^{(l)}(t, s_l) \mathcal{L}_{\Lambda_l}(s_l) \\ & \cdots \mathcal{U}^{(1)}(s_2, s_1) \mathcal{L}_{\Lambda_1}(s_1) \mathcal{U}^{(0)}(s_1) |A_R\rangle. \end{aligned} \quad (65)$$

Then taking norms, we have

$$\begin{aligned} & \left\| \prod_{i=1}^m \text{ad}_{\mathcal{O}_i} \sum_{M: T(M) \in [T]} \int_0^t ds_l \cdots \int_0^{s_2} ds_1 \mathcal{L}_{X_n}(s_l) \right. \\ & \quad \left. \cdots \mathcal{L}_{X_1}(s_1) |A_R\rangle \right\| \\ & \leq 2^m \int_0^t ds_l \cdots \int_0^{s_2} ds_1 \|\mathcal{L}_{\Lambda_l}(s_l)\| \cdots \|\mathcal{L}_{\Lambda_1}(s_1)\| \\ & \leq 2^m \frac{(2t)^l}{l!} \prod_{X \in \Lambda} \sup_{s \leq t} \|H_X(s)\|, \end{aligned} \quad (66)$$

which appropriately generalizes Eq. (37).

V. HAMILTONIANS WITH EXPONENTIAL TAILS

For some of our applications, it is important to extend our bounds to Hamiltonians with *exponentially decaying* interactions. The main issue is that we are no longer able to write each causal tree as a disjoint union of irreducible paths between region R and each S_i . In this section, we describe how to overcome this difficulty to generalize many of the results of the previous section to interactions with exponential tails. These results will also apply to local Hamiltonians, although they are sometimes weaker than the strongest bounds in the case of local interactions, and the proof methods are significantly more nuanced [53].

A. Review of LR bounds for quasi-local Hamiltonians

In the previous section regarding strictly local interactions, we defined locality via the factor graph. With the introduction of exponentially decaying interactions, this definition is no longer appropriate, because it is the strength of the interactions rather than their support which defines locality. We can still apply the irreducible path bounds from the previous section, but the degree of the factor graph is unbounded, and more care must be taken to correctly sum up the weights of all the potential paths.

In this section, our setting will be qudits on the sites of a graph $G = (V, E)$ with finite degree Δ . Unlike in the previous sections, we use the distance $d(x, y)$ to refer to the geodesic distance provided by this underlying graph G . The volume of connected sets in this graph will then constrain the weights of edges on the factor graph.

Definition 9. We say that a Hamiltonian H is quasi-local if

$$H = \sum_{S \subseteq V} H_S \text{ such that } S \text{ is connected and } \|H_S\| \leq h e^{-\kappa|S|} \quad (67)$$

for some constants h and κ . A subset of the graph is connected if, for each $u \in S$, there is some $v \in S$ such that $d(u, v) = 1$. Here H_S does not need to act non-trivially on every site in S , but decomposition (67) may only exist if most terms in H are associated with the smallest possible set S . In this paper, for convenience, we restrict to

$$\kappa > \log(\Delta) + 1. \quad (68)$$

The aim of this section is to generalize the following conventional LR bound in the same way we generalized bounds for local systems in the previous section.

Proposition 2 (Theorem 3.7 of Ref. [2]). Let H be quasi-local. Suppose that O is supported on a set $S \subset B$ and A is supported on a set R which is disjoint from but adjacent to B as in Fig. 2. Then there exist constants $c_{\text{LR}}, \mu, v_{\text{LR}}$ determined by h, κ, Δ such that

$$\frac{1}{2} \| [O, A(t)] \| \leq c_{\text{LR}} |\partial B| |\partial S| e^{-\mu d(R, S)} (e^{\mu v_{\text{LR}} t} - 1). \quad (69)$$

In the case of strictly local Hamiltonians, a path needed a minimum length, measured on the factor graph, to couple two regions. Since we can no longer do this for quasi-local Hamiltonians, we have to use a different strategy.

Conventional LR bounds follow from a slightly less restrictive condition than our criterion for quasi-locality [2]. A more standard definition of quasi-local in the literature is to require that there exist h, μ such that, for any two sites $u, v \in V$, we have $\sum_{X \ni u, v} \|H_X\| \leq h e^{-\mu d(u, v)}$.

Intuitively, it requires that Hamiltonian terms decay exponentially in their *radius* rather than their entire *support*. We can show that this weaker condition follows readily from our definition using the following proposition.

Proposition 3 (Adapted from Ref. [54]). Given a vertex $u \in V$ on a graph of maximal degree Δ , there are at most $(\Delta e)^m$ connected subsets S with $u \in S$ and $|S| = m$. Furthermore, these subsets can be enumerated classically in runtime $m \Delta (\Delta e)^m$.

This is obtained directly by Proposition 4.6 and Sec. 4.3 of Ref. [54].

Lemma 4. Let $u, v \in V$ be arbitrary. Then, for any $\alpha > 0$, there exist constants μ, h' such that the following holds:

$$\sum_{X \ni u, v} \|H_X\| \leq \frac{h' e^{-\mu d(u, v)}}{d(u, v)^\alpha}. \quad (70)$$

Proof. We begin by overbounding the sum with every set which includes u and has at least $d(u, v)$ members, as required of a connected set containing both u and v :

$$\begin{aligned} \sum_{X \ni u, v} \|H_X\| &\leq \sum_{m \geq d(u, v)} \sum_{\substack{X \ni u \\ |X| = m}} \|H_X\| \\ &\leq h \sum_{m \geq d(u, v)} (e\Delta)^m e^{-\kappa m} \\ &= \frac{h e^{-[\kappa + \log(e\Delta)]d(u, v)}}{1 - (e\Delta)} e^{-\kappa} \\ &=: C e^{-\mu' d(u, v)}. \end{aligned} \quad (71)$$

We have bounded the number of connected subsets X containing u of size m with $(e\Delta)^m$ by Proposition 3. Then the result is a geometric series, which converges due to Eq. (68). Now let $\alpha > 0$. Since an exponential decays much more rapidly than a power-law, we can find μ dependent on α such that

$$C e^{-\mu' d(u, v)} \leq h' \frac{e^{-\mu d(u, v)}}{d(u, v)^\alpha}, \quad (72)$$

where h' is another constant dependent on α and C , but not dependent on $d(u, v)$. This is the advertised result. ■

The reason for requiring the power-law in the denominator is that we need a way to iteratively apply this bound in order to constrain the weight of irreducible paths. This is called the *reproducing* property [55], and is encapsulated by the following lemma.

Lemma 5. For any $\alpha > d$, there exists a constant K independent of u, v such that

$$\sum_{k \neq u, v} \frac{e^{-\mu d(u, k)}}{d(u, k)^\alpha} \frac{e^{-\mu d(k, v)}}{d(k, v)^\alpha} \leq K \frac{e^{-\mu d(u, v)}}{d(u, v)^\alpha}. \quad (73)$$

This makes the function $G_\alpha(l) := e^{-\mu l} l^{-\alpha}$ with $\alpha > d$ reproducing for the given lattice. Bounding the weight of edges in the factor graph by a reproducing function of distance on the underlying graph is exactly the criterion which allows us to easily bound the sum of weights of irreducible paths through the factor graph by a function of the distance between the start and end points. We illustrate how this property is deployed to constrain the locality of dynamics by reproducing the proof of Proposition 2 found in Ref. [2].

Proof of Proposition 2. Applying the result from Theorem 6, the commutator is bounded by

$$\frac{1}{2} \| [O, A(t)] \| \leq \sum_{m=1}^{\infty} \frac{(2t)^m}{m!} \sum_{u \in \partial B, v \in \partial S} \sum_{\substack{\Gamma(u \rightarrow v) \\ |\Gamma|=m}} w(\Gamma). \quad (74)$$

Now we need to bound the sum over weighted subsets. First fix some $u \in \partial B, v \in \partial S$. Then consider the irreducible path $\Gamma = (X_1, X_2, \dots, X_m)$. We note that the sets X_i must form a connected path, i.e. $X_k \cap X_{k-1} \neq \emptyset$, so we can overbound the sum by summing over all such connected multisets. Furthermore, we can pick one point v_k in each intersection $X_k \cap X_{k-1}$. Since this is true for each such connected multiset, we can further overbound the sum by considering all such sequences of vertices $\vec{v} = (u, v_2, \dots, v_{m-1}, v)$, where $v_2, \dots, v_{m-1}$ are arbitrary, and sum over $\vec{X} = X_1, \dots, X_m$ satisfying $v_k \in X_k \cap X_{k-1}$:

$$\begin{aligned} \sum_{\substack{\Gamma(u \rightarrow v) \\ |\Gamma|=m}} w(\Gamma) &\leq \sum_{v_2, \dots, v_{m-1}} \sum_{X_1 \ni u, v_2} \|H_{X_1}\| \cdots \sum_{X_m \ni v_{m-1}, v} \|H_{X_m}\| \\ &\leq h^m \sum_{\vec{v}} \frac{e^{-\mu d(u, v_2)}}{d(u, v_2)^\alpha} \cdots \frac{e^{-\mu d(v_{m-1}, v)}}{d(v_{m-1}, v)^\alpha} \\ &\leq \frac{(h'K)^m e^{-\mu d(u, v)}}{K d(u, v)^\alpha}. \end{aligned} \quad (75)$$

Thus we have the bound

$$\sum_{u \in \partial R, v \in \partial S} \sum_{\substack{\Gamma(u \rightarrow v) \\ |\Gamma|=m}} w(\Gamma) \leq |\partial B| |\partial S| \frac{(h'K)^m e^{-\mu d(R, S)}}{K d(R, S)^\alpha}. \quad (76)$$

Inserting this into Eq. (74), we find that

$$\begin{aligned} C_S^R(t) &\leq |\partial B| |\partial S| \frac{e^{-\mu d(u, v)}}{K d(u, v)^\alpha} \sum_{m=1}^{\infty} \frac{(2t)^m}{m!} (h'K)^m \\ &= |\partial B| |\partial S| \frac{e^{-\mu d(u, v)}}{K d(u, v)^\alpha} [\exp(2tKh') - 1] \\ &= c |\partial B| |\partial S| e^{-\mu d(R, S)} (e^{\mu v_{LR} t} - 1), \end{aligned} \quad (77)$$

where we have introduced the constants $c = K^{-1}$ and $v_{LR} = 2Kh'\mu^{-1}$. ■

B. Extending to nested commutators: $m = 2$ case

Now we can extend this result to nested commutators. For an illustrative example, we first consider the case where there are only two regions B_1, B_2 . In this section, we illustrate the strategy for the general case by proving the following proposition.

Proposition 4. For operators $\mathcal{O}_1, \mathcal{O}_2$ supported on S_1, S_2 , respectively, and A supported in the complement of $B_1 \cup B_2$, with $\mathcal{O}(1)$ constants μ, v_{LR}, h, h' as defined in the previous section, we have

$$\begin{aligned} \frac{1}{4} \| [\mathcal{O}_1, [\mathcal{O}_2, A(t)]] \| &\leq \mu v_{LR} t (\mu v_{LR} t + e^{-2(\mu\chi - \kappa)}) \\ &\quad \prod_{i=1,2} b |\partial B_i| |\partial S_i| e^{\mu(v_{LR} t - d(S_i, B_i))}, \end{aligned} \quad (78)$$

where $b := h'/h$ and $\chi \leq d(B_1, B_2)/2$.

Before proceeding with the proof, we introduce a few definitions. The most important distinction with the single-region case in the previous section is that we now have factors which intersect both B_1 and B_2 , so we can no longer immediately factor the paths into products within each B_i , as in Corollary 2. However, we can reduce to the decoupled case by sorting the equivalence classes in the following way.

Definition 10. Let $[\Gamma] \in \mathcal{S}$ be an equivalence class with irreducible paths Γ_1 and Γ_2 connecting S_1 and S_2 to R , respectively. We say that Γ couples B_1 and B_2 if $\Gamma_1 \cap \Gamma_2 \neq \emptyset$.

Lemma 6. If Γ couples B_1 with B_2 then $\Gamma_1 \cap \Gamma_2 = \{X\}$, where X is a factor. Here X is the first element of both Γ_1 and Γ_2 , and the only element of either path that intersects R .

Proof. Let $X \in \Gamma_1 \cap \Gamma_2$. The first element of an irreducible path is the only element of said irreducible path that can intersect with R by construction. Since X is

required to be connected, it must intersect with R . Therefore it must be the first element in the irreducible path Γ_1 and similarly for Γ_2 . If $Y \in \Gamma_1 \cap \Gamma_2$ then X and Y are both the first elements of both paths, so $X = Y$. ■

This shows in particular that the irreducible skeletons in the coupled case can be reduced to the uncoupled case by removing the first term.

