# Disordered Lieb-Robinson Bounds in One Dimension

Christopher L. Baldwin,[1,*] Adam Ehrenberg,[1,2] Andrew Y. Guo,[1,2] and Alexey V. Gorshkov[1,2]

[1] *Joint Quantum Institute, National Institute of Standards and Technology (NIST)–University of Maryland, College Park, Maryland 20742, USA*

[2] *Joint Center for Quantum Information and Computer Science, National Institute of Standards and Technology (NIST)–University of Maryland, College Park, Maryland 20742, USA*

(Received 5 September 2022; accepted 11 April 2023; published 22 June 2023)

By tightening the conventional Lieb-Robinson bounds to better handle systems that lack translation invariance, we determine the extent to which "weak links" suppress operator growth in disordered one-dimensional spin chains. In particular, we prove that ballistic growth is impossible when the distribution of coupling strengths $\mu(J)$ has a sufficiently heavy tail at small $J$ and we identify the correct dynamical exponent to use instead. Furthermore, through a detailed analysis of the special case in which the couplings are genuinely random and independent, we find that the standard formulation of Lieb-Robinson bounds is insufficient to capture the complexity of the dynamics—we must distinguish between bounds that hold for *all* sites of the chain and bounds that hold for a *subsequence* of sites and we show by explicit example that these two can have dramatically different behaviors. All the same, our result for the dynamical exponent is tight, in that we prove by counterexample that there cannot exist any Lieb-Robinson bound with a smaller exponent. We close by discussing the implications of our results, both major and minor, for numerous applications ranging from quench dynamics to the structure of ground states.

DOI: 10.1103/PRXQuantum.4.020349

## I. INTRODUCTION

Lieb-Robinson (LR) bounds, named after Ref. [1], have proven to be a valuable mathematical tool for many-body physics and quantum information theory [2–14]. Conceptually, they provide hard constraints on the extent to which correlations of any type can spread through a many-body lattice system (broadly termed "operator spreading"). Numerous applications, from rigorous results on many-body ground states to lower bounds on the run time of quantum protocols, can be found in the above references.

Given their utility, LR bounds have been both generalized and specialized in multiple ways: leveraging the commutativity and graph structure of interactions [15–17], extending to bosonic systems [18–21], allowing for long-range interactions [22–26], and considering open systems [27–30], to name a few. However, one ingredient that has been noticeably absent is disorder or, more generally, a lack of translation invariance. To be fair, the conventional bounds do allow for non-translation-invariant interactions, but there have been no studies assessing the tightness of the resulting bounds (and we find that they are far from tight). Certain works have considered related topics—Refs. [31,32] study the effects of disordered local terms, albeit with uniform interactions, and Ref. [33] derives bounds for random Hamiltonians with statistical translation invariance (although see Ref. [34]). Others have calculated bounds for *specific* (often free-fermion-integrable) systems [35–38]. While interesting on their own merits, none of these quite address the question with which we concern ourselves here—to what extent is operator spreading, as constrained by LR bounds, necessarily suppressed by non-translation-invariant interactions?

Non-translation-invariant systems are known to exhibit a variety of phenomena not found in their translation-invariant counterparts. Examples include spin glass phases (both static and dynamic) [39,40], localization [41,42], and Griffiths effects [43]. Disordered fermionic models have been used to help understand quantum dots and strongly correlated metals as well [44–46]. Of particular (and somewhat controversial) interest is the phenomenon of many-body localization (MBL) [47–50], the existence of which remains under debate [51–55]. Given the challenges inherent in studying not only MBL but disordered quantum many-body systems in general, it is all the more important to identify rigorous constraints such as those that LR bounds supply (although to be clear, our

*cbaldwi3@umd.edu

results and LR bounds in general are not strong enough to resolve the questions surrounding MBL, as we explain in Sec. II).

In the present work, we initiate the study of non-translation-invariant LR bounds by considering arguably the simplest (but still quite rich) situation: one-dimensional (1D) chains with nearest-neighbor interactions. An essential feature of such systems is the importance of "weak links," i.e., atypically weak interactions. We develop the machinery for analyzing these systems, prove that the LR bounds thus obtained are in a certain sense optimal, and use the results to place constraints on various physical properties and processes.

Section II summarizes our results in conceptual terms. We aim for it to be sufficiently self-contained so that readers primarily interested in our conclusions and willing to forgo the derivations should be able to read Sec. II on its own. Section III then gives precise definitions of all quantities involved in our analysis. Section IV derives LR bounds for general non-translation-invariant systems and Sec. V specializes to the case of random couplings. Lastly, Sec. VI discusses some implications of our results, related in particular to quench dynamics, topological order, heating rates, ground-state correlations, and machine learning of local observables.

## II. SUMMARY OF RESULTS

We determine the extent to which operator spreading in 1D nearest-neighbor chains is suppressed by weak links. To be precise, we consider arbitrary $d$-state degrees of freedom ("qudits") interacting via any Hamiltonian of the form $H(t) = \sum_l H_l(t)$, where the sum is over links of the chain and $H_l(t)$ acts on the two sites connected by link $l$ (although our analysis in fact allows for arbitrary local terms as well [56]) See Fig. 1. Unlike previous works, we assume that $\|H_l(t)\| \leq J_l$, where $J_l$ *varies* from link to link ($\|\cdot\|$ denotes the operator norm throughout). The "weak" links are those on which $J_l$ is much smaller than the typical value.

For arbitrary local operators $A_0$ and $B_r$ supported on sites 0 and $r$, respectively, and with $A_0^t$ denoting the evolution of $A_0$ over time $t$, our goal is to bound the quantity $\|[A_0^t, B_r]\|$ as tightly as possible, making use of the weak links in the set $\{J_l\}$. We focus in particular on the asymptotic behavior at large $r$ and $t$. The commutator $[A_0^t, B_r]$ is a standard and useful measure of operator spreading (as we describe in Sec. III B) but it is by no means the only such measure. In fact, most of our calculations involve a different quantity, from which bounds on the commutator easily follow.

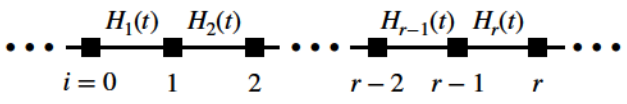


FIG. 1. The geometry of the systems studied in the present work, i.e., 1D chains with nearest-neighbor interactions. Sites are indicated by black squares, labeled by $i$. Terms of the Hamiltonian are indicated by solid lines, with $H_l(t)$ acting on the two sites connected by that link. Given a set $\{J_l\}$, we require that $\|H_l(t)\| \leq J_l$.

LR bounds are best viewed as applying to *families* of Hamiltonians. For each set of couplings $\{J_l\}$, we use $\mathcal{H}_J$ to denote the set of all Hamiltonians as described above consistent with $\{J_l\}$ (i.e., $\|H_l(t)\| \leq J_l$ for all links $l$ at all times $t$). Our bounds are uniform among $H \in \mathcal{H}_J$, in the sense that they make no reference to any property of the Hamiltonian beyond the couplings $\{J_l\}$. Thus, while the results are extremely general, they may not be very tight for one specific system. The "tightness" of LR bounds referred to in this paper is instead the existence of *some* $H \in \mathcal{H}_J$ that saturates the bound.

As a consequence of their generality, LR bounds have little to say regarding, e.g., the existence of MBL—the complete lack of operator spreading in strongly disordered

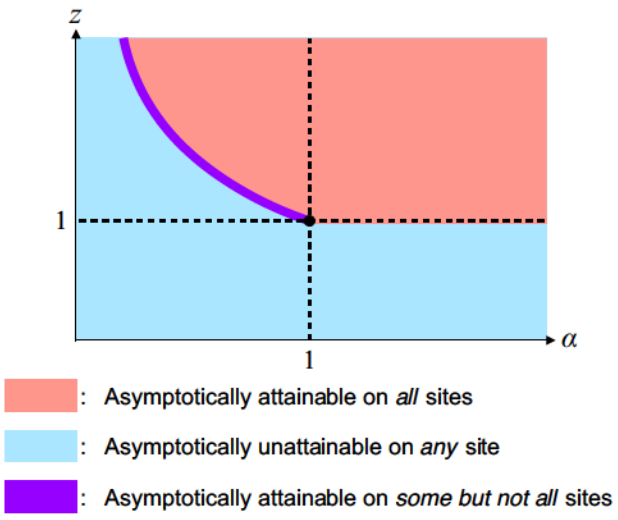


FIG. 2. The possibility of a Hamiltonian causing operator spreading at dynamical exponent $z$, i.e., in a time growing asymptotically as $t(r) \propto r^z$ to reach a large distance $r$. The horizontal axis is the exponent $\alpha$ characterizing the number of weak links in the couplings $\{J_l\}$—the fraction of links between sites 0 and $r$ having $J_l \leq J$ is assumed to go as $J^\alpha$ at small $J$ and large $r$. The vertical axis is $z$. The blue region is where no Hamiltonian with such $\alpha$ can reach *any* large-distance site at such $z$. The red region is where a Hamiltonian exists with such $\alpha$ that reaches *every* large-distance site at (or faster than) such $z$. The boundary between the two is given by $z_c(\alpha) = \max[1/\alpha, 1]$. In the special case of independent random couplings, the $\alpha > 1$ portion of the boundary is included in the red region, while the $\alpha < 1$ portion (shown in purple) is where a Hamiltonian exists that reaches a *subsequence* of sites faster than $z$ but where no Hamiltonian can reach *every* site at such $z$. Lastly, at point $(\alpha, z) = (1, 1)$, no Hamiltonian can reach every site ballistically but it is unknown whether any Hamiltonian can reach some sites ballistically.

time-independent systems [47–50]. The same applies to other phenomena involving slow dynamics under a fixed Hamiltonian, such as activated processes in spin glasses [57–59]. Those $H \in \mathcal{H}_J$ that saturate our bounds will instead tend to be highly time-*dependent* Hamiltonians, specifically designed to transmit information and better viewed as quantum circuits. These are the types of systems for which LR bounds give a reasonably full picture.

Our first result, which holds for any possible set of couplings $\{J_l\}$, is an improvement on the conventional LR bound (see also Ref. [60]). Whereas the standard analysis leads to the result

$$\left\|\left[A_0^t, B_r\right]\right\| \leq C\|A_0\|\|B_r\|\left(\prod_{l=1}^{r} 4J_l\right)\frac{t^r}{r!}, \tag{1}$$

we show that one further has

$$\left\|\left[A_0^t, B_r\right]\right\| \leq C\|A_0\|\|B_r\| \min_{\lambda}\left[\left(\prod_{l\in\lambda} 4J_l\right)\frac{t^{|\lambda|}}{|\lambda|!}\right], \tag{2}$$

where the minimization is over all subsets $\lambda$ of the links between sites 0 and $r$, with $|\lambda|$ denoting the size of the subset. Even though much of what follows is concerned with the asymptotic behavior, Eq. (2) holds for all $r$ and $t$. In these equations and throughout the entire paper, we use $C$ to denote any constant that does not depend on $r$ or $t$ and for which the precise value is irrelevant to our conclusions. Its value often changes between expressions (and primes and subscripts differentiate such constants within the same expression).

As an explicit example of the improvement that Eq. (2) can provide over Eq. (1), consider a four-site chain having $J_1 = J_2 = 1/4$ and $J_3 = 1/4000$. For operator spreading from site 0 to site 3, Eq. (1) gives the bound $t^3/6000$, while Eq. (2) gives additional bounds such as $t/1000$ (corresponding to $\lambda = \{3\}$). Note that the former reaches 1 at $t = 10\sqrt[3]{6} \approx 18$, whereas the latter does not reach 1 until $t = 1000$, later by multiple orders of magnitude.

After establishing Eq. (2), we next derive more explicit bounds by considering the "empirical distribution" $\mu_r(J)$, defined (for a given set $\{J_l\}$) as the fraction of links between sites 0 and $r$ having $J_l \leq J$:

$$\mu_r(J) \equiv \frac{1}{r}\sum_{l=1}^{r} \delta_{J_l \leq J}, \tag{3}$$

where $\delta_{J_l\leq J} \equiv 1$ if $J_l \leq J$ and 0 otherwise. We assume that $\mu_r(J)$ converges as $r \to \infty$ (in a sense defined in Sec. IV C) to a function $\mu(J)$ and that the latter behaves as a power law at small $J$: $\mu(J) \sim \mu_0 J^\alpha$ with $\alpha > 0$. The exponent $\alpha$ characterizes the prevalence of weak links, with smaller $\alpha$ implying more weak links. Our analysis in fact applies for more general forms of $\mu(J)$ (as we discuss in Sec. IV C) but the power-law behavior is particularly convenient and representative. Note that the convergence of $\mu_r(J)$ to $\mu(J)$ does not assume anything regarding the arrangement of weak links in space—many of our results hold regardless of where the weak links are located.

An essential feature of an LR bound is the shape of the "front," i.e., the space-time curve $t(r)$ that separates the region in which the bound is small from that in which the bound is large (and thus vacuous). The dynamical exponent $z$ and generalized LR velocity $v$ are defined by the asymptotic behavior $vt(r) \sim r^z$ at large $r$. Keep in mind that $v$ has units of velocity only when $z = 1$—we stick to the term "generalized velocity" for $z \neq 1$ (see also Ref. [61]). Whereas the conventional LR bound, Eq. (1), has a ballistic front ($z = 1$) for all $\alpha > 0$, we show that the improved bound, Eq. (2), instead has dynamical exponent

$$z_c(\alpha) = \max\left[\alpha^{-1}, 1\right]. \tag{4}$$

The curve $z_c(\alpha)$ is sketched in Fig. 2.

One consequence of the above is that there cannot exist any $H \in \mathcal{H}_J$ for which the operator-spreading front grows with a dynamical exponent $z < z_c(\alpha)$. In particular, it is impossible to have a ballistic front if $\alpha < 1$. We shade this region blue in Fig. 2 and label it as "asymptotically unattainable on any site"—it is impossible to construct a Hamiltonian having that value of $\alpha$ which, at large distances, reaches any site at that value of $z$.