Proof of Proposition 4. We now divide the equivalence classes into uncoupled paths $\mathcal{S}_d$ and coupled paths $\mathcal{S}_c$. Then the commutator bound breaks into a sum:

$$\frac{1}{4} \|[\mathcal{O}_1, [\mathcal{O}_2, A(t)]]\| \leq \sum_{\Gamma \in \mathcal{S}_d} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} w(\Gamma) + \sum_{\Gamma \in \mathcal{S}_c} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} w(\Gamma). \quad (79)$$

By Corollary 2, we can bound the first term in this equation by

$$\sum_{\Gamma \in \mathcal{S}_d} \frac{(2|t|)^{|\Gamma|}}{|\Gamma|!} w(\Gamma) \leq \prod_{i=1,2} c|\partial B_i| |\partial S_i| e^{-\mu d(S_i, B_i)} (e^{\mu v_{\text{LRF}}} - 1). \quad (80)$$

We are left with just the second term to bound. For convenience, let

$$A(X \rightarrow Y; l) := \{(U_1, \dots, U_l) : U_1 \cap X \neq \emptyset, U_i \cap U_{i-1} \neq \emptyset, U_l \cap Y \neq \emptyset\}, \quad (81)$$

where $X, Y, U_i \subseteq V$. Although not all $X, Y \in F$, we bound the sum by summing over all *possible* subsets. Accordingly, given $\vec{U} = (U_1, \dots, U_l)$ where $U_i \subseteq V$, we define

$$\prod_i \|H_{U_i}\| \delta_{U_i \in F} \leq \prod_i h e^{-\kappa |U_i|} := W(\vec{U}). \quad (82)$$

Then we write as a shorthand for our bound on the single commutator

$$\sum_{\substack{\Gamma(R \rightarrow S_i) \\ |\Gamma|=l}} w(\Gamma) \leq \sum_{\vec{U} \in A(R \rightarrow S_i; l)} W(\vec{U}) \leq |\partial B_i| |\partial S_i| \frac{(h'K)^l e^{-\mu d(R, S_i)}}{K d(R, S_i)^\alpha} := \frac{(h'K)^l}{K} Q_i. \quad (83)$$

Note that in the first inequality, we have relaxed the bound by summing over non-irreducible (e.g. backtracking) paths $\vec{U}$ as well. By Proposition 6, we can break each path up into the first factor X which couples the two balls, and sequences of factors $\vec{U}, \vec{V}$ which then couple X to S_1 and S_2 , respectively. Then we account for all the ways that the elements of $\vec{U}, \vec{V}$ can be permuted without changing the ordering, which contributes a factor of $(l_1 + l_2)!/(l_1! l_2!)$:

$$\begin{aligned} \sum_{\substack{\Gamma \in \mathcal{S}_c \\ |\Gamma|=l}} w(\Gamma) &\leq \sum_{l_1+l_2=l-1} \sum_{\substack{P: P \cap B_1 \neq \emptyset \\ P \cap B_2 \neq \emptyset}} W(P) \sum_{\substack{\vec{U} \in A(P \rightarrow S_1; l_1) \\ \vec{V} \in A(P \rightarrow S_2; l_2)}} \sum_{\text{ordering}\{U_i\} \cup \{V_i\}} W(\vec{U} \oplus \vec{V}) \\ &= \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1! l_2!} \sum_{\substack{P: P \cap B_1 \neq \emptyset \\ P \cap B_2 \neq \emptyset}} W(P) \sum_{\vec{U} \in A(P \rightarrow S_1; l_1)} W(\vec{U}) \sum_{\vec{V} \in A(P \rightarrow S_2; l_2)} W(\vec{V}). \end{aligned} \quad (84)$$

Here we used “ $\oplus$ ” to emphasize that $\vec{U}$ and $\vec{V}$ are concatenated together as sequences of sets. Now we recognize that if P intersects both B_1 and B_2 then we can pick $u \in \partial B_1 \cap P$ and $v \in \partial B_2 \cap P$ and write $P = X \setminus \{u, v\} \sqcup Y \sqcup Z$, where $Y = P \cap B_1$, $Z = P \cap B_2$, and $X = P \cap R \cup \{u, v\}$. Then we can bound the sum over such P by summing over X, Y, Z satisfying $u, v \in X, u \in Y, v \in Z$. This is illustrated in Fig. 6. We have

$$\begin{aligned} \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1! l_2!} \sum_{\substack{P: P \cap B_1 \neq \emptyset \\ P \cap B_2 \neq \emptyset}} W(P) \sum_{\vec{U} \in A(P \rightarrow S_1; l_1)} W(\vec{U}) \sum_{\vec{V} \in A(P \rightarrow S_2; l_2)} W(\vec{V}) \\ \leq \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1! l_2!} \sum_{\substack{u \in \partial B_1 \\ v \in \partial B_2}} \sum_{X \ni u, v} \sum_{Y \cap X = u} \sum_{Z \cap X = v} W(X \cup Y \cup Z) \sum_{\vec{U} \in A(Y \rightarrow S_1; l_1)} W(\vec{U}) \sum_{\vec{V} \in A(Z \rightarrow S_2; l_2)} W(\vec{V}) \\ \leq \frac{e^{2\kappa}}{h^2} \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1! l_2!} \sum_{\substack{u \in \partial B_1 \\ v \in \partial B_2}} \sum_{X \ni u, v} W(X) \sum_{\vec{U} \in A(\{u\} \rightarrow S_1; l_1+1)} W(\vec{U}) \sum_{\vec{V} \in A(\{v\} \rightarrow S_2; l_2+1)} W(\vec{V}). \end{aligned} \quad (85)$$

We have abused notation slightly in writing

$$W(X \cup Y \cup Z) = W(X \setminus \{u, v\} \sqcup Y \sqcup Z) = \frac{e^{2\kappa}}{h^2} W(X) W(Y) W(Z). \quad (86)$$

Applying Lemma 4, we bound $\sum_{X \ni u, v} w(X) \leq h' e^{-\mu d(u, v)} \leq h' e^{-2\mu\chi}$. The remaining sums over paths are exactly the sums appearing in Lemma 5:

$$\begin{aligned} & \frac{e^{2\kappa}}{h^2} \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1!l_2!} \sum_{\substack{u \in \partial B_1 \\ v \in \partial B_2}} \sum_{X \ni u, v} w(X) \sum_{\vec{U} \in A(\{u\} \rightarrow S_1; l_1+1)} w(\vec{U}) \sum_{\vec{V} \in A(\{v\} \rightarrow S_2; l_2+1)} w(\vec{V}) \\ & \leq h' \frac{1}{h^2} e^{2\kappa} e^{-2\mu\chi} \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1!l_2!} \frac{(h'K)^{l_1+1}}{K} Q_1 \frac{(h'K)^{l_2+1}}{K} Q_2 \\ & = h' \left(\frac{h'}{h}\right)^2 e^{2\kappa} e^{-2\mu\chi} Q_1 Q_2 \sum_{l_1+l_2=l-1} \frac{(l-1)!}{l_1!l_2!} (h'K)^{l_1+l_2}. \end{aligned} \quad (87)$$

Now we can plug this bound into the left-hand side of Eq. (90):

$$\sum_{\Gamma \in \mathcal{S}} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} W(\Gamma) \leq h' \sum_{l=1}^{\infty} \sum_{l_1+l_2=l-1} \frac{(2t)^l}{l!} \frac{(l-1)!}{l_1!l_2!} (h'K)^{l_1} (h'K)^{l_2} \prod_{i=1,2} b e^{-\mu\chi+\kappa} Q_i. \quad (88)$$

We can similarly rewrite this sum in terms of sums over l_1, l_2 alone, i.e.

$$\begin{aligned} & \sum_{l=1}^{\infty} \sum_{l_1+l_2=l-1} \frac{(2t)^l}{l!} \frac{(l-1)!}{l_1!l_2!} (h'K)^{l_1} (h'K)^{l_2} = 2t \sum_{l_1=0}^{\infty} \sum_{l_2=0}^{\infty} \frac{(2h'Kt)^{l_1}}{l_1!} \frac{(2h'Kt)^{l_2}}{l_2!} \frac{1}{(l_1+l_2+1)} \\ & \leq 2t \left(\sum_{l_1=0}^{\infty} \frac{(2h'Kt)^{l_1}}{l_1!} \right) \left(\sum_{l_2=0}^{\infty} \frac{(2h'Kt)^{l_2}}{l_2!} \right) = 2te^{2\mu v_{LR}t}, \end{aligned} \quad (89)$$

where we used $v_{LR} = 2Kh'/\mu$. Now with $K > 1$, our bound simplifies to

$$\begin{aligned} & \sum_{\Gamma \in \mathcal{S}_c} \frac{(2t)^{|\Gamma|}}{|\Gamma|!} w(\Gamma) \leq \mu v_{LR} t e^{-2\mu\chi+2\kappa} \\ & \prod_{i=1,2} b |\partial B_i| |\partial S_i| e^{\mu(v_{LR}t - d(S_i, B_i))}. \end{aligned} \quad (90)$$

This gives a bound on the second term in Eq. (79). To obtain a simple expression for the sum, we observe that $e^t - 1 \leq te^t$, which follows from a comparison of their Taylor series. Then we have $b > 1$ because $h' > h$, and $c = 1/K < 1$ by construction, so

$$\begin{aligned} & \prod_{i=1,2} c |\partial B_i| |\partial S_i| e^{-\mu d(S_i, B_i)} (e^{\mu v_{LR}t} - 1) \\ & \leq \prod_{i=1,2} b |\partial B_i| |\partial S_i| e^{-\mu d(S_i, B_i)} (\mu v_{LR}t) e^{\mu v_{LR}t}. \end{aligned} \quad (91)$$

Adding this to Eq. (90) gives the desired bound. $\blacksquare$

We remark that, since $\kappa \sim \mu$, if $\mu\chi > \kappa$ then, as $\kappa \rightarrow \infty$, the factor $e^{2(\kappa-\mu\chi)}$ suppresses the contribution of the terms in $\mathcal{S}_c$, and we recover the product bound for the strictly local Hamiltonian.

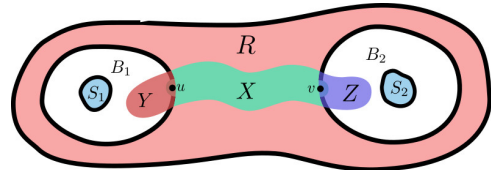


FIG. 6. Example of a long-range coupling term supported on $P = X \cup Y \cup Z$. Since P must intersect the boundary of both B_1 and B_2 by the connectedness requirement, we pick two “anchor” points u, v on the respective boundaries of these regions so that $u \in X, Y$ and $v \in X, Z$.

C. Nested commutators: the general case

Now the goal is to generalize this proof to arbitrary m . Most of the manipulations are similar to those from the previous section. The problem now is that we need to account for long-range coupling between all possible subsets of the regions. If we were to naively sum the long-range couplings, the number of such subsets would grow much faster than the exponential suppression. In order to overcome this difficulty, we need to capture the fact that there are only very few small connected subsets coupling the regions, and the contribution of the large connected subsets are suppressed exponentially in the distance between these subsets. In order to formalize this notion, we note that in d spatial dimensions, we can tile the plane with Euclidean boxes such that every box is adjacent to $3^d - 1$ boxes. We specialize to the case where each box contains at most one region $B_1, \dots, B_m$, and the distance between B_i and any other box is at least χ . This is straightforwardly generalized to any way $\mathbb{R}^d$ can be broken into a course-grained lattice by tiling it with arbitrary shapes.

Definition 11. If $[T] \in \mathcal{S}_c$ is an equivalence class of trees with irreducible paths $\Gamma_1, \dots, \Gamma_m$ such that $\Gamma_1 \cap \Gamma_2 \cap \dots \cap \Gamma_k \neq \emptyset$ and $\Gamma_i \cap \Gamma_j = \emptyset$ for $i \leq k$ and $j > k$, then we say that $[T]$ couples $B_1, \dots, B_k$.

Proposition 5. Let T be a causal tree coupling $B_1, \dots, B_k$. Let $\Gamma_1, \dots, \Gamma_k$ be the irreducible paths coupling $S_1, \dots, S_k$ to R . Then $\Gamma_1 \cap \dots \cap \Gamma_k = \{X\}$, where X intersects with $B_1, \dots, B_k$, and X is the first element of $\Gamma_1, \dots, \Gamma_k$. Furthermore, all elements of $\Gamma_i \setminus \{X\}$ are contained within B_i for each $i \leq k$.

Proof. By the algorithm used to construct causal trees, the first element in each irreducible path is the only one

to intersect R . Since X must be connected and couples the regions, it must intersect with R . Therefore, removing X from Γ_i creates a path that lies entirely within B_i . ■

Definition 12. We call a connected set of boxes a *connected cluster*. For any $k \leq m$, a connected cluster containing $B_1, \dots, B_k$ which has minimal volume is called a *minimal cluster* corresponding to $B_1, \dots, B_k$. An example is illustrated in Fig. 7.

We then proceed by induction on m . If a number of regions are coupled together as shown in Fig. 7 by a long-range term in a particular equivalence class then, by applying Corollary 2, we can factorize the system into two disjoint subsets of smaller size, which will establish the recursion. We first derive the contribution from an equivalence class which couples k regions using techniques from the previous section.

Lemma 7. Suppose that $\mathcal{S}_c$ is the set of equivalence classes coupling $B_1, B_2, \dots, B_k$ with minimal cluster of order $M_{\min}$ and irreducible skeleton of length l . Then we can carry out the weighted sum over $\mathcal{S}_c$ with the bound

$$\sum_{l=1}^{\infty} \sum_{\substack{\Lambda \in \mathcal{S}_c \\ |\Lambda|=l}} \frac{(2t)^l}{l!} w(\Lambda) \leq (\mu v_{\text{LR}} t) e^{-\mu \chi M_{\min}} \times \prod_i e^{\kappa} b |\partial B_i| |\partial S_i| e^{\mu(v_{\text{LR}} l - d(S_i, B_i))}. \quad (92)$$

Proof. Fix an equivalence class $\mathcal{S}_c$ that couples $B_1, \dots, B_k$. As in the proof of Proposition 4, we can bound the sum by

$$\begin{aligned} \sum_{\substack{\Lambda \in \mathcal{S}_c \\ |\Lambda|=l}} w(\Lambda) &\leq \sum_k \sum_{l_k=l-1} \frac{(l-1)!}{l_1! \dots l_k!} \sum_{X \cap B_1 \dots X \cap B_k \neq \emptyset} \|H_X\| \sum_{\tilde{U}_1 \in A(X \rightarrow S_1; l_1)} W(\tilde{U}_1) \dots \sum_{\tilde{U}_k \in A(X \rightarrow S_k; l_k)} W(\tilde{U}_k) \\ &\leq \sum_k \sum_{l_k=l-1} \frac{(l-1)!}{l_1! \dots l_k!} \sum_{u_1 \in \partial B_1 \dots u_k \in \partial B_k} \sum_{X \ni u_1, \dots, u_k} W(X \setminus \{u_1, \dots, u_k\}) \sum_{\tilde{U}_1 \in A(\{u_1\} \rightarrow S_1; l_1)} W(\tilde{U}_1) \dots \sum_{\tilde{U}_k \in A(\{u_k\} \rightarrow S_k; l_k)} W(\tilde{U}_k) \\ &\leq h' \left(\frac{h'}{h} \right)^k e^{\kappa k} e^{-\mu M_{\min} \chi} Q_1 \dots Q_k \sum_{\sum_i l_i=l-1} \frac{(l-1)!}{l_1! \dots l_k!} (h'K)^{l-1}. \end{aligned} \quad (93)$$

We have skipped a few steps because the manipulations are the same as in Proposition 4. The significant difference is that each path needs to pass through a connected cluster, and the minimal cluster captures the number of boxes that

it must either pass through or connect to the ball at the center. Therefore if the minimal cluster is of size $M_{\min}$, the weight of this path is bounded by $W(X \setminus \{u_1, \dots, u_k\}) \leq h e^{\kappa k} e^{-\mu \kappa M_{\min}}$, from which the bound on the summation

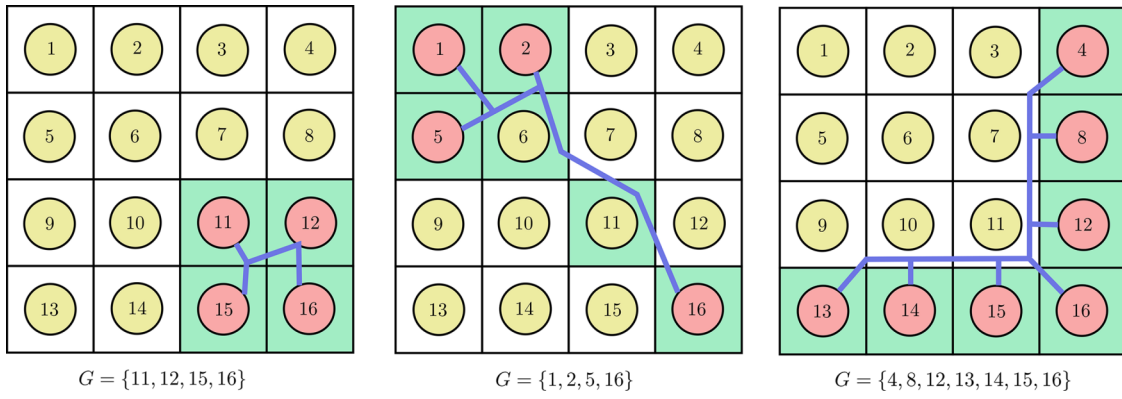


FIG. 7. Here $\mathbb{R}^2$ is covered with Euclidean boxes, and at most one region of interest B_i is situated within each box with a distance at most χ from the edge of the box. Subset G is a subset of the regions representing the ways region 16 can be coupled by a single term in a given irreducible path to the rest of the system, as shown in red. This coupling term is illustrated in blue. Since the coupling term is connected, it must pass through boxes that form a connected cluster; G itself does not need to be connected. The green boxes show a minimal connected cluster containing G , which captures the exponential suppression of the terms coupling together the regions. We denote the number of green boxes by $M_{\min}(G)$. Our formalism accounts for terms like the one shown on the left that couple nearby balls as well as ones like the two shown on the right, which couple regions far away.

of all possible subsets containing $u_1 \in \partial B_1, \dots, u_k \in \partial B_k$ follows as in Lemma 4. With this, we have

$$\begin{aligned} & \sum_{l=1}^{\infty} \sum_{\substack{\Lambda \in \mathcal{S} \\ |\Lambda|=l}} \frac{(2t)^l}{l!} w(\Lambda) \\ & \leq h' e^{-\mu \chi M_{\min}} e^{k\kappa} \left(\prod_i b |\partial B_i| |\partial S_i| e^{-\mu d(S_i, B_i)} \right) \\ & \times \sum_{l=1}^{\infty} \frac{(2t)^l}{l!} \left(\sum_{l_1 + \dots + l_k = l-1} \frac{(l-1)!}{l_1! l_2! \dots l_k!} \prod_i (h' K)^{l_i} \right). \end{aligned} \quad (94)$$

Then we can bound the compound sum with a product:

$$\begin{aligned} & \sum_{l=1}^{\infty} \frac{(2t)^l}{l!} \sum_{l_1 + \dots + l_k = l-1} \frac{(l-1)!}{l_1! l_2! \dots l_k!} \prod_i (h' K)^{l_i} \\ & = 2th' \sum_{l_1 + \dots + l_k = l-1} \frac{(2t)^{\sum_i l_i}}{l_1! \dots l_k!} \frac{1}{1 + \sum_i l_i} \prod_i (h' K)^{l_i} \\ & \leq \mu v_{\text{LR}} t \left(\sum_{l=0}^{\infty} \frac{(2h' K t)^l}{l!} \right)^k \\ & \leq \mu v_{\text{LR}} t e^{k\mu v_{\text{LR}} t}. \end{aligned} \quad (95)$$

Inserting this into Eq. (94) gives the desired result. $\blacksquare$

Theorem 8. Consider a quasi-local interaction defined on a graph in d dimensions. Suppose that we tile the plane with Euclidean boxes, where each box has $3^d - 1$ adjacent neighbors. Assume that $B_1, \dots, B_m$ are regions, each contained within a separate box such that the distance between B_i and any other box is at least χ , where $\mu \chi > \max(\log(2) + d \log(3) + 2, \kappa)$, and $S_i \subset B_i$. Then the nested commutator bound

$$\begin{aligned} & \frac{1}{2^m} \| [\mathcal{O}_m, [\dots, [\mathcal{O}_1, A(t)] \dots]] \| \leq \mu v t (e^{-\gamma + \kappa} \\ & + \mu v t)^{m-1} \prod_{i=1}^m |\partial B_i| |\partial S_i| b e^{\mu(v_{\text{LR}} t - d(S_i, R))} \end{aligned} \quad (96)$$

holds, where A is supported in R which is the complementary region to $\bigcup_i B_i$, $\mathcal{O}_i$ is supported within S_i , $\gamma = \mu \chi - \log(2e3^d) > 1$, and $v = 4e3^d v_{\text{LR}}$.

Proof. Let $B_1, \dots, B_m$ be defined as in the theorem statement. We then induct on m . The base case was proved in Sec. V A. We then consider an additional region B_{m+1} . We sort the equivalence classes by considering all the possible ways B_{m+1} can be coupled to $B_1, \dots, B_m$. For any $k \leq m$, let $\Lambda(\{i_1, \dots, i_k\})$ denote the set of irreducible skeletons connecting $S_{i_1}, \dots, S_{i_k}$ to R and let $\Gamma(\{i_1, \dots, i_k\})$ denote the irreducible skeletons coupling $i_1, \dots, i_k$. As we have shown in Eq. (40), we can bound the nested commutator

with

$$\begin{aligned}
& \frac{1}{2^m} \| [\mathcal{O}_{m+1}, [\dots, [\mathcal{O}_1, A(t)] \dots]] \| \\
& \leq \sum_{\Lambda \in \Lambda(1, \dots, m)} \frac{(2|t|)^{|\Lambda|}}{|\Lambda|!} w(\Lambda) \\
& \leq \sum_{\substack{G \subseteq \{1, \dots, m, m+1\} \\ m+1 \in G}} \left(\sum_{\Gamma \in \Gamma(G)} \frac{(2|t|)^{|\Gamma|}}{|\Gamma|!} w(\Gamma) \right) \\
& \times \left(\sum_{\Lambda \in \Lambda(G^c)} \frac{(2|t|)^{|\Lambda|}}{|\Lambda|!} w(\Lambda) \right). \quad (97)
\end{aligned}$$

The decomposition above is illustrated in Fig. 7, with G represented by the red circles (all connected to B_{m+1}) and G^c being the yellow circles (not connected to B_{m+1} , not necessarily all connected to each other). Let $C(G) := \sum_{\Lambda \in \Lambda(G)} ((2|t|)^{|\Lambda|}/|\Lambda|!) w(\Lambda)$ with $C(\emptyset) = 1$. Applying Lemma 7, we have a bound for the first term:

$$\begin{aligned}
& C(\{1, \dots, m, m+1\}) \\
& \leq \sum_{\substack{G \subseteq \{1, \dots, m, m+1\} \\ m+1 \in G}} \mu v_{\text{LR}} t e^{-\mu \chi (M_{\min}(G)-1)} C(G^c) \\
& \times \prod_{i \in G} |\partial B_i| |\partial S_i| b e^{\kappa(|G|-1)} \exp(\mu(v_{\text{LR}} t - r_i)) \quad (98)
\end{aligned}$$

with $r_i = d(S_i, R)$ for short. Since we assumed that $\mu \chi > \kappa$, and $|G| \leq M_{\min}(G)$, we relaxed $e^{\kappa|G|} e^{-\mu \chi M_{\min}(G)}$ to $e^{\kappa(|G|-1)} e^{-\mu \chi (M_{\min}(G)-1)}$. This choice greatly simplifies the induction because it makes Lemma 7 consistent with the $m = 1$ case that we proved in Sec. V A, thereby serving as a base case. Then we introduce $D(G)$ such that

$$C(G) = \prod_{i \in G} b |\partial B_i| |\partial S_i| \exp(\mu(v_{\text{LR}} t - r_i)) D(G). \quad (99)$$

Then $D(G)$ satisfies the relationship

$$\begin{aligned}
& D(\{1, \dots, m, m+1\}) \leq (\mu v_{\text{LR}} t) \\
& \times \sum_{\substack{G \subseteq \{1, \dots, m, m+1\} \\ m+1 \in G}} e^{(|G|-1)\kappa} e^{-\mu \chi (M_{\min}(G)-1)} D(G^c). \quad (100)
\end{aligned}$$

Now let

$$Q(k) = \sup_{\substack{G \subseteq \{1, \dots, m, m+1\} \\ |G|=k}} D(G). \quad (101)$$

Note that $Q(m+1) = D(\{1, \dots, m, m+1\})$ and $Q(0) = 1$. This gives us the simpler but weaker bound

$$\begin{aligned}
& Q(m+1) \leq (\mu v_{\text{LR}} t) \\
& \times \sum_{\substack{G \subseteq \{1, \dots, m, m+1\} \\ m+1 \in G}} e^{(|G|-1)\kappa} e^{-\mu \chi (M_{\min}(G)-1)} Q(m+1-|G|). \quad (102)
\end{aligned}$$

We can then bound the sum above by summing instead over connected clusters of boxes and then all possible subsets of those clusters:

$$\begin{aligned}
& \sum_{\substack{G \subseteq \{1, \dots, m, m+1\} \\ G \ni m+1}} e^{(|G|-1)\kappa} e^{-\mu \chi (M_{\min}(G)-1)} Q(m+1-|G|) \\
& = e^{\mu \chi} \sum_{k=1}^{m+1} e^{(k-1)\kappa} \sum_{\substack{M=k \\ \text{connected } X \ni m+1 \\ |X|=M}}^{\infty} \sum_{|X|=M} \\
& \times \sum_{\substack{G \subseteq X \\ |G|=k}} e^{-\mu \chi M} Q(m+1-k) \\
& = e^{\mu \chi} \sum_{k=1}^{m+1} e^{(k-1)\kappa} \sum_{M=k}^{\infty} (e^3)^M \binom{M}{k} \\
& \times e^{-\mu \chi M} Q(m+1-k) \\
& \leq e^{\mu \chi} \sum_{k=1}^{m+1} e^{(k-1)\kappa} \sum_{M=k}^{\infty} e^{-M\gamma} Q(m+1-k) \\
& \leq 4e^{3d} \sum_{k=0}^m e^{-k(\gamma-\kappa)} Q(m-k). \quad (103)
\end{aligned}$$