On the other hand, we also identify an $H \in \mathcal{H}_J$ for which the front grows faster than any $z > z_c(\alpha)$, again requiring only that $\mu_r(J) \to \mu(J)$ in a suitably strong sense. The Hamiltonian is rather straightforward, consisting simply of a sequential series of SWAP gates as shown in Fig. 3. Once all gates have been applied, $A_0^t$ is supported on site $r$ and thus will generically fail to commute (by an $r$-independent amount) with $B_r$. In order to satisfy $\|H_l(t)\| \leq J_l$, the total run time of the circuit is proportional to $\sum_{l=1}^{r} 1/J_l$—analysis of this sum gives the behavior of the front.

We shade the region $z > z_c(\alpha)$ red in Fig. 2 and label it as "asymptotically attainable on all sites"—it is possible (and we do so) to construct a Hamiltonian having that value of $\alpha$ which reaches every large-distance site at that value of $z$. In this sense, Eq. (4) is *the* dynamical exponent for a given $\alpha$, i.e., Eq. (4) is tight.

It is rather striking that our result for $z_c(\alpha)$ agrees exactly with the value predicted on physical grounds in Ref. [62]. The results of Ref. [62] are based on a coarse-grained description of 1D disordered systems, in which it is postulated that a region $l$ can be characterized by an effective "growth rate" $\Gamma_l$, setting the rate at which operators spread across the region. The authors assume that $\Gamma_l$ is power-law distributed with exponent $\alpha$ (although they work with

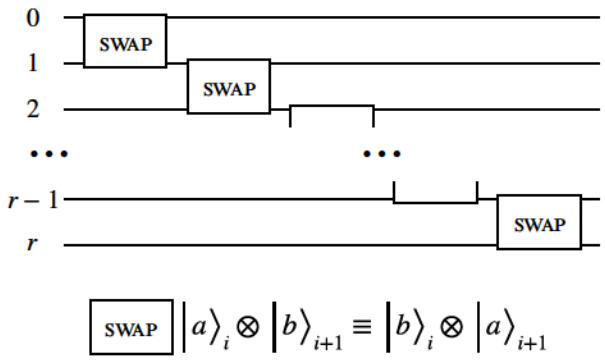


FIG. 3. The circuit diagram of a simple protocol that transfers the state on site 0 to site $r$ (here, the sites are arranged vertically and the solid lines indicate the absence of any unitary). The definition of the SWAP gate acting on sites $i$ and $i+1$ is given at the bottom of the figure—$a$ and $b$ label basis states of the local Hilbert space (each taking values in $\{1,\ldots,d\}$).

the probability *density* having exponent $\alpha - 1$) and ultimately deduce precisely Eq. (4). One important difference is that we consider *microscopic* weak links rather than effective weak links emerging at long wavelengths, but this nonetheless provides a rigorous foundation for many of the concepts at work in Ref. [62]. It would be of great interest to investigate whether the other phenomena discussed in Ref. [62], such as entanglement growth and transport, can be placed on similar rigorous grounds.

Returning to Fig. 2, the situation becomes much more complicated on the boundary $z = z_c(\alpha)$, i.e., when considering dynamics on the scale set by $z_c(\alpha)$, both for our LR bounds and for our example Hamiltonians. We demonstrate this by a detailed analysis of the special case in which the couplings are drawn independently from a literal probability distribution $\mu(J)$. After proving that $\mu_r(J) \to \mu(J)$ in the required sense with probability 1 and thus that the portion of Fig. 2 away from $z_c(\alpha)$ does indeed hold, we find that the generic behavior on the boundary cannot be described by a single bound—we must introduce (at least) *two types* of LR bounds:

(a) "Almost-always" (a.a.) bounds are those that hold for all sites $r$, excepting at most a finite number of sites. In other words, there exists a distance $R$ such that the bound holds for all $r > R$.
(b) "Infinitely-often" (i.o.) bounds are those that hold for an infinite subsequence of sites $\{r_k\}$ but need not hold outside of those sites. In other words, for any distance $R$, there exists some site $r > R$ that is subject to the bound.

One can imagine situations in which either of the above two bounds is more relevant. For example, suppose that Alice is manipulating one site of a spin chain and wants to be confident that her actions do not disturb distant regions in a certain amount of time. In this case, a.a. bounds provide the desired guarantee. On the other hand, suppose that Bob intends to transmit a signal along the spin chain. If it is important that his signal reach every site faster than a certain rate, then i.o. bounds place the heaviest restrictions on what can be achieved.

In our case (still at $z = z_c(\alpha)$ and still assuming independent random couplings), we find different behaviors depending on how $\alpha$ compares to 1. If $\alpha > 1$, the results are straightforward: our example Hamiltonian has a ballistic front that spreads to every site with finite velocity (note that $z_c(\alpha) = 1$ for $\alpha > 1$). We include this portion of the boundary with the red region in Fig. 2 to indicate that it is also "asymptotically attainable on all sites" (albeit with a maximum allowed velocity).

However, we show that if $\alpha < 1$, then an i.o. bound having $z = z_c(\alpha)$ and arbitrarily small (generalized) velocity holds, while concurrently, our example Hamiltonian does reach a subsequence $\{r_k\}$ asymptotically faster than $z_c(\alpha)$ (hence no such a.a. bound can hold). Both statements hold with probability 1. Thus it is impossible to have a front that reaches *every* site at dynamical exponent $z_c(\alpha)$ but it is possible (and we do so) to construct a Hamiltonian that reaches a *subsequence* of sites at $z_c(\alpha)$. We draw this portion of the boundary purple in Fig. 2 and accordingly label it "asymptotically attainable on some but not all sites." Keep in mind that we prove this final statement only for the special case of independent random couplings. Nonetheless, it is a highly nontrivial example that makes clear the importance of distinguishing between a.a. and i.o. bounds.

Interestingly, the lone point $(\alpha, z) = (1, 1)$ is the only portion of Fig. 2 in which we are unable to give a definite answer. An i.o. bound with vanishing velocity still holds but the front in our example Hamiltonian is now subballistic for every site. It may be that a more complicated Hamiltonian exists that does reach a subsequence $\{r_k\}$ at finite velocity, yet it may instead be that a more sophisticated mathematical technique can produce an a.a. bound with vanishing velocity. Further investigation is clearly warranted.

Lastly, we discuss the implications of our results for various applications. The LR bounds themselves have physical content—the statement derived here that ballistic spreading is impossible for $\alpha < 1$ can be considered an application in and of itself. That said, our results have broader consequences as well. Applications can roughly be grouped into two classes: those that follow from the existence of the front and those that follow from the "tail" (i.e., the rapid decay of the LR bound at large distances outside the front). Since we obtain a significantly altered front for $\alpha < 1$, our results have a qualitative impact on the former class. However, while we do find a more complicated tail than in the conventional bound, the behavior at the largest distances turns out to be unmodified and thus our results have only a minor impact on the latter class.

In the remainder of the paper, we make precise and prove the above statements. Section III establishes notation and the formalism within which we work. Section IV, after reviewing the conventional LR bound, derives Eq. (2) for generic non-translation-invariant systems and then makes use of the empirical distribution $\mu_r(J)$ to derive Fig. 2. Section V considers the case of independent random couplings $\{J_l\}$ in more detail, first proving that the requirements of the preceding section are met and then examining behavior on the boundary $z_c(\alpha)$, with a particular focus on the distinction between a.a. and i.o. behavior. Lastly, Section VI discusses the consequences of the above for various applications.

## III. DEFINITIONS AND NOTATION

### A. Geometry

In this work, we consider an $N$-site lattice in 1D, where each site hosts a $d$-state degree of freedom. In other words, the Hilbert space is a tensor product of $N$ local $d$-dimensional Hilbert spaces. Let $\Omega$ denote the set of all $N$ sites and $\Lambda$ denote the set of all $N-1$ links. Here, we consider only *nearest-neighbor* Hamiltonians on this lattice, i.e., Hamiltonians of the form $H(t)=\sum_{l\in\Lambda} H_l(t)$, where $H_l(t)$ is supported only on the sites connected by link $l$ (although, as noted above, our results hold for Hamiltonians with arbitrary local terms as well [56]). These features are illustrated in Fig. 1.

Given a set of couplings $\{J_l\}_{l\in\Lambda}$, let $\mathcal{H}_J$ be the family of all nearest-neighbor Hamiltonians for which

$$\left\|H_l(t)\right\| \le J_l. \tag{5}$$

We even allow for a nonvanishing fraction of $\{J_l\}$ to be infinite, meaning that there is no restriction on the corresponding terms. Pick an operator of interest $A_0$ supported only on site 0 and similarly $B_r$ on site $r$ (although it is clear from the proof that $A_0$ can be supported anywhere to the left of 0 and $B_r$ anywhere to the right of $r$ as well). For any $H\in\mathcal{H}_J$, let $A_0^t$ be the time evolution of $A_0$, i.e., the solution to

$$\partial_t A_0^t = i\big[H(t), A_0^t\big], \qquad A_0^0 = A_0 \tag{6}$$

(see also Ref. [63]). Even though $A_0$ is supported on site 0, $A_0^t$ will (barring trivial cases) be supported throughout the entire chain for any $t>0$. The purpose of LR bounds is to place a bound on the quantity

$$D(r,t) \equiv \left\|\left[A_0^t, B_r\right]\right\|, \tag{7}$$

*which holds uniformly for all* $H\in\mathcal{H}_J$. Since only the "portion" of $A_0^t$ that acts nontrivially on site $r$ can possibly fail to commute with $B_r$ (see Eq. (13) below), LR bounds considered as functions of $r$ and $t$ constrain the extent to which local operators "spread" throughout the system.

It is important to note that the Hamiltonians in $\mathcal{H}_J$ can have arbitrary time dependence, as long as Eq. (5) is obeyed at all times. Thus it is perhaps more informative to refer to any individual $H\in\mathcal{H}_J$ as a "protocol," since it can be a quantum circuit designed to perform a specific task. As discussed in Sec. II, this distinction sheds light on the limitations of LR bounds.

The choice to use the operator norm in Eq. (7) has long been standard, as it enters naturally in many applications [1,4,5,7]. There are situations in which alternative norms—in particular, the Frobenius norm, defined as $\|O\|_2 \equiv [d^{-N}\mathrm{Tr}O^\dagger O]^{1/2}$—may be more relevant and might behave quite differently [64]. However, the operator norm is itself an upper bound on a wide family of norms including Frobenius (see Appendix A). Furthermore, the transfer protocol shown in Fig. 3 leads to a commutator $[A_0^t, B_r]$ that is $O(1)$ using any of these norms. Thus, we exclusively consider the operator norm in this work and the bounds obtained are automatically tight (at least regarding the dynamical exponent) for the other norms as well.

### B. Basis strings

The set of Hermitian operators acting on the Hilbert space is itself a real vector space and thus we can express any operator as a linear combination of certain basis operators. First, consider a single site $i$ and pick a Hermitian basis $\{X_i^{(\nu_i)}\}$ ($\nu_i\in\{0,1,\ldots,d^2-1\}$). For the entire chain, we use the tensor-product basis $\{X^{(\nu)}\}$:

$$X^{(\nu)} \equiv \bigotimes_{i=1}^{N} X_i^{(\nu_i)}, \tag{8}$$

where $\nu\equiv(\nu_1,\ldots,\nu_N)$. We assume (without loss of generality) that $\{X_i^{(\nu_i)}\}$ is chosen to be orthonormal with respect to the trace product, meaning that the tensor-product basis is orthonormal as well:

$$\begin{aligned} d^{-N}\mathrm{Tr}X^{(\nu)}X^{(\nu')} &= \prod_{i=1}^{N} d^{-1}\mathrm{Tr}X_i^{(\nu_i)}X_i^{(\nu_i')} \\ &= \prod_{i=1}^{N}\delta_{\nu_i\nu_i'} = \delta_{\nu\nu'}. \end{aligned} \tag{9}$$

We also take $X_i^{(0)}$ to be the identity $I_i$. Beyond this, any choice of basis works equally well.

We often refer to the basis elements as "strings" and define the support of a string to be the set of sites on which it does not have the identity:

$$\mathrm{supp}X^{(\nu)} \equiv \big\{i : \nu_i \neq 0\big\}. \tag{10}$$

An important superoperator acting on the space of Hermitian operators is that which projects onto basis strings

for which the support contains site $i$, i.e., strings that act nontrivially on site $i$. We denote this superoperator by $\mathcal{P}_i$:

$$\mathcal{P}_i X^{(\nu)} \equiv \left(1 - \delta_{\nu_i 0}\right) X^{(\nu)}. \tag{11}$$

Similarly, for any subset of sites $\omega \subseteq \Omega$, we define $\mathcal{P}_\omega$ to project onto basis strings that act nontrivially somewhere (not necessarily everywhere) within $\omega$. A useful inequality (see Appendix A) is that for any $\omega$ and any operator $O$,

$$\left\|\mathcal{P}_\omega O\right\| \leq 2\left\|O\right\|. \tag{12}$$

Also note that $[A_0^t, B_r] = [\mathcal{P}_r A_0^t, B_r]$, and so for $D(r,t)$ in Eq. (7), we have the trivial bound

$$D(r,t) \leq 2\left\|\mathcal{P}_r A_0^t\right\|\left\|B_r\right\|. \tag{13}$$

In what follows, we focus on bounding $\|\mathcal{P}_r A_0^t\|$, with a bound on $D(r,t)$ following automatically by Eq. (13).

The next important superoperator is the generator of time evolution under $H(t)$,

$$\mathcal{L}(t)O \equiv i\big[H(t), O\big], \tag{14}$$

and so [Eq. (6)]

$$\partial_t A_0^t = \mathcal{L}(t) A_0^t. \tag{15}$$

We also need the generator corresponding to a subset of terms in the Hamiltonian. For any subset of links $\lambda \subseteq \Lambda$, define

$$\mathcal{L}_\lambda(t)O \equiv i\sum_{l\in\lambda}\big[H_l(t), O\big]. \tag{16}$$

Clearly, $\mathcal{L}(t) = \sum_{l\in\Lambda}\mathcal{L}_l(t)$.