Here we set $\gamma = \mu \chi - \log(2e^{3d})$ and require that $\gamma > 1$, which we can always ensure by choosing χ large enough. To obtain the third line, we used Proposition 3 to bound the number of clusters, and in the fourth line we bounded $\binom{M}{k} \leq 2^M$. With $Q(0) = 1$ and $Q(1) \leq \mu v_{\text{LR}} t \leq \mu v t$ already proven, for our inductive hypothesis, we assume that, for $0 < k \leq m$,

$$Q(k) \leq \mu v t (\mu v t + e^{-(\gamma-\kappa)})^{k-1}. \quad (104)$$

For convenience, let $t' := \mu vt$ and $a := e^{-(\gamma-\kappa)}$. Then

$$\begin{aligned}
Q(m+1) &\leq t' a^m + t' \sum_{k=1}^m a^{m-k} t' (t' + a)^{k-1} \\
&= t' a^m + \sum_{k=1}^m a^{m-k} \sum_{j=0}^{k-1} \binom{k-1}{j} t'^{j+2} a^{k-1-j} \\
&= t' a^m + \sum_{j=0}^{m-1} a^{m-j-1} t'^{j+2} \sum_{k=j+1}^m \binom{k-1}{j} \\
&= t' a^m + \sum_{j=0}^{m-1} a^{m-j-1} t'^{j+2} \binom{m}{m-j-1} \\
&= t' \sum_{j=0}^m a^{m-j} t'^j \binom{m}{m-j} \\
&= t' (t' + a)^m.
\end{aligned} \tag{105}$$

This completes the induction. We used the “hockey stick” relation $\sum_{m=k}^n \binom{m}{k} = \binom{n+1}{k+1}$ for the binomial coefficients. Unraveling the definition of Q ,

$$\begin{aligned}
&\frac{1}{2^m} \|[\mathcal{O}_m, [\dots, [\mathcal{O}_1, A(t)] \dots]]\| \\
&\leq Q(m) \prod_i b |\partial B_i| |\partial S_i| \exp(\mu(v_{\text{LR}} t - r_i)) \\
&\leq \mu vt (\mu vt + e^{-\gamma+\kappa})^{m-1} \prod_i b |\partial B_i| |\partial S_i| \\
&\quad \times \exp(\mu(v_{\text{LR}} t - r_i)).
\end{aligned} \tag{106}$$

This is the desired result. $\blacksquare$

We remark on some of the properties of the above formula. Firstly, we see that, unlike the strictly local case, this bound is $\Theta(t)$ for small t . This reflects the fact that there are now terms at first order that couple the m regions to R . The term in the sum proportional to t has a prefactor of $e^{-(m-1)(\gamma-\kappa)}$. The $m-1$ instead of m comes from the fact that we needed to relax the suppression in χ slightly to obtain this simple formula, but this reflects the fact that any term which directly couples all m regions is exponentially suppressed in m . For example, expanding $Q(3) = t^3 + 2t^2 a + a^2 t'$, we can identify the first term as decoupled, the second as originating from a coupling of two regions with the binomial coefficient counting the number of such couplings, and the last term representing a coupling between all three. Next, when $\mu\chi \gg \kappa$, $\log(2e3^d)$, the prefactor approaches $(\mu vt)^m$. This can be seen as the decoupled limit, and reflects the fact that one has to go to m th order to connect each S_i to R , because irreducible skeletons Λ with $|\Lambda| < m$ are exponentially suppressed.

This also happens when we take $\kappa \rightarrow 0$, and, since $\mu \sim \kappa$, the requirement that $\mu\chi > \kappa$ is well behaved in this limit for fixed χ . Lastly, we note that this bound is uniform in the number of sites in the system $|V|$, which permits us to pass to the thermodynamic limit.

Although the above bound is stronger, we can also obtain a simplified result that more closely resembles standard LR bounds.

Corollary 6 (Theorem 8 simplified). Under the same assumptions as in Theorem 8, the bound

$$\begin{aligned}
&\frac{1}{2^m} \|[\mathcal{O}_m, [\dots, [\mathcal{O}_1, A(t)] \dots]]\| \\
&\times \leq \prod_{i=1}^m b |\partial B_i| |\partial S_i| e^{\mu(vt - d(S_i, R))}
\end{aligned} \tag{107}$$

holds, where A is supported in R which is the complementary region to $\bigcup_i B_i$, $\mathcal{O}_i$ is supported within S_i , $\gamma = \mu\chi - \log(2e3^d) > 1$, and $v = 8e3^d v_{\text{LR}}$.

Proof. We note that $e^{-\gamma+\kappa} < 1$, so we can apply the bound $(e^{-\gamma+\kappa} + \mu vt)^{m-1} < e^{(m-1)\mu vt}$ to Theorem 8. Using $\mu vt < e^{\mu vt}$ and $v = 4e3^d v_{\text{LR}} > v_{\text{LR}}$ (with v defined as in Theorem 8) gives the simplified bound. $\blacksquare$

The conclusion of Corollary 5 also holds for quasi-local interactions.

Corollary 7. Suppose that H is a quasi-local Hamiltonian. Let B_R denote a metric ball of radius R . There exist constants $\gamma, c_{\text{LR}} > 0$ such that we can choose $v_1, \dots, v_m \in B_R$ (where m depends on R) for which

$$C_{v_1, \dots, v_m}^{B_R^c}(t) \leq c_{\text{LR}} \exp\left(-\gamma \frac{(R - v_{\text{LR}} t)^d}{(v_{\text{LR}} t)^{d-1}}\right) \tag{108}$$

for any $v_{\text{LR}} t > 1$.

Proof. First, consider the interaction graph embedded in $\mathbb{R}^d$. Let $\chi := \max(\kappa, \log[2e^2 3^d])/\mu$. There exists a constant B such that, for $\xi > 1$, we can tile $\mathbb{R}^d$ with Euclidean boxes where each box contains a ball of radius $\xi + \chi$ and is contained within a ball of radius $B\xi$. Let C_1 be a constant such that $|\partial B_\xi| < (C_1 \xi)^{d-1}$ for all $\xi > 1$, and let $\xi := \alpha vt$ for some $\alpha > 1$. As in Corollary 5, there is a constant C_2 such that, for each $\alpha > 1$, we can find $K \gg 1$ such that, for all $R > K v_{\text{LR}} t$, we can find $m > C_2 R^d / \xi^d$ of these boxes

lying entirely within B_R . Then from Corollary 6, we have

$$C_{\tilde{v}}^R \leq ((C_1 \xi)^{d-1} e^{\mu(2vt-\xi)})^{C_2 \frac{R^d}{\xi^d}}. \quad (109)$$

The bound above can be re-written as

$$((C_1 \xi)^{d-1} e^{\mu(2vt-\xi)})^{C_2 \frac{R^d}{\xi^d}} = \left((C_1 \alpha vt)^{\frac{d-1}{\alpha vt}} e^{\mu \frac{(2-\alpha)}{\alpha}} \right)^{C_2 \frac{R^d}{(\alpha vt)^{d-1}}}. \quad (110)$$

Taking the logarithm of the term in parentheses,

$$\begin{aligned} & \log((C_1 \alpha vt)^{(d-1)/\alpha vt} e^{\mu(2-\alpha)/\alpha}) \\ & \times \leq \frac{1}{\alpha} \left[\frac{d-1}{e} + (d-1) \log(C_1 \alpha) + 2\mu \right] - \mu := -\beta. \end{aligned} \quad (111)$$

Since the first term on the right-hand side vanishes as $\alpha \rightarrow \infty$, it is clear that we can choose α large enough such that $\beta > \mu/2$. Thus the bound becomes

$$C_{\tilde{v}}^R \leq e^{-\frac{C_2 \mu}{2\alpha^{d-1}} \frac{R^d}{(vt)^{d-1}}}. \quad (112)$$

With this demonstrated for $R > K v_{\text{LR}} t$, the rest of the proof is the same as in Corollary 5, save for the fact that we must choose a smaller μ to achieve the same area prefactor in Proposition 2 as in Theorem 5. ■

VI. ACCURACY OF CLASSICAL SIMULATIONS OF QUANTUM DYNAMICS

In random unitary circuits [39,40], the dynamics are restricted by a light cone with volume $\sim t^d$. The complexity of simulating these dynamics by truncating the operator outside of the light cone is exactly bounded by $\exp(t^d)$. This intuition breaks down for locally generated continuous-time evolution, where the light cone is not sharp. As we discussed in the Introduction, one then needs to specify an error tolerance ε to which the simulation must be performed. Using classical LR bounds, one expects that the truncation algorithm with tolerance ε will produce a quasi-polynomial error growth $\sim \exp(\log(\varepsilon)^d)$, which originates from failing to control the magnitude of the operator tails outside the light cone. In Ref. [42], it was shown that, for product initial states and strictly local interactions, a simulation algorithm based on truncating the support of an operator to a large subset achieves the optimal error scaling $\varepsilon^{-\exp(t)}$, which is polynomial in ε , but seemingly suboptimal in t in finite dimensions.

The bound achieved by Wild *et al.* in [42] draws on previous cluster expansion techniques from statistical physics. Such approximation methods involve an asymptotic expansion in the local parameters of the Hamiltonian, which work as long as the Hamiltonian is local on a

bounded-degree graph. By contrast, we use the machinery developed above to revisit classical simulation of quasi-local dynamics from the perspective of an expansion in *time*, and show that our algorithm is able to optimally leverage spatial locality and the finite dimensionality of the lattice.

A. Operator expansion with exponential-in-volume tails

In order to bound the simulability of quantum systems with quasi-local interactions (which also includes strictly local interactions), we wish to have better control over the fraction of an operator which is supported on *any* large set S containing $\gg (v_{\text{LR}} t)^d$ vertices. We have already made progress along these lines. In Corollaries 5 and 7, we made a weaker statement about the tail of an operator supported on vertices within a *specific* large set S . What remains is to show that, even though there are $\sim \exp(|S|)$ possible choices of connected sets with the volume of S —a factor which is large enough to easily overwhelm the volume-law LR bound in either corollary—the *net* contribution of all such large operators is still exponentially small in volume. This is achieved by the following theorem, where, for technical convenience, we restrict to dynamics on a square lattice.

Theorem 9. Let H be a quasi-local Hamiltonian on a d -dimensional square lattice V of qudits, whose LR bound is given by Proposition 2 with parameters μ, v_{LR} . Consider a local operator A with $\|A\| = 1$ acting on a connected region of size $O(1)$, which is evolved to $A(t)$ at time t by H . It can be expanded to connected subsets S that contain the support of A :

$$A(t) = \sum_{S \subset V} \tilde{A}(S; t). \quad (113)$$

Here $\tilde{A}(S; t)$ is supported in S , such that all contributions from $|S| > M$ decay exponentially with the volume cutoff M : for any $M \geq 1$,

$$\tilde{A}(t) = \sum_{\substack{S \subset V: |S| \leq M \\ \text{supp}(A) \subseteq S}} \tilde{A}(S; t) \quad (114)$$

approximates $A(t)$ with error

$$\|A(t) - \tilde{A}(t)\| \leq c_d \exp \left[4\mu v_{\text{LR}} t - c'_d \mu \frac{M}{(v_{\text{LR}} t + c_{\text{box}})^{d-1}} \right], \quad (115)$$

where $c_d, c'_d, c_{\text{box}}$ are constants determined by d, μ, v_{LR} .