Denote the evolution superoperator itself by $\mathcal{U}(t)$, i.e., $A_0^t \equiv \mathcal{U}(t)A_0$ (and define $\mathcal{U}_\lambda(t)$ analogously). We can express the action of $\mathcal{U}(t)$ (and $\mathcal{U}_\lambda(t)$) in terms of a time-ordered exponential:

$$\begin{aligned}\mathcal{U}(t)A_0 &= \left(\mathcal{T}e^{\int_0^t ds\mathcal{L}(s)}\right)A_0 \\ &= \left(\mathcal{T}e^{i\int_0^t dsH(s)}\right)A_0\left(\mathcal{T}e^{i\int_0^t dsH(s)}\right)^\dagger,\end{aligned} \tag{17}$$

where $\mathcal{T}$ denotes time-ordering (note the ordering in the bottom line of Eq. (17)—earlier times appear *inside* later times). Note that $\|\mathcal{U}(t)A_0\| = \|A_0\|$.

Lastly, suppose that $\lambda \subseteq \Lambda$ contains every link that connects a subset of sites $\omega$ to its complement $\Omega/\omega$ (see Fig. 4). It is intuitively clear that evolution under $\mathcal{L}_{\Lambda/\lambda}(t)$ alone cannot transform a basis operator that acts trivially on $\omega$ into one that acts nontrivially, or vice versa. Put precisely,

$$\mathcal{P}_\omega \mathcal{U}_{\Lambda/\lambda}(t) = \mathcal{U}_{\Lambda/\lambda}(t)\mathcal{P}_\omega. \tag{18}$$

We give a proof of Eq. (18) in Appendix B.

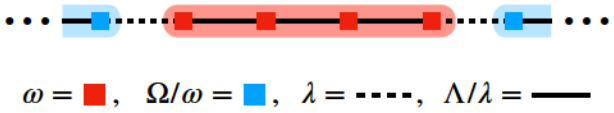


FIG. 4. The visual interpretation of Eq. (18). $\omega$ consists of the sites in red and $\Omega/\omega$ the sites in blue. $\lambda$ consists of the dashed links and $\Lambda/\lambda$ the solid links. Evolution under only $H_{\Lambda/\lambda}$ cannot transform a string having the identity on $\omega$ into one with a nonidentity element, nor vice versa.

### C. Types of Lieb-Robinson bounds

To reiterate, the purpose of LR bounds is to place an upper limit on $\|\mathcal{P}_r A_0^t\|$ (and thus $D(r,t)$) that applies to every $H \in \mathcal{H}_J$ simultaneously. The bounds we construct are of the form

$$\left\|\mathcal{P}_r A_0^t\right\| \leq \left\|A_0\right\| f\left[r^\gamma\left(1 - \frac{vt}{r^z}\right)\right], \tag{19}$$

where $\gamma > 0$ and the function $f(x)$ decays to zero as $x \to \infty$ and remains finite (or even diverges) as $x \to -\infty$. A simple and common example is $f(x) = \exp[-\kappa x]$ for some $\kappa > 0$. Although we do not indicate so explicitly, note that all quantities here are functions of the couplings $\{J_l\}$ defining $\mathcal{H}_J$.

In the large-$r$ and large-$t$ limit, one identifies two important features from Eq. (19):

(a) There is a "front" defined by $vt = r^z$. For $vt > r^z$, the right-hand side of Eq. (19) is large (and thus the bound is vacuous), while for $vt < r^z$, $\|\mathcal{P}_r A_0^t\|$ must be small. Thus, the space-time curve $vt = r^z$ constitutes an envelope that constrains the expansion of $A_0^t$. We refer to $z$ as the "dynamical exponent" and $v$ as the "generalized velocity" of the bound (only when $z = 1$ do we speak simply of the "velocity"). Of particular interest are the *largest* value of $z$ and *smallest* value of $v$ for which a bound as in Eq. (19) holds.

(b) At fixed $t$, there is the "tail" behavior as $r \to \infty$, characterized by the exponent $\gamma$ and the large-$x$ behavior of $f(x)$ (e.g., exponential or power-law). Even though $\|\mathcal{P}_r A_0^t\|$ need never be identically zero, the tail describes how rapidly it must decay at large distances.

As mentioned in Sec. II, we find it necessary to distinguish between different types of LR bounds, based on whether a statement such as Eq. (19) holds for *all* sites or merely a *subsequence* of sites:

(a) We call Eq. (19) an "almost-always" (a.a.) bound if there exists an $R$ such that it holds for all $r > R$.

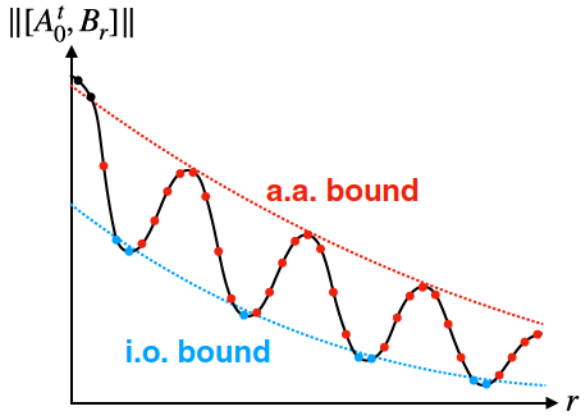


FIG. 5. A sketch of a possible spatial profile $\|[A_0^t, B_r]\|$ as a function of $r$, indicated by dots (the black line merely connects the points). The red dashed line shows an a.a. bound consistent with the profile—the red and blue dots indicate points contained within the bound. There are implied to be only a finite number of black dots. The blue dashed line shows an i.o. bound—the blue dots indicate the subsequence $\{r_k\}$ to which the bound applies.

This is the sense in which all past works of which we are aware have derived and discussed LR bounds (although see Ref. [65]).

(b) We call Eq. (19) an "infinitely-often" (i.o.) bound if for every $R$, it holds for some $r > R$. This is equivalent to saying that there exists a subsequence $\{r_k\}_{k=1}^{\infty}$ on which the bound holds.

Figure 5 illustrates the distinction between the two, showing a hypothetical curve $\|[A_0^t, B_r]\|$ versus $r$ alongside consistent a.a. and i.o. bounds. Note that a.a. bounds are automatically i.o. bounds but not vice versa. The distinction is likely unimportant for translation-invariant systems (although we are not aware of any works that compare the two to begin with) but it turns out to be essential in non-translation-invariant systems—as mentioned in Sec. II, there will be situations in which we can derive i.o. bounds while simultaneously proving that no corresponding a.a. bound can exist.

## IV. GENERAL BOUND FOR NON-TRANSLATION-INVARIANT SYSTEMS

### A. Review of the conventional bound

We first review the standard Lieb-Robinson bound. Start with a Hamiltonian $H \in \mathcal{H}_J$:

$$H(t) = \sum_{l\in\Lambda} H_l(t), \qquad \|H_l(t)\| \le J_l. \tag{20}$$

For any link $l$, pass to the interaction picture with respect to all other terms in the Hamiltonian by defining

$$A_{0I}^t \equiv \mathcal{U}_{\Lambda/l}(t)^\dagger A_0^t. \tag{21}$$

The equation of motion for $A_{0I}^t$ [see Eqs. (15) and (17)] is

$$\begin{aligned}\partial_t A_{0I}^t &= -\mathcal{U}_{\Lambda/l}(t)^\dagger \mathcal{L}_{\Lambda/l}(t) A_0^t + \mathcal{U}_{\Lambda/l}(t)^\dagger \mathcal{L}(t) A_0^t \\ &= \mathcal{U}_{\Lambda/l}(t)^\dagger \mathcal{L}_l(t) A_0^t,\end{aligned} \tag{22}$$

from which it follows that

$$A_{0I}^t = A_0 + \int_0^t ds\, \mathcal{U}_{\Lambda/l}(s)^\dagger \mathcal{L}_l(s) A_0^s. \tag{23}$$

Since we are considering a 1D chain, $l$ is the only link that connects the sites on its right (denoted $> l$) to the sites on its left (denoted $< l$). Thus, using Eq. (18),

$$\mathcal{P}_{>l} A_{0I}^t = \mathcal{P}_{>l} A_0 + \int_0^t ds\, \mathcal{U}_{\Lambda/l}(s)^\dagger \mathcal{P}_{>l} \mathcal{L}_l(s) A_0^s. \tag{24}$$

Furthermore, $\mathcal{P}_{>l} A_{0I}^t = \mathcal{U}_{\Lambda/l}(t)^\dagger \mathcal{P}_{>l} A_0^t$. Taking the norm then gives

$$\left\|\mathcal{P}_{>l} A_0^t\right\| \le \left\|\mathcal{P}_{>l} A_0\right\| + \int_0^t ds \left\|\mathcal{P}_{>l} \mathcal{L}_l(s) A_0^s\right\|. \tag{25}$$

Recall that $\mathcal{L}_l(s) A_0^s \equiv i[H_l(s), A_0^s]$. Since $H_l(s)$ is supported only on link $l$,

$$\mathcal{L}_l(s) A_0^s = \mathcal{L}_l(s) \mathcal{P}_{>l-1} A_0^s. \tag{26}$$

Together with Eq. (12), we thus have

$$\begin{aligned}\left\|\mathcal{P}_{>l} \mathcal{L}_l(s) \mathcal{P}_{>l-1} A_0^s\right\| &\le 2 \left\|\mathcal{L}_l(s) \mathcal{P}_{>l-1} A_0^s\right\| \\ &\le 4J_l \left\|\mathcal{P}_{>l-1} A_0^s\right\|\end{aligned} \tag{27}$$

and Eq. (25) becomes

$$\left\|\mathcal{P}_{>l} A_0^t\right\| \le \left\|\mathcal{P}_{>l} A_0\right\| + 4J_l \int_0^t ds \left\|\mathcal{P}_{>l-1} A_0^s\right\|. \tag{28}$$

Taking $r > 0$ for concreteness and supposing that $l$ lies between sites 0 and $r$, $\|\mathcal{P}_{>l} A_0\| = 0$ and we are left with

$$\left\|\mathcal{P}_{>l} A_0^t\right\| \le 4J_l \int_0^t ds \left\|\mathcal{P}_{>l-1} A_0^s\right\|. \tag{29}$$

To obtain a closed-form bound on $\|\mathcal{P}_r A_0^t\|$, first note that $\|\mathcal{P}_r A_0^t\| = \|\mathcal{P}_r \mathcal{P}_{\ge r} A_0^t\| \le 2\|\mathcal{P}_{\ge r} A_0^t\|$ [Eq. (12)], and then use Eq. (29) iteratively: bound $\|\mathcal{P}_{\ge r} A_0^t\|$ in terms of $\|\mathcal{P}_{\ge r-1} A_0^s\|$, then bound the latter in terms of $\|\mathcal{P}_{\ge r-2} A_0^{s'}\|$, and so on until reaching the origin (at which point do not

introduce $\mathcal{P}_{\geq 0}$—simply use that $\|A_0^s\| = \|A_0\|$). The result is

$$\left\|\mathcal{P}_r A_0^t\right\| \leq 2\left\|A_0\right\| \left(\prod_{l=1}^{r} 4J_l\right) \frac{t^r}{r!}. \tag{30}$$

From here, one usually takes all $J_l$ to equal a common value $J$. This gives (using that $r! \geq (r/e)^r$)

$$\begin{aligned}\left\|\mathcal{P}_r A_0^t\right\| &\leq 2\left\|A_0\right\| \exp\left[r \log \frac{4eJt}{r}\right] \\ &\leq 2\left\|A_0\right\| \exp\left[-r\left(1 - \frac{4eJt}{r}\right)\right].\end{aligned} \tag{31}$$

Equation (31) is of the form in Eq. (19), with $z = 1$, $\gamma = 1$, and $v = 4eJ$: we have a ballistic (a.a.) LR bound with an exponential tail and a velocity of order $J$.

One can certainly use Eq. (30) for non-translation-invariant $\{J_l\}$ as well. Yet it is easy to see that the result might be rather weak. Suppose that at large $r$, the empirical distribution of couplings [Eq. (3)] approaches a function $\mu(J)$ (in some sufficiently strong sense—we are only reasoning schematically for the moment). Then,

$$\prod_{l=1}^{r} 4J_l = \exp\left[\sum_{l=1}^{r} \log 4J_l\right] \sim \exp\left[r \int d\mu(J) \log 4J\right]. \tag{32}$$

As long as $\log J$ is integrable with respect to $\mu(J)$ (which, note, is true if $\mu(J) \sim J^\alpha$ for *any* $\alpha > 0$), we still obtain a ballistic LR bound:

$$\left\|\mathcal{P}_r A_0^t\right\| \leq 2\left\|A_0\right\| \exp\left[-r\left(1 - \frac{4e\overline{J}t}{r}\right)\right], \tag{33}$$

where $\log 4\overline{J} \equiv \int d\mu(J) \log 4J$. Clearly, this is a much weaker claim than Fig. 2. As we show in the following, taking the small-$J$ links more seriously proves that the actual dynamics must be sub-ballistic for any $\alpha < 1$.

### B. Improvements via integrating out links

We improve on the conventional bound by passing to a further interaction picture with respect to all "large" terms in the Hamiltonian, namely, all $H_l(t)$ for which the norm exceeds some threshold $\epsilon$ (see Refs. [66,67]). On the one hand, the remaining terms have a larger support in this interaction picture and we do not try to describe their structure within that support. Yet in return, *every* remaining term has norm less than $\epsilon$ and so no dynamics can occur on any scale faster than $\epsilon^{-1}$. The latter effect, which suppresses operator spreading, turns out to be the dominant one when there is a sufficient number of weak links. We sketch the situation in Fig. 6.

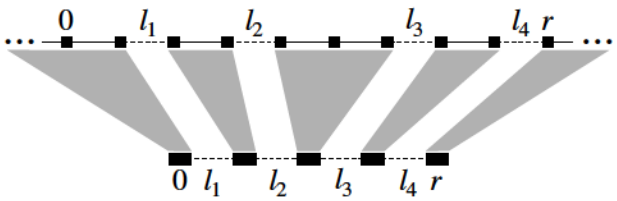


FIG. 6. The transformation from the original picture (top line) to the interaction picture with respect to "large" terms (bottom line). The dashed links (labeled $l_1$ through $l_4$) are those on which $\|H_l(t)\| \leq \epsilon$ and the solid links those on which $\|H_l(t)\| > \epsilon$. The squares on the top line represent individual sites but the rectangles on the bottom line represent the collection of all corresponding sites connected by solid links—the transformed interaction $\widetilde{H}_{l_t}^t$ is supported on the entire neighboring rectangles [but no further by virtue of Eq. (18)].