Proof. Similar to the previous section, we partition the lattice sites into non-overlapping boxes $V = b_0 \sqcup b_1 \sqcup \dots$,

where each box b is a cube of length r (thus containing r^d sites), and operator A is contained in one box b_0 . We choose r to be

$$r = 4(v_{\text{LR}}t + c_{\text{box}}) > \max \left[d, \frac{45}{4\mu} 3^d (\log 2 + d \log 3) \right], \quad (116)$$

where the inequality holds by choosing a sufficiently large $O(1)$ constant c_{box} . The boxes can be viewed as vertices of a coarse-grained d -dimensional square lattice, and form a graph $G_b = (V_b, E_b)$ where two boxes $b, b' \in V_b$ share an edge in E_b if and only if they are neighbors, because $d(b, b') \leq d$. Here $d(\cdot, \cdot)$ is the distance function on the original lattice V . Note that here we do not work with factor graphs directly. According to the Taylor expansion (20), we organize the time-evolved operator $A(t) = e^{t\mathcal{L}}A$ by the boxes it touches [or has touched in the past by taking commutators with factors X_j in the causal tree $(X_1, \dots, X_n)$]:

$$A(t) = \sum_{\text{connected } S_b \subset V_b: S_b \ni b_0} A(S_b; t) \quad (117)$$

with

$$|A(S_b; t)| := \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{X_1, \dots, X_n \in F: (b \in S_b \Leftrightarrow \exists X_j \text{ such that } X_j \cap b \neq \emptyset)} \times \mathcal{L}_{X_n} \cdots \mathcal{L}_{X_1} |A|. \quad (118)$$

The summation here is taken over sets of adjacent boxes $S_b \subset V_b$ that are connected in G_b , which we call a *connected cluster* [56]. In other words, we include the causal tree $(X_1, \dots, X_n)$ in $A(S_b; t)$ if and only if the factors touch all boxes in S_b but no boxes in its complement S_b^c . Observe that we have restricted to a subset of boxes S_b that are connected in G_b and that contain b_0 ; otherwise, $A(S_b; t) = 0$ simply because the operator grows in a connected way. ■

An equivalent way to express $A(S_b; t)$ is by induction: the starting point is

$$A(\{b_0\}; t) = e^{t\mathcal{L}_{b_0}} A, \quad (119)$$

where $\mathcal{L}_S := \sum_{X \in F: X \subset S} \mathcal{L}_X$. Then, given all $A(S_b; t)$ for $|S_b| \leq m$, any S'_b with size $|S'_b| = m + 1$ is given by

$$A(S'_b; t) = e^{t\mathcal{L}_{S'_b}} A - \sum_{S_b \subsetneq S'_b: S_b \ni b_0} A(S_b; t), \quad (120)$$

where $\mathcal{L}_{S'_b} := \mathcal{L}_{b_0 \cup b_1 \cup \dots \cup b_m}$ for $S'_b = \{b_0, b_1, \dots, b_m\}$. Equation (120) holds because, when Taylor-expanding

$e^{t\mathcal{L}_{S'_b}} A$, any causal tree $(X_1, \dots, X_n)$ that does not touch all boxes in S'_b is canceled by a corresponding term in exactly one $A(S_b; t)$ in the second term of Eq. (120). We note that this inclusion-exclusion principle also plays an important role in the linked-cluster expansion algorithm [57,58] for computing extensive properties in statistical mechanics.

We adapt the nested commutator bound in Theorem 8 to show that $A(S_b; t)$ decays exponentially with $|S_b|$. The proof of this result is delayed until Sec. VIC.

Lemma 8. There exist constants c_d determined by d , and c_{box} determined by d, μ, v_{LR} , such that Eq. (116) holds, and, for any connected cluster S_b ,

$$\|A(S_b; t)\| \leq c_d \exp[-\mu(3^{-d}|S_b| - 1)r/5]. \quad (121)$$

We then approximate $A(t)$ by an operator

$$\tilde{A}(t) = \sum_{\text{connected cluster } S_b: |S_b| \leq M_*} A(S_b; t) \quad (122)$$

that truncates the sum over clusters in Eq. (117) at size M_* (an integer). The truncation error is bounded by

$$\begin{aligned} \|A(t) - \tilde{A}(t)\| &\leq \sum_{\text{connected cluster } S_b: |S_b| > M_*} \|A(S_b; t)\| \\ &\leq c_d \sum_{m > M_*} (3^d e)^m \exp[-\mu(3^{-d}m - 1)r/5] \\ &\leq c_d \exp\left[\frac{\mu r}{5} - M_* \left(3^{-d} \frac{\mu}{5} r - d \log 3\right)\right] \\ &\quad \times \frac{1/2}{1 - 1/2} \leq c_d \exp[-\mu r(3^{-d-2}M_* - 1)]. \end{aligned} \quad (123)$$

Here in the second line, we have applied Lemma 8 and Proposition 3 on the number of connected clusters $\leq (3^d e)^m$ of size m . In the last line, we have used the fact that

$$3^{-d} \frac{\mu}{5} r - d \log 3 \geq \max(\log 2, 3^{-d-2} \mu r) \quad (124)$$

from Eq. (116).

The theorem then follows by identifying $\tilde{A}(S; t)$ with $A(S_b; t)$, where the S are the sites contained in S_b . Equation (115) comes from Eq. (123) with $M = M_* r^d$ and an updated constant c_d .

B. Simulatability bound

According to the expansion obtained in Theorem 9, one can classically simulate $A(t)$ by its truncation $\tilde{A}(t)$.

Corollary 8. Let H be a quasi-local Hamiltonian on a d -dimensional square lattice V of qudits, whose LR bound

is given by Proposition 2 with parameters c_{LR}, μ, v_{LR} . Consider a local operator A with $\|A\| = 1$ acting on a connected region of $O(1)$ sites, and an initial state ρ whose marginal ρ_S on any connected set $S \subset V$ can be obtained classically with complexity $\exp[O(|S|)]$. The expectation $\text{Tr}[A(t)\rho]$ with any $t > 0$ can be computed to error ε with complexity

$$\exp \left[O \left((1 + v_{LR}t)^{d-1} \left(v_{LR}t + \log \frac{1}{\varepsilon} \right) \right) \right]. \quad (125)$$

Here the complexity assumption on the initial state applies to, e.g. product states, thermal Gibbs states at high temperature [59], and gapped ground states that are connected to solvable states by a gapped path of Hamiltonians [30]. The reason is that, for those states, expectation values of local observables can be computed locally in an efficient way, so that the reduced density matrix ρ_S can be obtained by state tomography with an extra $\exp[O(|S|)]$ overhead. Note that our algorithm in the proof of Corollary 8 actually applies to more general initial states, which reduces the problem of computing local observables after time evolution to the problem of obtaining local marginals of the initial state.

Equation (125) achieves a polynomial dependence of $1/\varepsilon$ even for $d > 1$, which improves upon a quasipolynomial bound $\exp[O((v_{LR}t + \log(1/\varepsilon))^d)]$ from standard LR bounds (see, e.g. Proposition 4.4 of Ref. [2]). On the other hand, although a polynomial-in- $1/\varepsilon$ bound (Theorem 6 of Ref. [42]) has been obtained from cluster expansions in the restricted setting where ρ is a product state, its complexity $\exp[O(e^{\Theta(t)}(t + \log(1/\varepsilon)))]$ grows exponentially faster than Eq. (125) in terms of the t dependence. The physical reason is that the cluster expansion techniques in Ref. [42] work for any bounded-degree graph. Our bound takes full advantage of the additional locality properties of finite-dimensional quantum systems. Our simulation complexity (125) thus rules out a proposed super-polynomial quantum advantage in analog quantum simulators [60] while maintaining the correct scaling with t .

We argue that bound (125) is tight with respect to the dependence on both t and $1/\varepsilon$, because of the following. First, we argue that the $d = 1$ case of Eq. (125) is tight because the LR bound (18) can be (almost) saturated by a nearest-neighbor Hamiltonian [2], so the operator is effectively evolved in a truncated subsystem of $\Theta(v_{LR}t + \log(1/\varepsilon))$ qudits, whose simulation complexity is then exponential in the subsystem size and given in Eq. (125). Building on this one-dimensional example, we expect that the following time-dependent protocol $H(s)$ would saturate Eq. (125) for any constant dimension $d > 1$ (note that Theorem 8 generalizes to time-dependent cases). Here time s is backwards so that operator A first hits $H(0)$ instead of $H(t)$ in Heisenberg evolution. In the first half of

the protocol $0 < s < t/2$, $H(s)$ just evolves the initial local operator along a one-dimensional chain [e.g. $H(s)$ is a set of decoupled one-dimensional Hamiltonians], so that there is a tail A_{tail} of operator $A(t/2)$ that is supported in a chain of length $\sim v_{LR}t + \log(1/\varepsilon)$ with

$$\|A_{\text{tail}}\| \sim \varepsilon. \quad (126)$$

Such a tail exists due to the tightness of LR bound in one dimension discussed above. The second half of the protocol $t/2 < s < t$ then evolves in the horizontal directions of the original chain, which grows A_{tail} to a highly entangled operator A'_{tail} occupying a “cylinder” of height $\sim v_{LR}t + \log(1/\varepsilon)$ and radius $\sim v_{LR}t$. This second half can even be a quantum circuit with strict light cones. Since unitary evolution does not change operator norm (126) of this part of operator $\|A'_{\text{tail}}\| \sim \varepsilon$, one has to simulate dynamics in the whole cylinder to get precision ε , which requires complexity (125) that is an exponential of the cylinder volume.

Proof of Corollary 8. In the proof of Theorem 9, we expanded $A(t)$ into connected clusters S_b . Choosing

$$M_* = \left\lceil \frac{3^{d+2}}{\mu r} \log \left(\frac{2c_d}{\varepsilon} \right) \right\rceil + 3^{d+2} + 1, \quad (127)$$

the truncation error (123) leads to

$$|\text{Tr}[\rho(A(t) - \tilde{A}(t))]| \leq \|A(t) - \tilde{A}(t)\| \leq \frac{\varepsilon}{2}. \quad (128)$$

Therefore, to simulate $\text{Tr}[\rho A(t)]$, it suffices to compute

$$\text{Tr}[\rho \tilde{A}(t)] = \sum_{\text{connected cluster } S_b: |S_b| \leq M_*} a(S_b), \quad (129)$$

where

$$a(S_b) := \text{Tr}[\rho A(S_b; t)] = \text{Tr}[\rho_{S_b} A(S_b; t)] \quad (130)$$

only involves a marginal of the initial state $\rho_{S_b} := \rho_{b_0 \cup b_1 \cup \dots \cup b_m}$ for $S_b = \{b_0, b_1, \dots, b_m\}$. According to Eq. (120),

$$a(S_b) = \tilde{a}(S_b) - \sum_{S'_b \subsetneq S_b: S'_b \neq \emptyset} a(S'_b), \quad (131)$$

where

$$\tilde{a}(S_b) := \text{Tr}[\rho_{S_b} e^{t\mathcal{L}_{S_b}} A] = \text{Tr}[\rho_{S_b} e^{itH_{S_b}} A e^{-itH_{S_b}}]. \quad (132)$$

The simulation algorithm then works as follows.

For each $\tilde{a}(S_b)$, since it only involves a subsystem of $\leq M_* r^d$ qudits where the marginal ρ_S is assumed to be

ALGORITHM 2. A classical simulation algorithm for $\text{Tr}[\rho A(t)]$.

```

Generate all connected clusters  $S_b \ni b_0$  with  $|S_b| \leq M_*$ 
Create array  $a(S_b) \leftarrow 0$  for the generated clusters  $\{S_b\}$ 
for  $m = 1, 2, \dots, M_*$  do
  for all connected cluster  $S_b \ni b_0$  with  $|S_b| = m$  do
    Compute  $\tilde{a}(S_b)$  in (6.20) by exponentiating the matrix  $H_{S_b}$ 
    Update  $a(S_b)$  by (6.19) using  $\tilde{a}(S_b)$  and the stored values for  $a(S'_b)$  where  $|S'_b| < m$ 
  end for
end for

return  $\sum_{S_b} a(S_b)$ 

```

computable in exponential time, $\tilde{a}(S_b)$ can be computed accurately by

$$N = \exp[\mathcal{O}(M_* r^d)] \quad (133)$$

resources. This dominates the complexity for $a(S_b)$, as the sum over $a(S'_b)$ in Eq. (131) contains at most $\exp[\mathcal{O}(M_*)]$ terms [61]. Since there are at most $\exp[\mathcal{O}(M_*)]$ clusters and they can be enumerated by a similar cost according to Proposition 3, the total complexity of the algorithm is still of the scaling (133), which becomes Eq. (125) by plugging in Eqs. (116) and (127). The final error is bounded by ε because the M_* truncation error is bounded by $\varepsilon/2$ in Eq. (123), and the other operations like computing $\tilde{a}(S_b)$ are essentially exact using resources above. ■

C. Applying the nested commutator bound: proof of Lemma 8

For S_b that only contains boxes b_0 or its nearest neighbors (there are $3^d - 1$ of them), Eq. (121) holds by choosing a sufficiently large c_d determined by d . The reason is that $\|A(\{b_0\}; t)\| \leq 1$ due to Eq. (119), and the other $\|A(S_b; t)\|$ can be bounded by iterating Eq. (120): for example, let b_1 be one neighbor of b_0 ; then

$$\begin{aligned} \|A(\{b_0, b_1\}; t)\| &= \|e^{t\mathcal{L}_{\{b_0, b_1\}}} A - A(\{b_0\}; t)\| \\ &\times \leq \|e^{t\mathcal{L}_{\{b_0, b_1\}}} A\| + \|A(\{b_0\}; t)\| \leq 1 + 1 = 2. \end{aligned} \quad (134)$$

There are only finitely many S_b that only contain b_0 and its neighbors, so, by a finite number of iterations, they all satisfy $\|A(S_b; t)\| \leq c_d$ for some constant c_d determined by d , so that Eq. (121) is satisfied for them because the exponential factor is always ≤ 1 for these S_b with $|S_b| \leq 3^d$.

Beyond these finitely many connected clusters, any other S_b contains at least one box that is a distance $\geq r$ from the initial operator A . We can then find boxes $S_1, S_2, \dots, S_M \in S_b$ with $M \geq 1$ such that each S_m is surrounded by a larger box B_m of side length $2(r - \chi)$ (where $r/2 - \chi$ is an integer and $r, \chi \geq 1$ are chosen shortly) and the larger boxes $B_1, \dots, B_M$ are a distance $\geq 2\chi$ from each

other and from b_0 . See Fig. 8 for an illustration. More precisely, because we have excluded the finitely many S_b in the neighborhood of b_0 , we can always first find an $S_1 \in S_b$ whose larger box B_1 does not touch b_0 . Then we try to find an $S_2 \in S_b$ that is not a neighboring box of S_1 , and so on. We can always find such a set $S_1, S_2, \dots, S_M \in S_b$ containing at least

$$M \geq 3^{-d}|S_b| - 1 \quad (135)$$

boxes, because each S_m only forbids its $3^d - 1$ neighbors to be selected at the same time. Here we need the -1 in Eq. (135) to also avoid boxes neighboring b_0 .

For a given S_b , after finding these M boxes S_m together with their surrounding boxes B_m , we can then apply Theorem 8 to bound $\|A(S_b; t)\|$. The reason is that although Theorem 8 is stated directly for the nested commutator $C_{S_1, \dots, S_M}^R(t)$, its proof works by individually bounding each of the irreducible skeletons that contribute to $C_{S_1, \dots, S_M}^R(t)$. Here $\|A(S_b; t)\|$ is also bounded by a sum over irreducible

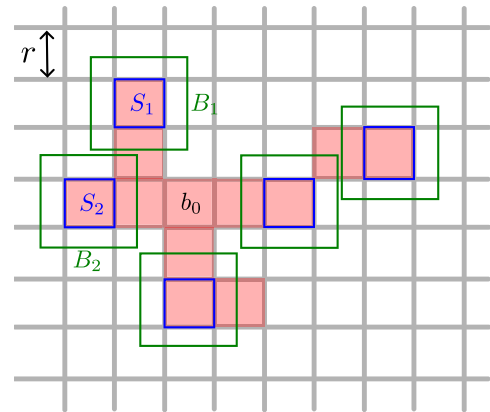


FIG. 8. To apply the nested commutator bound for a connected cluster S_b (red shaded region), we find boxes $S_1, S_2, \dots, S_M \in S_b$ (blue) that are well separated from each other and from the initial operator in b_0 . We can always find $M = \Theta(|S_b|)$ number of them, which are surrounded by larger boxes $B_1, \dots$ that we use in the proof of the nested commutator bound.

skeletons, with the only difference that here the causal forests in Eq. (117) corresponding to the irreducible skeletons are restricted to be contained in S_b . In other words, $\|A(S_b; t)\|$ is bounded by the bound on $C_{S_1, \dots, S_M}^R(t)$ if H only contains terms supported inside S_b ; but this truncated H is still a quasi-local Hamiltonian so we can just apply Eq. (96). Another small difference here is that $\{B_m\}$ do not lie in a square lattice. Theorem 8 generalizes to this case because its proof only uses the fact that the whole space can be tiled by $\{B_m\}$ and some other regions where the connectivity graph of the regions have bounded degree $\leq 3^d$; this property also holds here.

Equation (96) thus implies that

$$\begin{aligned} \|A(S_b; t)\| &\leq (c'(t+1) \cdot r^{2(d-1)} \cdot e^{\mu(v_{\text{LR}}t - r/2 + \chi)})^M \\ &\leq (c''r^{2d-1} \cdot e^{-\mu r/4})^M \leq (e^{-\mu r/5})^M, \end{aligned} \quad (136)$$

where c', c'' are constants. Here in the first line, we have used $|\partial S_i| \leq |\partial B_i| = O(r^2)$ and chosen a sufficiently large constant χ independent of t to get the factor $(t+1)$. In the second line of Eq. (136), we have used Eq. (116) with a sufficiently large constant c_{box} , so that $t+1 = O(r)$ and $v_{\text{LR}}t - r/2 + \chi = -r/4 - c_{\text{box}} + \chi \leq -r/5 - \mu^{-1} \log(c''r^{2d-1})$; this last condition is achievable because the function $f(r) := r/20 - \mu^{-1} \log(c''r^{2d-1})$ is bounded from below. Equation (121) follows by plugging Eq. (135) into Eq. (136).

VII. SPONTANEOUS SYMMETRY-BREAKING PHASES OF FINITE SYMMETRIES

We now turn to the second main application of our nested commutator bounds, and discuss the diagnosis of non-trivial phases of quantum matter. We focus on phases that exhibit spontaneous symmetry breaking (SSB) of finite symmetries. SSB can be detected by the presence of ground states with exponentially small energy splitting, along with order and disorder parameters computed in these nearly degenerate states. For example, in the Ising ferromagnetic phase, there are two degenerate (in the thermodynamic limit) ground states $|\psi_{\pm}\rangle$ with eigenvalues ± 1 under the global $\mathbb{Z}_2$ symmetry operator. These states are adiabatically connected to the $\pm$ GHZ states

$$|\pm\rangle = \frac{|\mathbf{0}\rangle \pm |\mathbf{1}\rangle}{\sqrt{2}}. \quad (137)$$

More concretely, it was shown in Ref. [30] using LR bounds that the splitting between the energies of $|\psi_{\pm}\rangle$ is at most exponentially small in the linear system size L :

$$\delta := |\langle \psi_+ | H | \psi_+ \rangle - \langle \psi_- | H | \psi_- \rangle| \leq e^{-O(L)}. \quad (138)$$

However, deep in the ferromagnetic phase, one expects that adding a small transverse field $h \sum_i X_i$ would lead to

$\delta \sim e^{-O(L^d)}$, because the two ground states $|\pm\rangle$ are distinguished only by the global operator $\prod_i X_i$. Intuitively, this expectation is because it requires going to $O(L^d)$, not $O(L)$, in perturbation theory, in order to connect $|\pm\rangle$ via perturbations.

Spontaneous symmetry breaking is also marked by a long-range order parameter $\lim_{|i-j| \rightarrow \infty} \langle \psi | Z_i Z_j | \psi \rangle \sim O(1)$ in states in the ground-state subspace, as well as a quickly decaying disorder parameter $\lim_{R \rightarrow \infty} \langle \psi | D_R | \psi \rangle$. Here, d is the spatial dimension, and $D_R = \prod_{i \in B_R(v)} X_i$ is the global $\mathbb{Z}_2$ symmetry operator restricted to a d -dimensional ball of radius R centered at any vertex v . Like with the splitting, conventional LR bounds give a far looser bound, saying that $\lim_{R \rightarrow \infty} \langle \psi_+ | D_R | \psi_+ \rangle \lesssim e^{-O(R)}$ rather than the bound one might expect from perturbation theory: $\lim_{R \rightarrow \infty} \langle \psi_+ | D_R | \psi_+ \rangle \lesssim e^{-O(R^d)}$. Indeed, this latter scaling is observed numerically [46].

A. Proofs of volume-law scaling

Using our nested commutator bound, we confirm the intuition raised in the two arguments above. If states $|\psi_{\pm}\rangle$ are connected to the GHZ fixed point states $|\pm\rangle$ by finite-time evolution generated by a quasi-local Hamiltonian (satisfying Definition 9), then the ground-state splitting is in fact at most exponentially small in the *volume* of the system. Similarly, we show that the disorder parameter is also at most exponentially small in the *volume* of $B_\ell(v)$. It was shown in Ref. [62] that, for a sufficiently small local perturbation in the vicinity of the ferromagnetic fixed point, such a quasi-local generator exists. The usual quasi-adiabatic generator [63–65] that can be constructed within the entire adiabatic phase decays almost exponentially in diameter [to be precise, exponentially in $R/\log^2(R)$], so outside of this vicinity of the fixed point, it is not known whether an adiabatic generator satisfying our stronger definition of quasi-local exists. Therefore, at least for states in the vicinity of the ferromagnetic fixed point, our nested commutator results give bounds on the ground-state energy splitting and disorder parameter that are much tighter than those given by the usual LR bounds. It seems reasonable to conjecture that such a quasi-local generator might exist within the entire phase, not just in the vicinity of the fixed point. This would imply that these bounds hold within the entire ferromagnetic phase.

In the following, we specialize to properties of the Ising ferromagnetic phase on a lattice in d dimensions, with a constant finite density of sites with respect to the Euclidean distance, where each site hosts a qubit. However, our results are straightforward to generalize to spontaneous symmetry-breaking phases of other finite symmetries, in systems with more general local Hilbert spaces. The following two theorems prove the two conjectures stated

above, given the assumption of the exponential-in-volume-tailed quasi-adiabatic generator.

Theorem 10. If $|\psi_{\pm}\rangle$ are ground states of a gapped, quasi-local Hamiltonian satisfying Eq. (9), that are connected to the $\pm$ GHZ states $|\pm\rangle$ defined in Eq. (137) in spatial dimension d by finite-time evolution generated by a quasi-local Hamiltonian, then the splitting δ defined in Eq. (138) obeys

$$\delta \leq cL^d \exp\left(-\gamma \frac{(L-vt)^d}{(vt)^{d-1}}\right) \quad (139)$$

for $L \gg vt$, where c, γ are $\mathcal{O}(1)$ constants, L is the linear system size, and v and t are $\mathcal{O}(1)$ constants describing the finite-time evolution and its quasi-local generator.

Theorem 11. If $|\psi\rangle = U(t)|\mathbf{0}\rangle$, where $U(t)$ describes finite-time evolution in d dimensions generated by a local, $\mathbb{Z}_2$ symmetric Hamiltonian, then the disorder parameter $\lim_{R \rightarrow \infty} \langle \psi | D_R | \psi \rangle$ for a d -dimensional ball of radius R is upper bounded by

$$\lim_{R \rightarrow \infty} \langle \psi | D_R | \psi \rangle \leq c_{LR} \exp\left(-\gamma \frac{(R-vt)^d}{(vt)^{d-1}}\right), \quad (140)$$

where again c_{LR} and γ are $\mathcal{O}(1)$ constants and v, t describe the time evolution.

One subtlety about the above theorems is that in any given state $|t\rangle$ may be very large, even $\mathcal{O}(\log L)$. Therefore, to probe the scaling of the disorder parameter and truly rule out the possibility that a state is connected to the ferromagnetic fixed point by finite [or $\mathcal{O}(\log L)$] time evolution generated by a quasi-local Hamiltonian, we need to compute expectation values of extensive operators and compare very small values (exponentially small in L versus exponentially small L^d). In Sec. VII B, we present an example of a state $|\psi(\beta)\rangle$ that fails to satisfy Theorem 11. However, due to the subtlety above, we can obtain another state $|\tilde{\psi}(\beta)\rangle$ whose overlap with $|\psi(\beta)\rangle$ goes to 1 in the thermodynamic limit, and does not violate Theorem 11. It is likely that $|\tilde{\psi}(\beta)\rangle$ can be connected to the ferromagnetic fixed point if we allow $|t|$ to be $\mathcal{O}(\log L)$. This point is discussed more in depth in Ref. [66]. However, for realistic systems where t is an $\mathcal{O}(1)$ constant, the scaling of the disorder parameter provides a reliable diagnostic for the ferromagnetic phase.

To prove the above two theorems, we make use of the following lemma. The rough idea of the lemma is that an expectation value of the form $\langle \psi | D_{\text{tot}} O | \psi \rangle$, where $D_{\text{tot}} = \prod_i X_i$ is the global $\mathbb{Z}_2$ symmetry operator, can be written as the expectation value of a large nested commutator of various single-site unitaries Z_i with $O(t)$. Here, $|\psi\rangle = U(t)|\mathbf{0}\rangle$, where $U(t)$ is $\mathbb{Z}_2$ symmetric, and $O(t) = U(t)^\dagger O(0) U(t)$.

The polarized state is denoted by $|\mathbf{0}\rangle$, which is the unique eigenstate of every Z_i operator with eigenvalue 1. This rewriting of $\langle \psi | D_{\text{tot}} O | \psi \rangle$ will allow us to apply our nested commutator bound.