We now make this argument precise. Pick any subset of links $\lambda \equiv \{l_1, \ldots, l_n\}$ lying between 0 and $r$. Pass to the interaction picture with respect to $H_{\Lambda/\lambda}$:

$$\widetilde{A}_0^t \equiv \mathcal{U}_{\Lambda/\lambda}(t)^\dagger A_0^t, \quad \widetilde{H}_l(t) \equiv \mathcal{U}_{\Lambda/\lambda}(t)^\dagger H_l(t). \tag{34}$$

As in Eq. (22), the equation of motion for $\widetilde{A}_0^t$ is

$$\begin{aligned}\partial_t \widetilde{A}_0^t &= \mathcal{U}_{\Lambda/\lambda}(t)^\dagger \mathcal{L}_\lambda(t) \mathcal{U}_{\Lambda/\lambda}(t) \widetilde{A}_0^t \\ &\equiv \widetilde{\mathcal{L}}_\lambda(t) \widetilde{A}_0^t,\end{aligned} \tag{35}$$

where $\widetilde{\mathcal{L}}_\lambda(t)\widetilde{A}_0^t = i[\widetilde{H}_\lambda(t), \widetilde{A}_0^t]$ and $\widetilde{H}_\lambda(t) = \sum_{i=1}^n \widetilde{H}_{l_i}(t)$.

The transformed operator $\widetilde{H}_{l_i}(t)$ is supported on (potentially) all sites between $l_{i-1}$ and $l_{i+1}$, yet it has the same norm as $H_{l_i}(t)$, namely, bounded by $J_{l_i}$. Thus, we can apply the same procedure as in Sec. IV A to the operator $\widetilde{A}_0^t$, with $\lambda$ in place of $\Lambda$ and with $\widetilde{H}_\lambda(t)$ as the Hamiltonian. Following an identical derivation to that of Eq. (29), we obtain

$$\left\|\mathcal{P}_{>l_i} \widetilde{A}_0^t\right\| \leq 4J_{l_i} \int_0^t ds \left\|\mathcal{P}_{>l_{i-1}} \widetilde{A}_0^s\right\|. \tag{36}$$

We use that $\|\mathcal{P}_r A_0^t\| \leq 2\|\mathcal{P}_{>l_n} A_0^t\| = 2\|\mathcal{P}_{>l_n} \widetilde{A}_0^t\|$ to start the iteration and $\|\widetilde{A}_0^s\| = \|A_0\|$ to terminate it. Thus,

$$\left\|\mathcal{P}_r A_0^t\right\| \leq 2\left\|A_0\right\| \left(\prod_{i=1}^{n} 4J_{l_i}\right) \frac{t^n}{n!}. \tag{37}$$

Since $\lambda$ is arbitrary, we are free to use the subset that gives the tightest bound:

$$\left\|\mathcal{P}_r A_0^t\right\| \leq 2\left\|A_0\right\| \min_\lambda \left[\left(\prod_{i=1}^{|\lambda|} 4J_{l_i}\right) \frac{t^{|\lambda|}}{|\lambda|!}\right], \tag{38}$$

where the minimum is over all subsets of links between 0 and $r$. Equation (38) is our improved LR bound.

Although the minimization in Eq. (38) may seem computationally expensive due to the $2^r$ possible $\lambda$, it can be performed efficiently. For a fixed size of $\lambda$, the optimal choice is clearly those links having the $|\lambda|$ smallest values of $J_l$. Thus one need only sort $\{J_l\}_{l=1}^r$ beforehand and the minimization amounts to simply checking the $r$ possible values of $|\lambda|$.

### C. An explicit bound in terms of the distribution of couplings

We now make use of Eq. (38) to prove general results in terms of the "distribution" of couplings $\mu_r(J)$, by which we mean the fraction of links between 0 and $r$ for which $J_l \leq J$. Reproducing Eq. (3) (recall that $\delta_{J_l \leq J} \equiv 1$ if $J_l \leq J$ and 0 otherwise),

$$\mu_r(J) \equiv \frac{1}{r} \sum_{l=1}^{r} \delta_{J_l \leq J}. \tag{39}$$

From this definition, for any $\epsilon > 0$ such that $\mu_r(\epsilon) \neq 0$, choosing $\lambda$ to be the subset of links with $J_l \leq \epsilon$ yields

$$\begin{aligned} \left\| \mathcal{P}_r A_0^t \right\| &\leq 2 \left\| A_0 \right\| \frac{(4\epsilon t)^{r\mu_r(\epsilon)}}{[r\mu_r(\epsilon)]!} \\ &\leq 2 \left\| A_0 \right\| \exp\left[ -r\mu_r(\epsilon) \left( 1 - \frac{4e\epsilon t}{r\mu_r(\epsilon)} \right) \right]. \end{aligned} \tag{40}$$

Now suppose that $\mu_r(J)$ converges as $r \to \infty$ to a function $\mu(J)$ defined on $[0, \infty)$, with the small-$J$ ("tail") behavior

$$\mu(J) \sim \mu_0 J^\alpha, \tag{41}$$

for some $\mu_0 > 0$ (see Refs. [68–70]). The parameter $\alpha > 0$ plays an essential role in what follows. We do need to specify the precise sense in which $\mu_r(J)$ converges. It is not sufficient to require merely that $\lim_{r\to\infty} \mu_r(J) = \mu(J)$ at any fixed $J$ but also that $\mu_r(J)$ behave as $\mu(J)$ on scales that decrease with increasing $r$. This is expressed by the following conditions, which we assume to be met:

(a) For any $\beta \in [0, 1/\alpha)$ and $J > 0$,

$$\lim_{r\to\infty} \frac{\mu_r(Jr^{-\beta})}{\mu(Jr^{-\beta})} = 1. \tag{42}$$

(b) For any $\beta > 1/\alpha$,

$$\lim_{r\to\infty} \frac{\min_{l=1}^r J_l}{r^{-\beta}} = \infty. \tag{43}$$

As an example of why Eqs. (42) and (43) are necessary, rather than simply the condition $\lim_{r\to\infty} \mu_r(J) = \mu(J)$ (which is contained in Eq. (42) as the case $\beta = 0$), suppose that there is a single link on which $J_l = 0$. Clearly, there cannot be any operator spreading past link $l$ [note that this is captured by the general bound in Eq. (38)]. On the other hand, the value of a single coupling does not affect the *fraction* that are less than any value in the large-$r$ limit. Thus, $\lim_{r\to\infty} \mu_r(J)$ at any fixed value of $J$ does not identify *individual* anomalously weak links, whereas Eq. (43) does. In other words, Eqs. (42) and (43) are a precise way of stating that the *weakest* links between sites 0 and $r$ are also distributed in a manner behaving as $\mu(J)$ at large $r$.

Strictly speaking, we only need (and prove) Eqs. (42) and (43) to hold for a dense subset of $\beta$—say, all rational $\beta$. The exponent $1/\alpha$ appears because $r\mu(Jr^{-1/\alpha}) = O(1)$—one expects to find couplings with $\beta < 1/\alpha$ but not with $\beta > 1/\alpha$ between sites 0 and $r$.

We feel that these conditions are reasonable to expect in practice. We show in Sec. V that if each $J_l$ is chosen independently from a literal probability distribution $\mu(J)$ obeying Eq. (41), then Eqs. (42) and (43) are satisfied with probability 1—in this sense, *any* sufficiently disordered set of couplings meets our requirements. Whether Eqs. (42) and (43) hold in any specific situation obviously depends on the system under consideration but in any case (and perhaps more importantly), one can always return to Eq. (38) if needed.

Let us first consider $\alpha \geq 1$. From Eq. (41) and the fact that $\lim_{r\to\infty} \mu_r(J) = \mu(J)$, there exists $R$ and $\epsilon > 0$ such that for all $r > R$, $\mu_r(\epsilon) > C\epsilon^\alpha$ (recall our convention that $C$ is any $r$- and $t$-independent constant for which the value may change from line to line). Thus, Eq. (40) becomes (for $r > R$)

$$\left\| \mathcal{P}_r A_0^t \right\| \leq 2 \left\| A_0 \right\| \exp\left[ -C\epsilon^\alpha r \left( 1 - \frac{4et}{C\epsilon^{\alpha-1} r} \right) \right]. \tag{44}$$

This is simply a conventional LR bound with ballistic front and exponential tail (a.a. because it holds for *all* $r > R$), having velocity $4e/C\epsilon^{\alpha-1}$.

Now consider $\alpha < 1$. Pick any $\beta \in (0, 1/\alpha)$. From Eqs. (41) and (42), there exists $R$ such that for all $r > R$, $\mu_r(r^{-\beta}) > Cr^{-\beta\alpha}$. Taking $\epsilon = r^{-\beta}$ in Eq. (40), we thus have that for all $r > R$,

$$\left\| \mathcal{P}_r A_0^t \right\| \leq 2 \left\| A_0 \right\| \exp\left[ -Cr^{1-\beta\alpha} \left( 1 - \frac{4et}{Cr^{1+\beta-\beta\alpha}} \right) \right]. \tag{45}$$

This is an a.a. LR bound having exponents

$$z = 1 + \beta - \beta\alpha, \tag{46}$$

$$\gamma = 1 - \beta\alpha. \tag{47}$$

Note that the front is sub-ballistic precisely for $\alpha < 1$ (whereas it is no tighter than Eq. (44) for $\alpha \geq 1$).

By taking $\beta \nearrow 1/\alpha$, we can make the dynamical exponent $z$ arbitrarily close to $1/\alpha$. This actually implies that for any $z < 1/\alpha$, we have an a.a. LR bound *with arbitrarily small generalized velocity*. Fixing $z$ and for any $v > 0$, setting $\beta \in (z, 1/\alpha)$ and taking $r$ sufficiently large (so that $4er^{-\beta} < Cvr^{1-\beta\alpha-z}$) gives

$$\begin{aligned}\left\|\mathcal{P}_r A_0^t\right\| &\leq 2\left\|A_0\right\| \exp\left[-Cr^{1-\beta\alpha}\left(1-\frac{4et}{Cr^{1+\beta-\beta\alpha}}\right)\right] \\ &\leq 2\left\|A_0\right\| \exp\left[-Cr^{1-\beta\alpha}\left(1-\frac{vt}{r^z}\right)\right]. \end{aligned} \tag{48}$$

Interestingly, the tightest tail corresponds to the opposite limit of $\beta$. Setting $\beta = 0$ gives the standard exponential tail (albeit only at distances $r > vt$ for some $v$), whereas increasing $\beta$ gives an increasingly stretched-exponential tail. The optimal choice of $\beta$ depends on the specific application: one should take $\beta \nearrow 1/\alpha$ if constraining the shape of the front is most important but one should set $\beta = 0$ if constraining the tail is most important.

We can combine these results (Eq. (44) for $\alpha \geq 1$ and Eq. (48) for $\alpha < 1$) simply by saying that Eq. (48) holds for any $z < z_c(\alpha) \equiv \max[1/\alpha, 1]$. This accounts for the blue region in Fig. 2, with the upper boundary being given by $z_c(\alpha)$.

Note that this analysis can straightforwardly be extended to limiting distributions $\mu(J)$ that are not simple power laws, with the expected results. First, if $\mu(J)$ decays to 0 at small $J$ faster than any power law (e.g., $\mu(J) \sim \exp[-1/J]$), then a conventional LR bound as in Eq. (44) still holds. If $\mu(J)$ decays slower than any power law (e.g., $\mu(J) \sim 1/\log J^{-1}$), and if Eq. (42) is obeyed for all $\beta > 0$, then $z = \infty$ in that an LR bound with infinitesimal $v$ holds for any finite $z$. Lastly, our main result still applies if $\mu(J)$ scales not solely as $J^\alpha$ but as $J^\alpha p(J)$ for some sub-power-law function $p(J)$—an LR bound with arbitrarily small generalized velocity holds for any $z < z_c(\alpha)$, as in Eq. (48).

### D. Tightness of the bound

As discussed in Sec. II, LR bounds should be complemented by an understanding of their tightness, ideally by constructing an explicit protocol $H \in \mathcal{H}_J$ that saturates the bound. To that end, we consider the simple transfer protocol shown in Fig. 3. Denoting the total run time of the circuit by $T_r$, clearly $\mathcal{P}_r A_0^{T_r} = A_0$ and so any valid LR bound must have a front that encompasses the space-time point $(r, T_r)$. We focus on the dynamical exponent—if $T_r = O(r^z)$, then no LR bound can have a dynamical exponent larger than $z$.

Effecting a SWAP gate for arbitrary $d$-dimensional local Hilbert spaces is not entirely trivial but a construction is given in Ref. [71] (see also Ref. [72]). For completeness, we give the relevant details in Appendix C. The only interactions needed (per SWAP gate) are a finite number of controlled-$Z$ (CZ) gates. In our case, since Eq. (5) must be respected, the time per CZ gate across link $l$ is $O(1/J_l)$. Thus the total run time is

$$T_r = C\sum_{l=1}^{r}\frac{1}{J_l}. \tag{49}$$

We again assume that the distribution of couplings $\mu_r(J)$ satisfies Eqs. (42) and (43). As we demonstrate below, it then follows that for any $z > \max[1/\alpha, 1]$,

$$\lim_{r\to\infty}\frac{T_r}{r^z} = 0. \tag{50}$$

Let us first note that it is the same threshold exponent $z_c(\alpha) \equiv \max[1/\alpha, 1]$ that enters into both Eq. (50) and (48). We can thus say that the dynamical exponent is $z_c(\alpha)$ in the following sense:

(a) For any $z < z_c(\alpha)$, there is no $H \in \mathcal{H}_J$ that can generate correlations at *any* sufficiently large distance $r$ in a time of order $r^z$.
(b) For any $z > z_c(\alpha)$, we know of an explicit protocol $H \in \mathcal{H}_J$ that can generate correlations at *every* sufficiently large distance $r$ in a time that vanishes compared to $r^z$.

However, the behavior precisely at $z = z_c(\alpha)$ is far more complicated and system dependent. In particular, the distinction between a.a. and i.o. bounds becomes essential, as we demonstrate explicitly in Sec. V.