Lemma 9. Let O be an operator and let $|\psi\rangle = U(t)|\mathbf{0}\rangle$. Then, for any subset of sites $v_1, \dots, v_m$, we have

$$\begin{aligned} \langle \psi | D_{\text{tot}} O | \psi \rangle &= \frac{1}{2^m} \langle \mathbf{0} | D_{\text{tot}} [[\dots [O(t), Z_{v_1}], \dots Z_{v_{m-1}}], \\ &\times Z_{v_m}] | \mathbf{0} \rangle := \frac{1}{2^m} \langle \mathbf{0} | D_{\text{tot}} [O(t), Z_{\vec{v}}] | \mathbf{0} \rangle. \end{aligned} \quad (141)$$

Proof. We use the following observations: (1) the evolution operator $U(t)$ commutes with D_{tot} , (2) $|\mathbf{0}\rangle$ is an eigenstate of Z_{v_i} , and (3) Z_{v_i} anticommutes with D_{tot} . Putting these observations together, we get

$$\begin{aligned} \langle \psi | D_{\text{tot}} O | \psi \rangle &= \langle \mathbf{0} | D_{\text{tot}} O(t) | \mathbf{0} \rangle \\ &= \langle \mathbf{0} | Z_{v_1} D_{\text{tot}} O(t) | \mathbf{0} \rangle \\ &= -\langle \mathbf{0} | D_{\text{tot}} Z_{v_1} O(t) | \mathbf{0} \rangle \\ &= \langle \mathbf{0} | D_{\text{tot}} [O(t), Z_{v_1}] | \mathbf{0} \rangle - \langle \psi | D_{\text{tot}} O | \psi \rangle. \end{aligned} \quad (142)$$

Rearranging the above gives $\langle \psi | D_{\text{tot}} O | \psi \rangle = \frac{1}{2} \langle \mathbf{0} | D_{\text{tot}} [O(t), Z_{v_1}] | \mathbf{0} \rangle$. This is the base case. Now we assume that Eq. (9) holds for a subset of sites $\vec{v}' = (v_1, \dots, v_{m-1})$, and use the exact same reasoning to make the inductive step:

$$\begin{aligned} \langle \psi | D_{\text{tot}} O | \psi \rangle &= \frac{1}{2^{m-1}} \langle \mathbf{0} | D_{\text{tot}} [O, Z_{\vec{v}'}] | \mathbf{0} \rangle \\ &= -\frac{1}{2^{m-1}} \langle \mathbf{0} | D_{\text{tot}} Z_{v_m} [O, Z_{\vec{v}'}] | \mathbf{0} \rangle \\ &= \frac{1}{2^{m-1}} (\langle \mathbf{0} | D_{\text{tot}} [[O, Z_{\vec{v}'}], Z_{v_m}] | \mathbf{0} \rangle \\ &\quad - \langle \mathbf{0} | D_{\text{tot}} [O, Z_{\vec{v}'}] | \mathbf{0} \rangle) \\ &= \frac{1}{2^m} \langle \mathbf{0} | D_{\text{tot}} [O, Z_{\vec{v}}] | \mathbf{0} \rangle. \end{aligned} \quad (143)$$

Here $\vec{v} = (v_1, \dots, v_{m-1}, v_m)$ and we define $[O, Z_{\vec{v}}] := [[\dots [O, Z_{v_1}], \dots], Z_{v_m}]$. This completes the proof. ■

With this technical lemma in hand, we now prove the theorems.

Proof of Theorem 10. Using $|\psi_{\pm}\rangle = U(t)|\pm\rangle$, we obtain

$$\begin{aligned} \delta &= |\langle + | H(t) | + \rangle - \langle - | H(t) | - \rangle| \\ &= 2 |\text{Re} \{ \langle \mathbf{1} | H(t) | \mathbf{0} \rangle \}| \\ &= 2 |\text{Re} \{ \langle \mathbf{0} | D_{\text{tot}} H(t) | \mathbf{0} \rangle \}| \\ &\leq 2 \sum_{u \in V} \sum_{X \ni u} |\text{Re} \{ \langle \mathbf{0} | D_{\text{tot}} H_X(t) | \mathbf{0} \rangle \}|, \end{aligned} \quad (144)$$

where in the last line, we decomposed the $H(t)$ into terms on connected clusters X as in Eq. (9). Note that we did not need to assume that $U(t)$ is $\mathbb{Z}_2$ symmetric here because we did not need to apply the first step of Eq. (142).

If H is strictly local, then we just have a sum over $u \in V$ and each of the terms above is of the form $\langle 0 | D_{\text{tot}} O(t) | 0 \rangle$, so we can apply Lemma 9 and then Theorem 8 to bound each term. Then Theorem 10 follows directly from Corollary 7. There is an overall factor of L^d coming from the sum over $u \in V$.

If H is quasi-local then we divide the sum into clusters X of size $|X| < L^d/(B\xi)^{d-1}$ and $|X| \geq L^d/(B\xi)^{d-1}$, where $B\xi$ is the linear size of the m boxes that we divide the lattice into in Corollary 7, to fit in m balls containing the operators $Z_{v_1}, \dots, Z_{v_m}$. Roughly speaking, for every

$u \in V$, we sum over the connected clusters X containing u [of which there are $\leq (e\Delta)^{|X|}$ by Proposition 3] suppressed by their weight $he^{-\kappa|X|}$ and further suppressed by a nested commutator obtained by inserting as many balls as possible in a region of size $\sim L^d - |X|$. More precisely, the nested commutator can include all of the operators $Z_{v_1}, \dots, Z_{v_m}$ except for those in boxes that overlap with X . To accommodate for the worst-case scenario where the connected cluster X forms a net along the edges of the boxes of linear size $B\xi$ (see the definitions in Corollary 7), to touch as many boxes as possible for smallest $|X|$, we use a region of size $L^d - (B\xi)^{d-1}|X|$ rather than $L^d - |X|$. Once $|X| \sim L^d/(B\xi)^{d-1}$, the worst-case scenario gives no further suppression from the nested commutators. Putting the above observations together, we get

$$\begin{aligned} \delta &\leq 2hL^d \left(\sum_{\substack{X \ni u \\ |X| < L^d/(B\xi)^{d-1}}} (e\Delta)^{|X|} e^{-\kappa|X|} e^{-\frac{C_2\mu}{2\alpha^{d-1}} \frac{L^d - (B\xi)^{d-1}|X|}{(vt)^{d-1}}} + \sum_{\substack{X \ni u \\ |X| \geq L^d/(B\xi)^{d-1}}} (e\Delta)^{|X|} e^{-\kappa|X|} \right) \\ &\leq 2hL^d \left(e^{-\frac{C_2\mu}{2\alpha^{d-1}} \frac{L^d}{(vt)^{d-1}}} \sum_{\substack{X \ni u \\ |X| < L^d/(B\xi)^{d-1}}} \left(\Delta e^{1-\kappa} e^{\frac{C_2\mu}{2\alpha^{d-1}} \frac{(B\xi)^{d-1}}{(vt)^{d-1}}} \right)^{|X|} + \frac{(\Delta e^{1-\kappa})^{\frac{L^d}{(B\xi)^{d-1}}}}{1 - \Delta e^{1-\kappa}} \right). \end{aligned} \quad (145)$$

Performing the first sum and simplifying notation using $y = \Delta e^{1-\kappa}$, we have

$$\delta \leq 2hL^d \left(\frac{y^{\frac{L^d}{(B\alpha vt)^{d-1}}} - e^{-\frac{C_2\mu}{2\alpha^{d-1}} \frac{L^d}{(vt)^{d-1}}}}{ye^{\frac{C_2\mu}{2\alpha^{d-1}} \frac{(B\xi)^{d-1}}{(vt)^{d-1}}} - 1} + \frac{y^{\frac{L^d}{(B\alpha vt)^{d-1}}}}{1 - y} \right), \quad (146)$$

where we used $\xi = \alpha vt$ as in the proof of Corollary 7. Since $y < 1$ by Eq. (68), every term is exponentially decaying in L^d , and we can use the same steps as in Corollary 5 to obtain Theorem 10 with a modified γ and c from the strictly local case. ■

Proof of Theorem 11. We make the observation that the disorder operator can be expressed as $D_R(t) = D_{\text{tot}} \bar{D}_R^\dagger(t)$, where $\bar{D}_R^\dagger$ is the Hermitian conjugate of the disorder operator for the complement of $B_R(v)$. This follows from the fact that $U(t)$ commutes with $D_{\text{tot}} = D_R \bar{D}_R$, the global $\mathbb{Z}_2$ symmetry operator. Now we can apply Lemma 9 to fill the support of D_R , B_R , with operators at site v_i in balls B_i (see Fig. 2). Note that $\bar{D}_R^\dagger$ is supported only in the complement of B_R . Substituting $D_{\text{tot}} \bar{D}_R^\dagger(t)$ for $D_R(t)$, we can then apply

Lemma 9 to get

$$|\langle \psi | D_R | \psi \rangle| \leq \frac{1}{2^m} \left\| \left[Z_{\bar{v}}, \bar{D}_R^\dagger(t) \right] \right\| \leq C_v^{B_R^c}(t), \quad (147)$$

where we used a slight simplification due to the observation that $\|D_{\text{tot}}\| = 1$. We can then apply Corollary 7 to obtain Theorem 11 for sufficiently large R . ■

Note that the usual LR bounds, applied to Eq. (147) with a single commutator, upper bounds $\langle \psi | D_R | \psi \rangle$ by a quantity exponentially small in the diameter of S . In dimension $d > 1$, we get a much stronger bound by using a nested commutator, which is exponentially small in the volume of $B_R(v)$. From the proof above, we see an illustration of the flexibility to optimize the parameters used to get the bound. In $d = 1$, the optimal choice of commutator is simply a single commutator with an operator Z_v at the center of the interval of radius R .

B. Rokhsar-Kivelson states

Rokhsar-Kivelson (RK) states are quantum ground states that encode classical partition functions [47–49]. These states have been studied in the context of conformal quantum critical points (when the classical partition

function goes through a critical point) and have recently received renewed interest in the context of the decohered topological order [67]. In the latter context, RK states appear naturally in the thermofield double (Choi state) representation of the decohered density matrix.

Consider the Ising RK wave function

$$|\psi(\beta)\rangle = \frac{1}{\mathcal{N}} \prod_{\langle ij \rangle} e^{\beta Z_i Z_j / 2} |+\rangle, \quad (148)$$

where $\mathcal{N}$ is a normalization factor to ensure that $\langle \psi(\beta) | \psi(\beta) \rangle = 1$. Here, the product is over all nearest-neighbor vertices on the d -dimensional hypercubic lattice. Although $|\psi(\beta)\rangle$ is a d -dimensional quantum state, it also encodes a d -dimensional classical partition function. It is easy to see that $Z_i Z_j$ correlation functions in $|\psi\rangle$ can be identified with spin-spin correlation functions in the classical Ising model. It follows that

$$\langle \psi(\beta) | Z_i Z_j | \psi(\beta) \rangle \sim \begin{cases} \exp(-\mu|i-j|), & \beta < \beta_c, \\ \mathcal{O}(1), & \beta > \beta_c, \end{cases} \quad (149)$$

where β_c is the classical Ising critical temperature and μ is a finite inverse correlation length. For example, for $d = 2$, $\beta_c = \log(1 + \sqrt{2})/2$. Because there is a long-range order parameter for $\beta > \beta_c$, one might suspect that $|\psi(\beta > \beta_c)\rangle$ belongs in the $\mathbb{Z}_2$ ferromagnetic phase. However,

$$\langle \psi(\beta) | D_R | \psi(\beta) \rangle \sim e^{-2\beta c R^{d-1}}, \quad (150)$$

where cR^{d-1} is the surface area of the radius R ball in d dimensions. The above scaling of the disorder parameter holds for all values of β , including for $\beta > \beta_c$, which can be calculated by relating the disorder parameter to a ratio of two-dimensional classical Ising partition functions, where the only difference between the partition functions comes from deleting the spin-spin couplings along ∂R in the numerator. The scaling is then expected from extensivity of free energy $Z = e^{\beta F}$. Further details can be obtained from Appendix A2 of Sahay *et al.* [68]. Therefore, this state violates the bound in Theorem 11 for all finite β , in any dimension. We therefore conclude that $|\psi(\beta)\rangle$, despite the fact that it demonstrates a long-ranged order parameter for all $\beta > \beta_c$, is not in the vicinity of the ferromagnetic fixed point (where a quasi-adiabatic generator satisfying Definition 9 is guaranteed by [62]) for any finite $\beta > \beta_c$. We conjecture that this state is not in the ferromagnetic phase in any d ; an independent argument for this result in $d = 2$ is found in [68].

VIII. CONCLUSION

This paper has explored the locality of the time evolution of local operators, focusing on obtaining strong

tail bounds on large operators supported on a volume $V \gg (v_{\text{LR}} t)^d$ —the volume of the LR light cone. We found that, loosely speaking, such operators are suppressed as $\exp[-V/(v_{\text{LR}} t)^{d-1}]$, closing a conceptual gap between the LR bound [1,2] and bounds from cluster expansions [42]. Two immediate applications of such bounds were presented—one on the efficiency of classical simulations of quantum dynamics, and one on the classification of quantum phases of matter.

We hope that many further applications of these strong volume-tailed LR bounds can be identified. A natural future direction is to generalize our volume-tailed bounds to systems with bosons [18–25] or power-law interactions [7–17]. It is of great interest to generalize the volume-tailed bounds on the quasi-adiabatic generator, obtained in Ref. [62] in the vicinity of a “code” fixed point, to the entire phase of matter. Lastly, it would also be intriguing if these heavy-tailed LR bounds could help to prove the entanglement area-law for gapped phases in $d > 1$ dimensions, following recent progress [69].

ACKNOWLEDGMENTS

C.Y. thanks Alvaro Alhambra for pointing out the connection with Ref. [60]. C.Z. thanks Rahul Sahay, Ruben Verresen, and Curt von Keyserlingk for collaboration on related work [66,68], and Michael Levin for helpful discussions. C.Y. and A.L. were supported by the Department of Energy under Quantum Pathfinder Grant DE-SC0024324. C.Z. was supported by the Harvard Society of Fellows and the Simons Collaboration on Ultra Quantum Matter.

DATA AVAILABILITY

No data were created or analyzed in this study.