We now turn to the proof of Eq. (50). It is convenient to define $Y_l \equiv 1/J_l$, so that $T_r = \sum_{l=1}^{r} Y_l$. Note that $\mu_r(J)$, the fraction of links with $J_l \leq J$, is equivalently the fraction with $Y_l \geq 1/J$. Fix $\gamma > 0$ and define $a_0 \equiv 0$, $a_k \equiv r^{(k-1)\gamma}$ for $k \geq 1$ (writing $a_k$ instead of $a_k(r)$ for conciseness). Also define $p_k$ to be the fraction of links with $Y_l \in [a_k, a_{k+1})$, equivalently

$$p_k \equiv \mu_r\left(a_k^{-1}\right) - \mu_r\left(a_{k+1}^{-1}\right). \tag{51}$$

By definition, we have the bound

$$T_r \leq r\sum_{k=0}^{\infty} p_k a_{k+1}. \tag{52}$$

For $p_0$ and $p_1$, we simply use that $p_0 \leq 1$, $p_1 \leq 1$. For $p_k$ with $k \geq 2$, first note that the second term in Eq. (51) can be neglected relative to the first at large $r$. Thus, it follows from Eqs. (41) and (42) that for any $\eta > 0$, there exists $R_k$ such that for all $r > R_k$,

$$p_k \in \left((1-\eta)Cr^{-(k-1)\alpha\gamma}, (1+\eta)Cr^{-(k-1)\alpha\gamma}\right). \tag{53}$$

Furthermore, Eq. (43) implies that if we take $K$ to be the smallest integer greater than $1/\alpha\gamma$, then there exists

$R_\infty$ such that for all $r > R_\infty$, $\min_{l=1}^{r} J_l > r^{-K\gamma} = a_{K+1}^{-1}$ and therefore $p_k = 0$ for all $k \geq K+1$. All together, we have that for $r > \max[R_2, \ldots, R_K, R_\infty]$,

$$p_k \leq \begin{cases} 1, & k \leq 1 \\ (1+\eta)Cr^{-(k-1)\alpha\gamma}, & 2 \leq k \leq K \\ 0, & K+1 \leq k \end{cases}. \tag{54}$$

Eq. (52) becomes

$$T_r \leq r + r^{1+\gamma} + (1+\eta)Cr^{1+\alpha\gamma} \sum_{k=2}^{K} r^{k(1-\alpha)\gamma}. \tag{55}$$

First, consider $\alpha > 1$. The sum in Eq. (55) is then $O(1)$ with respect to $r$. Since $\gamma$ is arbitrary, Eq. (50) follows for any $z > 1$ (namely, choose $\gamma < (z-1)/\alpha$). Note that this conclusion also applies when $\mu(J)$ decays faster than a power law—in such a case, $p_k$ (for $k \geq 2$) is even smaller than for any finite $\alpha$ and thus Eq. (55) remains a valid bound.

Next, suppose that $\alpha \leq 1$. The sum now grows no faster than $O(r^{K(1-\alpha)\gamma})$ and thus

$$T_r \leq Cr^{1+\alpha\gamma+K(1-\alpha)\gamma} \leq Cr^{\frac{1}{\alpha}+\gamma}. \tag{56}$$

The latter inequality follows because, by definition, $1/\alpha\gamma < K \leq 1 + 1/\alpha\gamma$. Again, since $\gamma$ is arbitrary, Eq. (50) follows.

Incidentally, this line of reasoning puts our discussion regarding the failure of the conventional LR bound [Eq. (32) in particular] on firmer ground. As noted above, the couplings enter into the conventional bound via the sum $\sum_{l=1}^{r} \log J_l$. We have that

$$r \sum_{k=0}^{\infty} p_k \log a_{k+1}^{-1} \leq \sum_{l=1}^{r} \log J_l \leq r \sum_{k=0}^{\infty} p_k \log a_k^{-1}. \tag{57}$$

The sums over $k$ again terminate at $K$ but now the summands go as $r^{-k\alpha\gamma} \log r^{-k\gamma}$ and are dominated by small $k$ regardless of $\alpha$. More precisely, use of Eq. (54) and the analogous lower bound on $p_k$ gives

$$Cr \log r^{-\gamma} \leq \sum_{l=1}^{r} \log J_l \leq C'r. \tag{58}$$

Since $\gamma$ can be arbitrarily small, inserting into Eq. (30) gives an LR bound for which the front, while not necessarily quite ballistic, cannot have a dynamical exponent larger than 1. As we have now established, that bound is far from tight.

## V. DISORDERED LIEB-ROBINSON BOUNDS

As a nontrivial example of a situation in which the above results apply, here we suppose that each $J_l$ is drawn independently and identically distributed (IID) from a probability distribution $\mu(J)$ for which the small-$J$ behavior is given by Eq. (41). We first prove Eqs. (42) and (43), not merely in some average sense but with probability 1, using standard techniques. The results in Secs. IV C and IV D then follow.

We next consider the threshold case $z = \max[1/\alpha, 1]$ in more detail. For $\alpha > 1$, it follows immediately from the strong law of large numbers (see Refs. [73,74] for an introduction) that the transfer protocol in Sec. IV D reaches all sites ballistically. For $\alpha < 1$, on the other hand, the distinction between a.a. and i.o. bounds becomes important—we derive an i.o. bound with arbitrarily small generalized velocity, implying that no protocol can reach every site in time $T_r = O(r^{1/\alpha})$, but we also show that the above transfer protocol does reach an infinite *subsequence* of sites in time $T_r = O(r^{1/\alpha})$ (again with probability 1). Interestingly, $\alpha = 1$ is the only point at which we cannot give a definite answer. Our i.o. bound still applies but the transfer protocol now fails to reach *any* site in time $T_r = O(r)$. It could be that an a.a. bound with arbitrarily small velocity holds for $\alpha = 1$ but we have not succeeded in proving so.

Finally, we discuss some straightforward extensions of the above results.

### A. Convergence of the distribution

First, we prove Eq. (42). This requires some tools from probability theory which can be found in textbooks on the subject [73,74] but are likely not common knowledge among physicists. Here, we apply these tools without further comment for ease of presentation but we include a description of them in Appendix D for completeness.

To prove Eq. (42), pick any $\epsilon > 0$ and $\beta \in [0, 1/\alpha)$. Define the event $E_r$ to be

$$E_r \equiv \left\{ \frac{\mu_r(Jr^{-\beta})}{\mu(Jr^{-\beta})} \notin (1-\epsilon, 1+\epsilon) \right\}. \tag{59}$$

Abbreviating $\mu(Jr^{-\beta})$ by $\mu$ for conciseness, Eq. (59) is equivalently the event that the number of couplings less than $Jr^{-\beta}$ is not between $(1-\epsilon)r\mu$ and $(1+\epsilon)r\mu$. We can evaluate the latter directly:

$$\begin{aligned}\Pr[E_r] &= \sum_{n\notin((1-\epsilon)r\mu,(1+\epsilon)r\mu)} \binom{r}{n}\mu^n(1-\mu)^{r-n} \\ &\le \sum_{n\notin((1-\epsilon)r\mu,(1+\epsilon)r\mu)} \exp\left[n\log\frac{r\mu}{n}+(r-n)\log\frac{r(1-\mu)}{r-n}\right] \\ &\le r\exp\left[-(1+\epsilon)r\mu\log(1+\epsilon)-\left(1-\frac{\mu}{1-\mu}\epsilon\right)(1-\mu)r\log\left(1-\frac{\mu}{1-\mu}\epsilon\right)\right], \end{aligned} \tag{60}$$

where the final inequality follows because the summand is maximized at $n=(1+\epsilon)r\mu$. One can confirm that the right-hand side goes as $r\exp[-Cr^{1-\beta\alpha}]$ at large $r$, with $C$ positive. Since $1-\beta\alpha>0$, $\Pr[E_r]$ is therefore summable and the probability of $E_r$ occurring infinitely often is zero (see Appendix D). The probability that this occurs for any rational $\epsilon$ or $\beta$, i.e., that $\mu_r(Jr^{-\beta})/\mu(Jr^{-\beta})$ does not converge to 1 for any $\beta<1/\alpha$, is likewise zero.

We now prove Eq. (43). Pick any $M>0$ and $\beta>1/\alpha$, and consider the event

$$E_r \equiv \left\{\min_{l=1}^{r} J_l \le Mr^{-\beta}\right\}. \tag{61}$$

We have (see Appendix D for the definition of "$E_r$ i.o.")

$$\Pr[E_r \text{ i.o.}] = \lim_{R\to\infty}\Pr\left[\exists r>R:\min_{l=1}^{r} J_l \le Mr^{-\beta}\right] = 1-\lim_{R\to\infty}\Pr\left[\forall r>R:\min_{l=1}^{r} J_l > Mr^{-\beta}\right]. \tag{62}$$

Clearly, if $J_l > Mr^{-\beta}$, then $J_l > Ms^{-\beta}$ for all $s>r$. Thus the following events are equivalent:

$$\begin{aligned}\left\{\forall r>R:\min_{l=1}^{r} J_l > Mr^{-\beta}\right\} = &\left\{J_1>M(R+1)^{-\beta}\right\}\bigcap\left\{J_2>M(R+1)^{-\beta}\right\}\bigcap\cdots\bigcap\left\{J_{R+1}>M(R+1)^{-\beta}\right\} \\ &\bigcap\left\{J_{R+2}>M(R+2)^{-\beta}\right\}\bigcap\left\{J_{R+3}>M(R+3)^{-\beta}\right\}\bigcap\cdots. \end{aligned} \tag{63}$$

Since the couplings $\{J_l\}$ are independent, the probability of the right-hand side is straightforward to evaluate:

$$\begin{aligned}\Pr\left[\forall r>R:\min_{l=1}^{r} J_l > Mr^{-\beta}\right] &= \left(1-\mu\left(M(R+1)^{-\beta}\right)\right)^{R+1}\lim_{R'\to\infty}\prod_{r=R+2}^{R'}\left(1-\mu\left(Mr^{-\beta}\right)\right) \\ &\sim \left(1-C(R+1)^{-\beta\alpha}\right)^{R+1}\exp\left[\sum_{r=R+2}^{\infty}\log\left(1-Cr^{-\beta\alpha}\right)\right], \end{aligned} \tag{64}$$

where the sum in the lower line is convergent because $\beta\alpha>1$. Therefore,

$$\lim_{R\to\infty}\Pr\left[\forall r>R:\min_{l=1}^{r} J_l > Mr^{-\beta}\right] = 1 \tag{65}$$

and $\Pr[\min_{l=1}^{r} J_l \le Mr^{-\beta} \text{ i.o.}]=0$. The probability for any $M\in\mathbb{N}$ or rational $\beta$, i.e., the probability that $\min_{l=1}^{r} J_l/r^{-\beta}$ does not diverge, is likewise zero. This completes the proof of Eq. (43).

### B. Fluctuations at the threshold exponent

We now take the dynamical exponent $z$ to be the threshold value $\max[1/\alpha,1]$, still for the model in which all couplings $\{J_l\}$ are chosen IID from $\mu(J)$. We determine whether an LR bound with arbitrarily small $v$ holds, both in the a.a. and i.o. sense. In the situations where we can provide a decisive answer (which is all $\alpha\neq1$), this completes the diagram in Fig. 2.

First, suppose that $\alpha>1$. Then, $\int d\mu(J)J^{-1}$ is finite (note that $d\mu(J)\sim CJ^{\alpha-1}dJ$ at small $J$) and recall that we identified a specific protocol for which the run time is

$T_r = C\sum_{l=1}^{r} 1/J_l$. Since the couplings are IID, the strong law of large numbers (SLLN) [73,74] gives that with probability 1,

$$\lim_{r\to\infty}\frac{T_r}{r} = \int d\mu(J)J^{-1} < \infty. \tag{66}$$

Thus the above transfer protocol reaches *every* sufficiently large-distance site with a nonzero velocity (namely, $[\int d\mu(J)J^{-1}]^{-1}$).

Now, suppose that $\alpha < 1$. Return to Eq. (38) and take the subset $\lambda$ to be solely the link connecting $r-1$ and $r$:

$$\left\|\mathcal{P}_r A_0^t\right\| \le 8\left\|A_0\right\| J_r t. \tag{67}$$

Although a rather loose bound, the probability that even the right-hand side of Eq. (67) exceeds $vtr^{-1/\alpha}$ for *all* large $r$ vanishes, for any $v > 0$. In other words, we have an i.o. bound with arbitrarily small generalized velocity: with probability 1, there exists a subsequence $\{r_i\}$ for which

$$\left\|\mathcal{P}_{r_i} A_0^t\right\| \le \left\|A_0\right\| \frac{vt}{r_i^{1/\alpha}}. \tag{68}$$

To prove Eq. (68), simply compute the probability that for all $r > R, J_r > vr^{-1/\alpha}$:

$$\begin{aligned}
&\Pr\left[\forall r > R : J_r > vr^{-1/\alpha}\right] \\
&\quad = \lim_{R'\to\infty}\prod_{r=R+1}^{R'}\left(1-\mu\left(vr^{-1/\alpha}\right)\right) \\
&\quad \sim \lim_{R'\to\infty}\prod_{r=R+1}^{R'}\left(1-Cr^{-1}\right) = 0.
\end{aligned} \tag{69}$$

This establishes an i.o. bound with arbitrarily small $v$ but let us consider again the previous transfer protocol. One can show that with probability 1,

$$\liminf_{r\to\infty}\frac{T_r}{r^{1/\alpha}} = 0, \tag{70}$$

meaning that this protocol does reach a *subsequence* of sites $\{r_i\}$ in time of order (and, in fact, asymptotically smaller than) $r_i^{1/\alpha}$. The proof of Eq. (70), which we adapt from Ref. [75], begins by noting that $T_r r^{-1/\alpha}$ converges in distribution to a non-negative random variable $S$ for which the support includes 0 (a fact that is well established but by no means trivial; see, e.g., Ref. [74]). Thus, pick any $\epsilon > 0$. We have that $\Pr[S \ge \epsilon] < 1$ and, at the same time, we can always choose a sequence of numbers $\{C_k\}$ such that $\Pr[S \ge C_k] < 1/k^2 \to 0$.

The convergence in distribution of $T_r r^{-1/\alpha}$ to $S$ implies that we can construct a subsequence $\{r_k\}$ with the following properties:

(a) $r_{k+1} - r_k \to \infty$ as $k \to \infty$.
(b) For sufficiently large $k$,

$$\left|\Pr\left[T_{r_k} r_k^{-1/\alpha} \ge C_k\right] - \Pr\left[S \ge C_k\right]\right| < \frac{1}{k^2}. \tag{71}$$

(c) $\epsilon_k > \epsilon$, where we define (for later convenience)

$$\epsilon_k \equiv \frac{2\epsilon r_{k+1}^{1/\alpha} - C_k r_k^{1/\alpha}}{(r_{k+1}-r_k)^{1/\alpha}}. \tag{72}$$

From Eq. (71), it follows that $\Pr[T_{r_k} r_k^{-1/\alpha} \ge C_k] < 2/k^2$ and thus, with probability 1, there is some $K$ such that $T_{r_k} r_k^{-1/\alpha} < C_k$ for all $k > K$.

Consider, for $k > K$, the event

$$E_k \equiv \left\{\forall k' \ge k : T_{r_{k'}} \ge 2\epsilon r_{k'}^{1/\alpha}\right\}. \tag{73}$$

Since $T_{r_{k'}} < C_{k'} r_{k'}^{1/\alpha}$ with probability 1, we have that

$$\begin{aligned}
&\Pr\left[\forall k' \ge k : T_{r_{k'}} \ge 2\epsilon r_{k'}^{1/\alpha}\right] \\
&= \Pr\left[\forall k' \ge k : T_{r_{k'}} \ge 2\epsilon r_{k'}^{1/\alpha} \bigcap T_{r_{k'}} < C_{k'} r_{k'}^{1/\alpha}\right] \\
&\le \Pr\left[\forall k' \ge k : T_{r_{k'+1}} - T_{r_{k'}} > 2\epsilon r_{k'+1}^{1/\alpha} - C_{k'} r_{k'}^{1/\alpha}\right] \\
&= \Pr\left[\forall k' \ge k : \frac{T_{r_{k'+1}} - T_{r_{k'}}}{(r_{k'+1}-r_{k'})^{1/\alpha}} > \epsilon_{k'}\right] \\
&\le \Pr\left[\forall k' \ge k : \frac{T_{r_{k'+1}} - T_{r_{k'}}}{(r_{k'+1}-r_{k'})^{1/\alpha}} > \epsilon\right],
\end{aligned} \tag{74}$$

using Eq. (72). Note that the *differences* $T_{r_{k'+1}} - T_{r_{k'}}$ are mutually independent and thus the probability on the right-hand side factors:

$$\Pr[E_k] \le \lim_{K'\to\infty}\prod_{k'=k}^{K'}\Pr\left[\frac{T_{r_{k'+1}} - T_{r_{k'}}}{(r_{k'+1}-r_{k'})^{1/\alpha}} > \epsilon\right]. \tag{75}$$

Furthermore, since $r_{k+1} - r_k \to \infty$, the random variable $(T_{r_{k+1}} - T_{r_k})/(r_{k+1}-r_k)^{1/\alpha}$ itself converges in distribution to $S$ and so, for sufficiently large $k$,

$$\Pr\left[\frac{T_{r_{k+1}} - T_{r_k}}{(r_{k+1}-r_k)^{1/\alpha}} > \epsilon\right] < \frac{1+\Pr[S \ge \epsilon]}{2}. \tag{76}$$

The right-hand side is strictly less than one, meaning the infinite product in Eq. (75) evaluates to zero. Thus, $\Pr[E_k] = 0$ and therefore

$$\Pr\left[T_{r_k} < 2\epsilon r_k^{1/\alpha} \text{ i.o.}\right] = 1 - \lim_{k\to\infty}\Pr[E_k] = 1. \tag{77}$$

In other words, Eq. (70) holds with probability 1.

It remains only to consider $\alpha = 1$. The calculation in Eq. (69) still holds and thus an i.o. bound with arbitrarily small velocity exists with probability 1. Yet for the transfer protocol, we now have

$$\liminf_{r\to\infty} \frac{T_r}{r} = \infty, \tag{78}$$

i.e., *no* site is reached ballistically. Compare with Eq. (70). It may be that a more sophisticated protocol is able to reach a subsequence ballistically or it may be that no such protocol exists. We are unable to rule out either possibility.

To prove Eq. (78), following Ref. [74], pick any $M > 0$ and define truncated random variables $Y_l' \equiv \min[Y_l, M]$ (recall that $Y_l \equiv 1/J_l$). The expectation value of $Y_l'$, denoted $\overline{Y}(M)$, is finite and therefore the SLLN applies. Thus, with probability 1,

$$\liminf_{r\to\infty} \frac{T_r}{r} = \liminf_{r\to\infty} \frac{1}{r}\sum_{l=1}^{r} Y_l \geq \liminf_{r\to\infty} \frac{1}{r}\sum_{l=1}^{r} Y_l' = \overline{Y}(M). \tag{79}$$

Since $\overline{Y}(M) \to \infty$ as $M \to \infty$, Eq. (78) follows.

### C. Extensions

Lastly, we discuss some straightforward extensions of the above results. These are not meant to be exhaustive; nor do we expect them to be particularly tight bounds—we only wish to point out some generalizations that can be obtained with little additional work.

#### *1. Couplings with finite-range correlations*

As a first example, suppose that the couplings $\{J_l\}$ are correlated but that correlations exist only within a finite range $\xi$. By the latter, we mean that joint distributions $\mu^{(n)}$ factor only if all couplings involved are separated by at least $\xi$ sites, e.g.,

$$\begin{aligned}&\mu^{(n+1)}\left(J_l, J_{l+\xi}, \ldots, J_{l+n\xi}\right)\\ &\quad = \mu^{(1)}(J_l)\mu^{(1)}(J_{l+\xi})\cdots\mu^{(1)}(J_{l+n\xi}).\end{aligned} \tag{80}$$

Such correlations present no difficulties—simply first pass to the interaction picture with respect to all but every $\xi$'th link and then the previous analysis applies. All lengths are reduced by a factor of $\xi$ and thus the generalized LR velocity is increased by a factor of $\xi^z$ (which we again do not claim to be a particularly accurate estimate) but the diagram in Fig. 2 remains unmodified.

#### *2. Bounds for multiple energy scales*

Suppose that $J_l$ can only take the values $J$ and $\epsilon J$, where $0 < \epsilon \ll 1$. The couplings are still chosen independently and the probability of $J_l = \epsilon J$ is $\epsilon^\alpha$. This is a discrete analogue to the situation from the previous subsections, for which $J_l$ could take any value greater than zero. Here, any LR bound clearly has a ballistic front but one can still ask how the LR velocity $v$ compares to the two scales $J$ and $\epsilon J$. By taking $\lambda$ in Eq. (38) to be those links with $J_l = \epsilon J$ (the fraction of which approaches $\epsilon^\alpha$ at large $r$ with probability 1) and comparing to the conventional LR bound, we find that

$$v = 4e\epsilon^{\max[1-\alpha,0]}J. \tag{81}$$

Analogous to the previous results, $v \ll J$ for $\alpha < 1$. Furthermore, the dependence of $v$ on $\epsilon$ in Eq. (81) is tight—the average of $1/J_l$ is finite for all $\alpha$, namely, given by $(1 - \epsilon^\alpha + \epsilon^{\alpha-1})/J$, and so the SLLN again applies, as in Eq. (66). Taking $\epsilon \ll 1$, the velocity of the transfer protocol is therefore Eq. (81) up to prefactors.

We can easily generalize to there being an arbitrary (finite) number of widely separated energy scales. Suppose that $J_l = \epsilon^{\gamma_k}J$ with probability $\epsilon^{\alpha_k}$ for $k \in \{1, \ldots, K\}$ and $J_l = J$ otherwise. Assume that $0 < \gamma_1 < \cdots < \gamma_K$ and $0 < \alpha_1 < \cdots < \alpha_K$. The optimal LR velocity is now

$$\begin{aligned}v &= 4e\epsilon^{\zeta}J,\\ \zeta &\equiv \max[\gamma_K - \alpha_K, \ldots, \gamma_1 - \alpha_1, 0].\end{aligned} \tag{82}$$

The dependence on $\epsilon$ is again tight. While simple, this result does highlight that in general, neither the largest nor the smallest energy scale necessarily determines the relevant velocity for operator growth on its own.

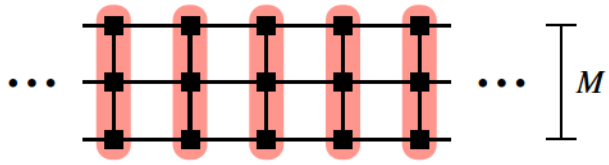


FIG. 7. An example of a "ladder," with $M = 3$ sites per rung. To apply our analysis, simply treat all $M$ sites connected vertically as a single "site," indicated by the red shading.

#### *3. Bounds for ladders*

Consider a system such as in Fig. 7, in which sites are labeled by $(i,j)$ with $i \in \{\cdots, -1, 0, 1, \ldots\}$ and $j \in \{1, \ldots, M\}$. The Hamiltonian is still given by a sum of terms for each link of the lattice. Interactions along vertical links are arbitrary and interactions along horizontal links (denoted $H_{ij}(t)$ for the link between $(i-1,j)$ and $(i,j)$) obey $\|H_{ij}(t)\| \leq J_{ij}$. Each $J_{ij}$ is again drawn independently from $\mu(J)$.

Since our analysis does not make any assumptions regarding the nature of the local Hilbert space,

we can simply identify each set of sites connected vertically—$\{(i,j)\}_{j=1}^{M}$ for fixed $i$—as comprising a single "local" Hilbert space. However, the interaction between neighboring $i$ is then $\sum_j H_{ij}(t)$, meaning that the coefficient $J_l$ entering into bounds such as Eq. (38) should be $\sum_j J_{lj}$. The probability of $\sum_j J_{lj} \leq J$ is bounded by

$$\mu(J/M)^M \leq \Pr\left[\sum_j J_{lj} \leq J\right] \leq \mu(J)^M. \tag{83}$$

Thus the correct exponent to use in our analysis is now $M\alpha$ and, in particular, $z_c(\alpha) = \max[1/M\alpha, 1]$. Obtaining an improvement over the conventional LR bound now requires $\alpha < 1/M$ but, nonetheless, we still have that $z_c \to \infty$ as $\alpha \to 0$.

In fact, we can adapt the construction of Ref. [76] to show that this result for $z_c(\alpha)$ is tight. We discuss this assuming two states per site ($d = 2$) labeled by $|0\rangle$ and $|1\rangle$—the same protocol extends to arbitrary $d$ simply by acting as the identity in the subspace orthogonal to $|0\rangle$ and $|1\rangle$, and it still produces an $O(1)$ commutator for generic operators $A_0$ and $B_r$ that act nontrivially on $|0\rangle$ and $|1\rangle$.

Taking cues from Ref. [76], consider starting in the product state with $a|0\rangle + b|1\rangle$ on site $(i-1,j)$ and $|0\rangle$ on all other sites of rungs $i-1$ and $i$. Since arbitrary vertical interactions are allowed, we can construct a unitary that takes this state to $(a|\overline{0}\rangle_{i-1} + b|\overline{1}\rangle_{i-1}) \otimes (|\overline{0}\rangle_i + |\overline{1}\rangle_i)/\sqrt{2}$ in arbitrarily short time, where $|\overline{0}\rangle_i$ and $|\overline{1}\rangle_i$ denote the states on rung $i$ with all sites in $|0\rangle$ and $|1\rangle$, respectively. Defining $H_i \equiv \sum_{j=1}^{M} J_{ij}|1\rangle\langle 1|_{(i-1,j)} \otimes |1\rangle\langle 1|_{(i,j)}$, we have

$$\begin{aligned} &e^{i\pi H_i/\sum_j J_{ij}}\left(a|\overline{0}\rangle_{i-1} + b|\overline{1}\rangle_{i-1}\right) \otimes \frac{|\overline{0}\rangle_i + |\overline{1}\rangle_i}{\sqrt{2}} \\ &= a|\overline{0}\rangle_{i-1} \otimes \frac{|\overline{0}\rangle_i + |\overline{1}\rangle_i}{\sqrt{2}} + b|\overline{1}\rangle_{i-1} \otimes \frac{|\overline{0}\rangle_i - |\overline{1}\rangle_i}{\sqrt{2}}. \end{aligned} \tag{84}$$

The states $(|\overline{0}\rangle_i + |\overline{1}\rangle_i)/\sqrt{2}$ and $(|\overline{0}\rangle_i - |\overline{1}\rangle_i)/\sqrt{2}$ can then be converted into $|\overline{0}\rangle_i$ and $|\overline{1}\rangle_i$, respectively, again in arbitrarily short time, using interactions solely on rung $i$. This procedure thus transforms the product state having $a|0\rangle + b|1\rangle$ on site $(i-1,j)$ and $|0\rangle$ otherwise into the generalized Greenberger-Horne-Zeilinger (GHZ) state $a|\overline{0}\rangle_{i-1} \otimes |\overline{0}\rangle_i + b|\overline{1}\rangle_{i-1} \otimes |\overline{1}\rangle_i$. Subsequently applying the procedure in reverse, albeit with the roles of rungs $i-1$ and $i$ exchanged, then takes this generalized GHZ state into the product state having $a|0\rangle + b|1\rangle$ on site $(i,j)$ and $|0\rangle$ otherwise. The net effect is that the state on site $(i-1,j)$ has been transferred to site $(i,j)$ in a time $2\pi/\sum_j J_{ij}$ [coming from Eq. (84)]. Repeating the transfer sequentially from rung 0 to $r$, we have a protocol analogous to Fig. 3 with run time $T_r = \sum_{l=1}^{r} 2\pi/\sum_j J_{lj}$.

Thus not only does our LR bound apply to the ladder of Fig. 7, with $\sum_j J_{lj}$ in place of $J_l$, but so does our analysis of the 1D transfer protocol, again using $\sum_j J_{lj}$ as an effective horizontal coupling. The result derived above that $z_c(\alpha) = \max[1/M\alpha, 1]$ is therefore tight. Note that if we restrict ourselves to bounded-strength vertical interactions but with bounds that are spatially uniform, then the dynamical exponent remains unaffected even once the time required to effect all single-rung transformations is incorporated. We leave the more complicated situation in which the vertical interactions themselves have weak links as a direction for future work.

The fact that $z_c(\alpha) \to 1$ as $M \to \infty$ for any $\alpha > 0$ suggests that our conclusions may not extrapolate to higher dimensions (and analogously to longer-range interactions). There are far more paths connecting any two sites in higher dimensions and it may be that transport remains ballistic for any power-law distribution of weak links. Of course, the analysis of the ladder presented here only accounts for weak links in one direction and so the behavior of truly multidimensional disordered systems remains an important open question.