-
- [1] Elliott H. Lieb and Derek W. Robinson, The finite group velocity of quantum spin systems, *Commun. Math. Phys.* **28**, 251 (1972).
 - [2] Chi-Fang (Anthony) Chen, Andrew Lucas, and Chao Yin, Speed limits and locality in many-body quantum dynamics, *Rep. Prog. Phys.* **86**, 116001 (2023).
 - [3] Tobias J. Osborne, Efficient approximation of the dynamics of one-dimensional quantum spin systems, *Phys. Rev. Lett.* **97**, 157202 (2006).
 - [4] Jeongwan Haah, Matthew B. Hastings, Robin Kothari, and Guang Hao Low, Quantum algorithm for simulating real time evolution of lattice Hamiltonians, *SIAM J. Comput.* **52**, FOCS18–250 (2021).
 - [5] S. Bravyi, M. B. Hastings, and F. Verstraete, Lieb-Robinson bounds and the generation of correlations and topological quantum order, *Phys. Rev. Lett.* **97**, 050401 (2006).
 - [6] Jens Eisert and Tobias J. Osborne, General entanglement scaling laws from time evolution, *Phys. Rev. Lett.* **97**, 150404 (2006).

- [7] Zhe-Xuan Gong, Michael Foss-Feig, Fernando G. S. L. Brandão, and Alexey V. Gorshkov, Entanglement area laws for long-range interacting systems, *Phys. Rev. Lett.* **119**, 050501 (2017).
- [8] Chi-Fang Chen and Andrew Lucas, Finite speed of quantum scrambling with long range interactions, *Phys. Rev. Lett.* **123**, 250605 (2019).
- [9] Minh C. Tran, Andrew Y. Guo, Yuan Su, James R. Garrison, Zachary Eldredge, Michael Foss-Feig, Andrew M. Childs, and Alexey V. Gorshkov, Locality and digital quantum simulation of power-law interactions, *Phys. Rev. X* **9**, 031006 (2019).
- [10] Tomotaka Kuwahara and Keiji Saito, Strictly linear light cones in long-range interacting systems of arbitrary dimensions, *Phys. Rev. X* **10**, 031010 (2020).
- [11] Minh C. Tran, Chi-Fang Chen, Adam Ehrenberg, Andrew Y. Guo, Abhinav Deshpande, Yifan Hong, Zhe-Xuan Gong, Alexey V. Gorshkov, and Andrew Lucas, Hierarchy of linear light cones with long-range interactions, *Phys. Rev. X* **10**, 031009 (2020).
- [12] Minh C. Tran, Andrew Y. Guo, Abhinav Deshpande, Andrew Lucas, and Alexey V. Gorshkov, Optimal state transfer and entanglement generation in power-law interacting systems, *Phys. Rev. X* **11**, 031016 (2021).
- [13] Yifan Hong and Andrew Lucas, Fast high-fidelity multi-qubit state transfer with long-range interactions, *Phys. Rev. A* **103**, 042425 (2021).
- [14] Tomotaka Kuwahara and Keiji Saito, Absence of fast scrambling in thermodynamically stable long-range interacting systems, *Phys. Rev. Lett.* **126**, 030604 (2021).
- [15] Minh C. Tran, Andrew Y. Guo, Christopher L. Baldwin, Adam Ehrenberg, Alexey V. Gorshkov, and Andrew Lucas, Lieb-Robinson light cone for power-law interactions, *Phys. Rev. Lett.* **127**, 160401 (2021).
- [16] Andrew Y. Guo, Minh C. Tran, Andrew M. Childs, Alexey V. Gorshkov, and Zhe-Xuan Gong, Signaling and scrambling with strongly long-range interactions, *Phys. Rev. A* **102**, 010401 (2020).
- [17] Chao Yin, Fast and accurate GHZ encoding using all-to-all interactions, [arXiv:2406.10336](https://arxiv.org/abs/2406.10336) [quant-ph].
- [18] Bruno Nachtergaele, Hillel Raz, Benjamin Schlein, and Robert Sims, Lieb-Robinson bounds for harmonic and anharmonic lattice systems, *Commun. Math. Phys.* **286**, 1073 (2008).
- [19] Norbert Schuch, Sarah K. Harrison, Tobias J. Osborne, and Jens Eisert, Information propagation for interacting-particle systems, *Phys. Rev. A* **84**, 032309 (2011).
- [20] J. Jünemann, A. Cadarso, D. Pérez-García, A. Bermudez, and J. J. García-Ripoll, Lieb-Robinson bounds for spin-boson lattice models and trapped ions, *Phys. Rev. Lett.* **111**, 230404 (2013).
- [21] Tomotaka Kuwahara and Keiji Saito, Lieb-Robinson bound and almost-linear light cone in interacting boson systems, *Phys. Rev. Lett.* **127**, 070403 (2021).
- [22] Chao Yin and Andrew Lucas, Finite speed of quantum information in models of interacting bosons at finite density, *Phys. Rev. X* **12**, 021039 (2022).
- [23] Jérémy Faupin, Marius Lemm, and Israel Michael Sigal, Maximal speed for macroscopic particle transport in the Bose-Hubbard model, *Phys. Rev. Lett.* **128**, 150602 (2022).
- [24] Jérémy Faupin, Marius Lemm, and Israel Michael Sigal, On Lieb–Robinson bounds for the Bose–Hubbard model, *Commun. Math. Phys.* **394**, 1011 (2022).
- [25] Tomotaka Kuwahara, Tan Van Vu, and Keiji Saito, Effective light cone and digital quantum simulation of interacting bosons, *Nat. Commun.* **15**, 2520 (2024).
- [26] Tan Van Vu, Tomotaka Kuwahara, and Keiji Saito, Optimal light cone for macroscopic particle transport in long-range systems: A quantum speed limit approach, *Quantum* **8**, 1483 (2024).
- [27] Marius Lemm, Carla Rubiliani, Israel Michael Sigal, and Jingxuan Zhang, Information propagation in long-range quantum many-body systems, *Phys. Rev. A* **108**, L060401 (2023).
- [28] Chao Yin and Andrew Lucas, Prethermalization and the local robustness of gapped systems, *Phys. Rev. Lett.* **131**, 050402 (2023).
- [29] Chao Yin, Federica M. Surace, and Andrew Lucas, Theory of metastable states in many-body quantum systems, *Phys. Rev. X* **15**, 011064 (2025).
- [30] M. B. Hastings and Xiao-Gang Wen, Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance, *Phys. Rev. B* **72**, 045141 (2005).
- [31] Sergey Bravyi, Matthew B. Hastings, and Spyridon Michalakis, Topological quantum order: Stability under local perturbations, *J. Math. Phys.* **51**, 093512 (2010).
- [32] Spyridon Michalakis and Justyna P. Zwolak, Stability of frustration-free Hamiltonians, *Commun. Math. Phys.* **322**, 277 (2013).
- [33] Matthew B. Hastings and Spyridon Michalakis, Quantization of Hall conductance for interacting electrons on a torus, *Commun. Math. Phys.* **334**, 433 (2015).
- [34] Matthew B. Hastings, An area law for one-dimensional quantum systems, *J. Stat. Mech.: Theory Exp.* **2007**, P08024 (2007).
- [35] Matthew B. Hastings and Tohru Koma, Spectral gap and exponential decay of correlations, *Commun. Math. Phys.* **265**, 781 (2006).
- [36] Bruno Nachtergaele and Robert Sims, Lieb-Robinson bounds and the exponential clustering theorem, *Commun. Math. Phys.* **265**, 119 (2006).
- [37] Chao Yin and Andrew Lucas, Heisenberg-limited metrology with perturbing interactions, *Quantum* **8**, 1303 (2024).
- [38] Chao Yin, Victor V. Albert, and Sisi Zhou, Small correlation is sufficient for optimal noisy quantum metrology, [arXiv:2408.00079](https://arxiv.org/abs/2408.00079) [quant-ph].
- [39] Adam Nahum, Sagar Vijay, and Jeongwan Haah, Operator spreading in random unitary circuits, *Phys. Rev. X* **8**, 021014 (2018).
- [40] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi, Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws, *Phys. Rev. X* **8**, 021013 (2018).
- [41] Tibor Rakovszky, C. W. von Keyserlingk, and Frank Pollmann, Dissipation-assisted operator evolution method for capturing hydrodynamic transport, *Phys. Rev. B* **105**, 075131 (2022).
- [42] Dominik S. Wild and Álvaro M. Alhambra, Classical simulation of short-time quantum dynamics, *PRX Quantum* **4**, 020340 (2023).

- [43] H. Araki, Gibbs states of the one-dimensional quantum spin chain, *Commun. Math. Phys.* **14**, 120 (1969).
- [44] Alexander Avdoshkin and Anatoly Dymarsky, Euclidean operator growth and quantum chaos, *Phys. Rev. Res.* **2**, 043234 (2020).
- [45] Chi-Fang Chen and Andrew Lucas, Operator growth bounds from graph theory, *Commun. Math. Phys.* **385**, 1273 (2021).
- [46] Jiarui Zhao, Zheng Yan, Meng Cheng, and Zi Yang Meng, Higher-form symmetry breaking at Ising transitions, *Phys. Rev. Res.* **3**, 033024 (2021).
- [47] Daniel S. Rokhsar and Steven A. Kivelson, Superconductivity and the quantum hard-core dimer gas, *Phys. Rev. Lett.* **61**, 2376- (1988).
- [48] Eddy Ardonne, Paul Fendley, and Eduardo Fradkin, Topological order and conformal quantum critical points, *Ann. Phys.* **310**, 493 (2004).
- [49] Claudio Castelnovo and Claudio Chamon, Quantum topological phase transition at the microscopic level, *Phys. Rev. B* **77**, 054433 (2008).
- [50] We note, however, that the optimal $\min(|\partial R|, |\partial S|)$ prefactor in Eq. (18) has not always been employed in the literature.
- [51] Zhiyuan Wang and Kaden R. A. Hazzard, Tightening the Lieb-Robinson bound in locally interacting systems, *PRX Quantum* **1**, 010303 (2020).
- [52] Christopher L. Baldwin, Adam Ehrenberg, Andrew Y. Guo, and Alexey V. Gorshkov, Disordered Lieb-Robinson bounds in one dimension, *PRX Quantum* **4**, 020349 (2023).
- [53] This kind of result also occurs for the usual single commutator Lieb-Robinson bounds, where the upper bounds from quasi-local Hamiltonian evolution also apply to local Hamiltonian evolution, but techniques specialized to local Hamiltonians can give a bound with slightly faster decay outside of the light cone [2].
- [54] Jeongwan Haah, Robin Kothari, and Ewin Tang, Learning quantum Hamiltonians from high-temperature Gibbs states and real-time evolutions, *Nat. Phys.* **20**, 1027 (2024).
- [55] Matthew B. Hastings, Quasi-adiabatic continuation for disordered systems: Applications to correlations, Lieb-Schultz-Mattis, and Hall conductance, [arXiv:1001.5280](https://arxiv.org/abs/1001.5280).
- [56] This definition may differ slightly from that used in the literature, where a cluster typically means a multiset of subsets of the graph. We use the term here to illustrate the connection between our algorithm and other cluster expansion methods.
- [57] Jaan Oitmaa, Chris Hamer, and Weihong Zheng, *Series Expansion Methods for Strongly Interacting Lattice Models* (Cambridge University Press, Cambridge, UK, 2006).
- [58] Baoming Tang, Ehsan Khatami, and Marcos Rigol, A short introduction to numerical linked-cluster expansions, *Comput. Phys. Commun.* **184**, 557 (2013).
- [59] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, Locality of temperature, *Phys. Rev. X* **4**, 031019 (2014).
- [60] Rahul Trivedi, Adrian Franco Rubio, and J. Ignacio Cirac, Quantum advantage and stability to errors in analogue quantum simulators, *Nat. Commun.* **15**, 6507 (2024).
- [61] Note that this sum can be handled more efficiently by storing intermediate results. However, this does not change the scaling of complexity for the whole algorithm.
- [62] Chao Yin and Andrew Lucas, Low-density parity-check codes as stable phases of quantum matter, *PRX Quantum* **6**, 030329 (2025).
- [63] M. B. Hastings, Lieb-Schultz-Mattis in higher dimensions, *Phys. Rev. B* **69**, 104431 (2004).
- [64] Tobias J. Osborne, Simulating adiabatic evolution of gapped spin systems, *Phys. Rev. A* **75**, 032321 (2007).
- [65] Sven Bachmann, Spyridon Michalakis, Bruno Nachtergaele, and Robert Sims, Automorphic equivalence within gapped phases of quantum lattice systems, *Commun. Math. Phys.* **309**, 835 (2012).
- [66] Rahul Sahay, Carolyn Zhang, Curt von Keyserlingk, and Ruben Verresen, Is the decohered toric code long-range entangled beyond its error threshold? (to appear).
- [67] Yimu Bao, Ruihua Fan, Ashvin Vishwanath, and Ehud Altman, Mixed-state topological order and the errorfield double formulation of decoherence-induced transitions, [arXiv:2301.05687](https://arxiv.org/abs/2301.05687).
- [68] Rahul Sahay, Curt von Keyserlingk, Ruben Verresen, and Carolyn Zhang, Enforced gaplessness from states with exponentially decaying correlations, [arXiv:2503.01977](https://arxiv.org/abs/2503.01977).
- [69] Anurag Anshu, Itai Arad, and David Gosset, in *52nd Annual ACM SIGACT Symposium on Theory of Computing* (2020).