## VI. APPLICATIONS

In Sec. IV, we derived a modified LR bound for non-translation-invariant systems—Eq. (45)—requiring only that the empirical distribution $\mu_r(J)$ converge to a function $\mu(J) \sim CJ^\alpha$ [as formalized by Eqs. (42) and (43)]. For $\alpha < 1$, the modified bound gives a significant improvement over the conventional bound and even guarantees that operator spreading is sub-ballistic. Here, we consider the consequences of this result for various applications (assume that $\alpha < 1$ throughout).

On the one hand, the manner in which LR bounds are used in the following is quite similar from case to case. Yet we see that different contexts come with different caveats, some more than others, and we discuss the many open directions for future work.

Note as well that we are working with the general bound of Sec. IV rather than the more detailed results of Sec. V. In particular, the implications of the "a.a."-"i.o." distinction for the following applications warrant further investigation.

### A. Growth of correlations

LR bounds directly place limitations on the extent to which correlations can develop following a quench. Consider the correlation function

$$G(t) \equiv \langle A^t B^t\rangle - \langle A^t\rangle\langle B^t\rangle, \tag{85}$$

where the expectation value is in a product state $|\Psi\rangle$, time evolution is under a Hamiltonian $H(t) \in \mathcal{H}_J$, and the operators $A^0$ and $B^0$ are supported on sites to the left of 0 and to the right of $r$, respectively (note that our analysis in Sec. IV applies equally well to such operators even if they are

not strictly local). Thus $G(0) = 0$ and one would like to understand how $G(t)$ grows in time.

The authors of Ref. [5] show that one can bound (assuming $\|A\| = \|B\| = 1$ for simplicity)

$$|G(t)| \leq 4\|\mathcal{P}_{\geq r/2}A^t\| + 4\|\mathcal{P}_{\leq r/2}B^t\|. \tag{86}$$

Supposing that $r$ is sufficiently large, we can immediately apply Eq. (45):

$$|G(t)| \leq \exp\left[-Cr^{1-\beta\alpha} + \frac{C't}{r^\beta}\right], \tag{87}$$

where $\beta$ can take any value in $(0, 1/\alpha)$. The fact that one can choose an optimal $\beta$ depending on $t$ makes Eq. (87) slightly more interesting than the conventional bound.

However, care must be taken in varying $\beta$. We have shown that for any $\beta \in (0, 1/\alpha)$, there exists $R(\beta)$ such that the bound holds for $r > R(\beta)$, but we have not shown that the convergence is uniform in $\beta$ (i.e., that $R(\beta)$ can be made independent of $\beta$). Rather than impose an additional requirement on the convergence of $\mu_r(J)$ and attempt to verify it in nontrivial situations, here we simply choose a finite set $\{\beta_i\}_{i=0}^n$ with $0 \equiv \beta_0 < \cdots < \beta_n < 1/\alpha$ (note that $\beta = 0$ simply recovers the conventional LR bound). The precise statement of Eq. (87) is that for $r/2 > \max[R(\beta_0), \ldots, R(\beta_n)]$,

$$|G(t)| \leq \exp\left[\min_{i=0}^{n}\left[-Cr^{1-\beta_i\alpha} + \frac{C't}{r^{\beta_i}}\right]\right]. \tag{88}$$

The behavior of Eq. (88) is shown in the dashed red line of Fig. 8. At large $r$, it is given by $\exp[-Cr + C't]$ until $t = O(r)$, then by $\exp[-Cr^{1-\beta_1\alpha} + C't/r^{\beta_1}]$ until $t = O(r^{1+\beta_1(1-\alpha)})$ and, in general,

$$|G(t)| \leq \exp\left[-Cr^{1-\beta_i\alpha} + \frac{C't}{r^{\beta_i}}\right], \quad r^{1+\beta_{i-1}(1-\alpha)} \ll t \ll r^{1+\beta_i(1-\alpha)}. \tag{89}$$

Of course, we are free to choose as large a set $\{\beta_i\}$ as we like (although this may increase the distance required for the bound to hold). One should heuristically think of Eq. (88) as minimizing over all $\beta \in (0, 1/\alpha)$ for any value of $r$, shown as the solid red line in Fig. 8.

### B. Creation of topological order

A related application is lower bounds on the time needed to create topological order. Again following Ref. [5], we say that two states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ in a 1D system of size $N$ are "topologically ordered" (relative to each other) if there exist constants $c_1 \in (0, 1)$ and $c_2 > 0$ such that, for every

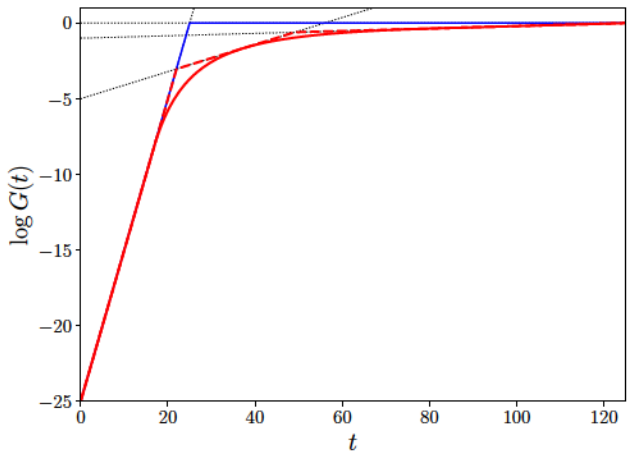


FIG. 8. The plot of Eq. (88), using $\alpha = 2/3$ for concreteness. The dashed red line corresponds to $n = 2$ with $\beta_i \in \{0, 3/4, 3/2\}$ (each of the individual curves being minimized among in Eq. (88) is shown as a dotted black line). The solid red line corresponds to $n = 30$ and $\beta_i$ equally spaced by $1/20$, which closely approximates the curve at large $n$. The blue line is the conventional LR bound for comparison. To reproduce the exact numbers in the plot, set $r = 25$ and $C = C' = 1$.

observable $O$ supported on a set with diameter $c_1N$ or less,

$$\begin{aligned}\left|\langle\Psi_1|O|\Psi_1\rangle - \langle\Psi_2|O|\Psi_2\rangle\right| &< 2e^{-c_2N},\\ \left|\langle\Psi_1|O|\Psi_2\rangle\right| &< e^{-c_2N}.\end{aligned} \tag{90}$$

In words, no "local" operator (even one supported on a nonvanishing fraction of the system) can distinguish between or couple such states.

Suppose that $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are topologically ordered and have been prepared from states $|\Phi_1\rangle$ and $|\Phi_2\rangle$ via time evolution under a Hamiltonian $H(t) \in \mathcal{H}_J$:

$$|\Psi_i\rangle = \mathcal{T}e^{-i\int_0^t ds H(s)}|\Phi_i\rangle. \tag{91}$$

Reference [5] shows that, for any operator $O$ with diameter less than $c_1N/2$,

$$\left|\langle\Phi_1|O|\Phi_1\rangle - \langle\Phi_2|O|\Phi_2\rangle\right| \leq 2e^{-c_2N} + 2\|\mathcal{P}_{\geq c_1N}O^t\| \tag{92}$$

and analogously for the off-diagonal matrix elements. We again use that

$$\|\mathcal{P}_{\geq c_1N}O^t\| \leq \exp\left[\min_{i=0}^{n}\left[-CN^{1-\beta_i\alpha} + \frac{C't}{N^{\beta_i}}\right]\right]. \tag{93}$$

Note that for $t \ll N$, Eq. (93) is exponentially small in $N$. Returning to Eq. (92), $|\Phi_1\rangle$ and $|\Phi_2\rangle$ then satisfy the definition of topological order. In other words, for times

less than $O(N)$, it is impossible to prepare topologically ordered states ($|\Psi_1\rangle$ and $|\Psi_2\rangle$) from any states that are not themselves topologically ordered ($|\Phi_1\rangle$ and $|\Phi_2\rangle$).

The above conclusion is identical to that of the conventional case but note that Eq. (93) in fact remains small for much longer times, until $t = O(N^{1/\alpha})$. The bound is no longer exponential (rather stretched-exponential) and so the definition in Eq. (90) is not strictly met but $|\Phi_1\rangle$ and $|\Phi_2\rangle$ exhibit a slightly looser sense of topological order nonetheless. In this sense, we have that for non-translation-invariant systems with $\alpha < 1$, it is impossible to prepare ordered states from unordered in times less than $O(N^{1/\alpha})$.

### C. Heating in periodically driven systems

The authors of Ref. [7] consider energy absorption in weakly driven systems, for which the Hamiltonian is of the form

$$H(t) = \sum_{l\in\Lambda} H_l - g\cos\omega t \sum_i O_i. \tag{94}$$

We limit ourselves to 1D, although Ref. [7] treats a general dimension. Within linear response theory [77], i.e., to leading order at small $g$, the energy absorption rate is proportional to $\sigma(\omega) \equiv \sum_{ij} \sigma_{ij}(\omega)$, where

$$\sigma_{ij}(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} dt e^{i\omega t}\left\langle\left[O_i^t, O_j^0\right]\right\rangle. \tag{95}$$

The expectation value denoted by $\langle\cdot\rangle$ is taken in the initial (potentially mixed) state of the system. Strictly speaking, $\sigma_{ij}(\omega)$ is a distribution and should be integrated against test functions to obtain meaningful results.

The authors first derive a bound

$$\left|\sigma_{ij}(\omega)\right| \le Ce^{-\kappa|\omega|}, \tag{96}$$

assuming a translation-invariant constraint $\|H_l\| \le J$. There is already the potential for tightening Eq. (96) in non-translation-invariant systems, since the constants $C$ and $\kappa$ involve sums over connected paths with factors of $J_l$, much as in Sec. IV A. However, we suspect that a more careful analysis of the path sum would not yield a significant improvement [see Eq. (32)], and so we continue to use Eq. (96).

LR bounds enter into the analysis of Ref. [7] as a means of bounding $|\sigma(\omega)|$ by $O(N)$ ($N$ being the number of sites in the system—note that naively bounding $|\sigma(\omega)| \le \sum_{ij} |\sigma_{ij}(\omega)|$ would give a bound $O(N^2)$). In particular, consider two sites $i$ and $j$ separated by a distance $r > r^*$, with $r^*$ to be chosen later. Assuming a conventional LR bound with velocity $v$, Eq. (95) can be bounded by two contributions:

$$\begin{aligned}\left|\sigma_{ij}(\omega)\right| &\le C\int_0^{r/v} dt e^{-\frac{t^2}{\delta t^2}} e^{-a(r-vt)} + C'\int_{r/v}^{\infty} dt e^{-\frac{t^2}{\delta t^2}} \\ &\le Ce^{-ar} + C'e^{-\frac{r^2}{v^2\delta t^2}}\end{aligned} \tag{97}$$

(for details, including the appearance of the Gaussian factor, see Ref. [7]). Since the terms of Eq. (97) decay exponentially with $r$ or faster, the sum over all $i$ and $j$ with $|i-j| > r^*$ is indeed $O(N)$ and scales as $\exp[-ar^*]$. For summing over $|i-j| < r^*$, simply use Eq. (96). Thus, by setting $r^* = O(|\omega|)$, both contributions decay exponentially with $|\omega|$, and therefore

$$\left|\sigma(\omega)\right| \le CNe^{-\kappa'|\omega|}. \tag{98}$$

The exponential decay with $|\omega|$ is the main result of Ref. [7].

Let us consider whether this conclusion is altered in non-translation-invariant systems by the use of our modified LR bound. As in the preceding subsections, we choose a set $\{\beta_i\}_{i=0}^n$ and minimize Eq. (45) over $\beta_i$. The integral over $t$ in Eq. (95) splits into multiple terms [compare with Eq. (97)]:

$$\begin{aligned}\left|\sigma_{ij}(\omega)\right| \le{}& C_0\int_0^{r/v_0} dt e^{-\frac{t^2}{\delta t^2}} e^{-a_0(r-v_0t)} \\ &+ C_1\int_{r/v_0}^{r^{1+\beta_1(1-\alpha)}/v_1} dt e^{-\frac{t^2}{\delta t^2}} e^{-a_1(r^{1-\beta_1\alpha}-v_1tr^{-\beta_1})} \\ &+ C_2\int_{r^{1+\beta_1(1-\alpha)}/v_1}^{r^{1+\beta_2(1-\alpha)}/v_2} dt e^{-\frac{t^2}{\delta t^2}} e^{-a_2(r^{1-\beta_2\alpha}-v_2tr^{-\beta_2})} \\ &+ \cdots + C_n\int_{r^{1+\beta_n(1-\alpha)}/v_n}^{\infty} dt e^{-\frac{t^2}{\delta t^2}}.\end{aligned} \tag{99}$$

Although the latter terms are indeed much smaller than in the translation-invariant case, note that the first term is unaffected. Thus $|\sigma_{ij}(\omega)|$ still scales as $\exp[-ar]$ at large $r$, we are still led to take $r^* = O(|\omega|)$, and the final result in Eq. (98) is unchanged.

The lack of any significant reduction in the heating rate can be traced back to the fact that it is the *tail* of the LR bound that constrains $|\sigma(\omega)|$ and the tail of the non-translation-invariant LR bound is no tighter than that of the translation-invariant case. Of course, this is only a statement about the bounds—the physics involved in any specific system very well may imply a dramatically slower heating rate.

Strictly speaking, the above comments only apply within linear response theory. To go beyond linear response, one could perform a Magnus-like expansion along the lines of Refs. [23,78,79] to derive an effective

Floquet Hamiltonian governing the dynamics. Since the terms of the Magnus expansion have norms given by products of couplings along paths [78] and we have established that disorder has a relatively modest effect on such expressions, the naive expectation is that the ultimate heating rate would not be significantly reduced according to Floquet-Magnus theory either. All the same, it would be quite interesting to investigate the effect of non-translation-invariant interactions on dynamics under the Floquet Hamiltonian (as distinct from the question of the time scale over which the Floquet Hamiltonian remains valid). Since the Floquet Hamiltonian involves higher-body interactions (i.e., terms supported on more than two sites), the results of this paper do not immediately apply and so we leave this for future work.

### D. Ground-state correlations

One of the most well-known applications of LR bounds is for the proof that gapped ground states have exponentially decaying correlations. Considering a 1D nearest-neighbor Hamiltonian $H$ for concreteness (although this result holds much more generally), the statement is that if there is a nonvanishing gap $\Delta E$ between the ground-state and first-excited-state energies, then for any local observables $A_0$ and $B_r$,

$$G(r) \equiv \langle A_0 B_r \rangle - \langle A_0 \rangle \langle B_r \rangle \leq C e^{-\kappa r}, \tag{100}$$

where the expectation values are in the ground state of $H$ (for full details of the proof, see Refs. [2,4]). Here, we only discuss the steps at which LR bounds enter.

The authors of Refs. [2,4] show that one can bound $G(r)$ by (assuming $\|A_0\| = \|B_r\| = 1$ for simplicity)

$$|G(r)| \leq \int_{-\infty}^{\infty} dt \frac{1}{|t|} e^{-\frac{t^2 \Delta E^2}{2q}} \left\| [A_0^t, B_r] \right\| + 2e^{-\frac{q}{2}}, \tag{101}$$

with $q$ to be chosen later. Much as we describe in the previous subsection, assuming a conventional LR bound with velocity $v$, split the integral into one over $|t| < r/2v$ and one over $|t| > r/2v$. Using the LR bound in the former and the trivial bound $\|[A_0^t, B_r]\| \leq 2$ in the latter, we find [80]

$$|G(r)| \leq C e^{-\frac{ar}{2}} + C' e^{-\frac{r^2 \Delta E^2}{8 v^2 q}} + 2e^{-\frac{q}{2}}. \tag{102}$$

Thus, if $\Delta E > 0$, taking $q = O(r)$ gives Eq. (100).

For non-translation-invariant systems, using our modified LR bound, we split the integral over $t$ into additional terms as in the previous subsection [see Eq. (99)]. Yet we again do not obtain any significant improvement over Eq. (102), since we still have a term scaling as $\exp[-ar/2]$.

In fact, it is unclear whether Eq. (102) itself applies to the systems considered here—Ref. [81] proves that ensembles of Hamiltonians are generically *gapless* if the norms of the interactions have continuous distributions extending to zero. However, the lack of a gap in this case is due to the existence with high probability of nearly disconnected local regions hosting low-energy excitations [81]. Since those excitations are decoupled from the larger system, one does not expect them to give rise to long-range correlations on physical grounds. To our knowledge, it remains a significant open question whether (and under what conditions) the ground states of such disordered systems have rapidly decaying correlations.

### E. Predicting properties of gapped ground states

An interesting recent application in which LR bounds enter is classical machine-learning algorithms for predicting properties of quantum many-body ground states. The authors of Ref. [82] consider a family of time-independent Hamiltonians $H(x)$ parametrized by a continuous (potentially multidimensional) variable $x$, with corresponding ground states $\rho(x)$. They obtain rigorous results on the ability to predict $\mathrm{Tr}\rho(x)O$ for a certain $x$ from knowledge of $\{\mathrm{Tr}\rho(x_i)O\}_{i=1}^n$ for other parameter values $\{x_i\}_{i=1}^n$, where $O$ is any observable that can be written as a sum of local terms.

A central ingredient is the result that if $H(x)$ has a nonzero spectral gap uniformly in $x$, then one can bound the size of the gradient $\nabla_x \mathrm{Tr}\rho(x)O$. Specializing to 1D chains with $H(x) = \sum_l H_l(x)$ and $O = \sum_i O_i$ (although Ref. [82] treats more general systems), the authors show that

$$\left\| \nabla_x \mathrm{Tr}\rho(x)O \right\| \leq \int_{-\infty}^{\infty} dt W(t) \sum_{il} \left\| \left[ O_i, \hat{u} \cdot \nabla_x H_l^t(x) \right] \right\|, \tag{103}$$

where $\hat{u}$ is an arbitrary unit vector in the parameter space and $W(t)$ is a filter function that decays faster than any polynomial as $|t| \to \infty$.

Much as before, the conventional LR bound enters by dividing the terms on the right-hand side into two groups, one in which $i$ and $l$ are separated by a distance less than $vt$ and the other in which they are separated by greater than $vt$. Use the trivial bound on the commutator for the former and the LR bound for the latter. Since there are $O(|t|)$ terms in the former and the sum over the latter is $O(1)$, Eq. (103) reduces to

$$\left\| \nabla_x \mathrm{Tr}\rho(x)O \right\| \leq C \int_{-\infty}^{\infty} dt W(t)|t|. \tag{104}$$

The remaining integral is finite (since $W(t)$ decays sufficiently fast) and thus the gradient is bounded. This property is then used in Ref. [82] to establish the efficiency of algorithms capable of predicting $\mathrm{Tr}\rho(x)O$.

In a certain sense, our modified LR bound for non-translation-invariant systems provides an improvement,

since it allows us to replace the factor of $|t|$ in Eq. (104) by $|t|^\alpha$. However, the important feature is merely that the resulting integral over $t$ is finite. Thus, while our modified bound does tighten the numerical value of the gradient, it does not seem to give any dramatic changes. Furthermore, since the analysis of Ref. [82] requires that the Hamiltonians $H(x)$ be gapped, the caveats from our discussion of ground-state correlations apply here as well. The more substantive question is therefore whether the conclusions of Ref. [82] apply to disordered systems in the first place.

## ACKNOWLEDGMENTS

It is a pleasure to acknowledge valuable and stimulating discussions with several colleagues: A. Chandran, P. J. D. Crowley, A. Deshpande, C. R. Laumann, D. M. Long, A. Polkovnikov, and M. C. Tran in particular. We would also like to thank MathOverflow users "Iosif Pinelis" and "mike" for guiding us through a proof of Eq. (70). We acknowledge funding by the U.S. Department of Energy (DoE) Advanced Scientific Computing Research (ASCR) Quantum Testbed Pathfinder program (Award No. DE-SC0019040), the Army Research Office (ARO) Multidisciplinary University Research Initiatives (MURI) program, the Air Force Office of Scientific Research (AFOSR), the DoE Quantum Systems Accelerator (QSA), the National Science Foundation (NSF) Quantum Leap Challenge Institutes (QLCI) (Award No. OMA-2120757), the DoE Advanced Scientific Computing Research (ASCR) Accelerated Research in Quantum Computing program (Award No. DE-SC0020312), the NSF Practical Fully-Connected Quantum Computer Challenge (PFCQC) program, the AFOSR MURI, the Defense Advanced Research Projects Agency (DARPA) SAVaNT ADVENT, and the Minta Martin and Simons Foundations.

## APPENDIX A: SOME PROPERTIES OF MATRIX NORMS

In this paper, we use $\|\cdot\|$ to denote the operator norm, i.e., the largest eigenvalue of the operator in question (technically the largest singular value but we only apply the norm to Hermitian operators). It is equivalently the $p \to \infty$ limit of the Schatten $p$-norm, defined for a generic $M \times M$ matrix $O$ as

$$\|O\|_p \equiv \left(\frac{1}{M}\mathrm{Tr}|O|^p\right)^{\frac{1}{p}} = \left(\frac{1}{M}\sum_{i=1}^{M}|\lambda_i|^p\right)^{\frac{1}{p}}. \tag{A1}$$

The Frobenius ($p = 2$) norm is often of independent interest (see, e.g., Ref. [64]). Let us first note that $\|O\|_p$ is a nondecreasing function of $p$ and thus the bounds on $\|O\|$ obtained in this work are automatically bounds on $\|O\|_p$ for all $p \in (0, \infty)$. This follows from the fact that for $p < q$, the function $f(x) = x^{q/p}$ is convex on $[0, \infty)$. Thus,

$$\left(\frac{1}{M}\sum_{i=1}^{M}|\lambda_i|^p\right)^{\frac{q}{p}} \le \frac{1}{M}\sum_{i=1}^{M}|\lambda_i|^q \tag{A2}$$

and exponentiating both sides by $1/q$ gives $\|O\|_p \le \|O\|_q$.

Now restrict ourselves to the operator norm and to tensor-product Hilbert spaces, i.e., $M = d^N$. A useful inequality is Eq. (12): for any Hermitian operator $O$ and any subset $\omega$ of the sites, $\|\mathcal{P}_\omega O\| \le 2\|O\|$, where $\mathcal{P}_\omega$ projects onto basis strings that act nontrivially somewhere in $\omega$.

To prove this, define $\mathcal{Q}_\omega \equiv I - \mathcal{P}_\omega$ to be the projector onto strings that *do* act trivially throughout $\omega$. We have that

$$\|\mathcal{P}_\omega O\| \le \|O\| + \|\mathcal{Q}_\omega O\|. \tag{A3}$$

Note that $\mathcal{Q}_\omega O$ is a tensor product between $\omega$ and its complement: $\mathcal{Q}_\omega O = I_\omega \otimes O_{\Omega/\omega}$ for a certain operator $O_{\Omega/\omega}$ acting on $\Omega/\omega$. Furthermore, $\|\mathcal{Q}_\omega O\| = \|O_{\Omega/\omega}\|$. Pick a state $|\psi\rangle$ such that $|\langle\psi|O_{\Omega/\omega}|\psi\rangle| = \|O_{\Omega/\omega}\|$ and define the normalized density matrix

$$\rho \equiv d^{-|\omega|} I_\omega \otimes |\psi\rangle\langle\psi|, \tag{A4}$$

so that $|\mathrm{Tr}\rho\mathcal{Q}_\omega O| = \|\mathcal{Q}_\omega O\|$. Also note that $\mathrm{Tr}\rho\mathcal{P}_\omega O = 0$ (since $\mathcal{P}_\omega O$ by definition consists only of basis strings orthogonal to the identity on $\omega$); thus $\|\mathcal{Q}_\omega O\| = |\mathrm{Tr}\rho O|$. Lastly, by the variational principle for density matrices, $|\mathrm{Tr}\rho O| \le \|O\|$. Returning to Eq. (A3), we have Eq. (12).

## APPENDIX B: PROOF THAT PROJECTION COMMUTES WITH DECOUPLED EVOLUTION

Recall the notation of Eq. (18): $\omega$ denotes any subset of sites and $\lambda$ denotes a subset of links such that $\omega$ and $\Omega/\omega$ are disconnected under $\Lambda/\lambda$. In words, removing $\lambda$ from the lattice disconnects $\omega$ from everything else. Equation (18), reproduced here,

$$\mathcal{P}_\omega \mathcal{U}_{\Lambda/\lambda}(t) = \mathcal{U}_{\Lambda/\lambda}(t)\mathcal{P}_\omega, \tag{B1}$$

is the statement that interactions on $\Lambda/\lambda$ alone cannot convert an operator supported outside $\omega$ into one for which the support intersects $\omega$ or vice versa. While obvious on physical grounds, it is worth confirming that this follows from the formal definitions in Sec. III B.

We need only make the following observations, all of which are clear from the explicit expression for $\mathcal{U}_{\Lambda/\lambda}(t)$ given by Eq. (17):

(a) $\mathcal{U}_{\Lambda/\lambda}(t)$ factors into $\mathcal{U}_\omega(t) \otimes \mathcal{U}_{\Omega/\omega}(t)$, where the first factor involves only the interactions within $\omega$ and the second only those outside $\omega$.

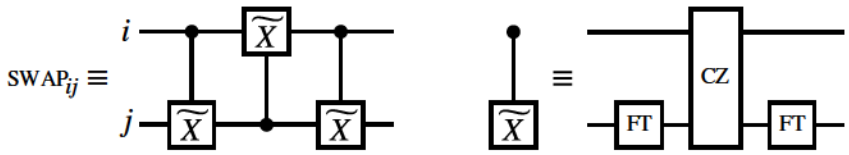


FIG. 9. The implementation of a SWAP gate between sites $i$ and $j$. The $\widetilde{X}_{ij}$ gate is given in Eq. (C1), the $\mathrm{FT}_j$ gate is given in Eq. (C3), and the $\mathrm{CZ}_{ij}$ gate is given in Eq. (C4).

(b) For any product operator $O \equiv O_\omega \otimes O_{\Omega/\omega}$,

$$\mathcal{U}_{\Lambda/\lambda}(t)O = \mathcal{U}_\omega(t)O_\omega \otimes \mathcal{U}_{\Omega/\omega}(t)O_{\Omega/\omega}. \tag{B2}$$

(c) The action of $\mathcal{U}(t)$ takes the identity string to itself and does not take any other basis string onto the identity (this statement is always true for any system).

Consider an arbitrary operator $O$, which can always be decomposed as

$$O = \sum_{\nu_\omega} X_\omega^{(\nu_\omega)} \otimes O_{\Omega/\omega}^{(\nu_\omega)}, \tag{B3}$$

where $\nu_\omega$ denotes a set of basis indices on $\omega$ and $O_{\Omega/\omega}^{(\nu_\omega)}$ denotes the partial projection of $O$ onto basis string $X_\omega^{(\nu_\omega)}$. As a result of the above, we have that

$$\begin{aligned}\mathcal{P}_\omega \mathcal{U}_{\Lambda/\lambda}(t)O &= \mathcal{P}_\omega \sum_{\nu_\omega} \mathcal{U}_\omega(t)X_\omega^{(\nu_\omega)} \otimes \mathcal{U}_{\Omega/\omega}(t)O_{\Omega/\omega}^{(\nu_\omega)} \\ &= \sum_{\nu_\omega \neq 0} \mathcal{U}_\omega(t)X_\omega^{(\nu_\omega)} \otimes \mathcal{U}_{\Omega/\omega}(t)O_{\Omega/\omega}^{(\nu_\omega)} \\ &= \mathcal{U}_{\Lambda/\lambda}(t) \sum_{\nu_\omega \neq 0} X_\omega^{(\nu_\omega)} \otimes O_{\Omega/\omega}^{(\nu_\omega)} \\ &= \mathcal{U}_{\Lambda/\lambda}(t)\mathcal{P}_\omega O. \end{aligned} \tag{B4}$$

## APPENDIX C: AN EXPLICIT STATE TRANSFER PROTOCOL

The transfer protocol $H(t) \in \mathcal{H}_J$ illustrated in Fig. 3, which we use to assess the tightness of our LR bounds in the main text, is given there in terms of SWAP gates. While perfectly sufficient on its own (see Ref. [72]), it is satisfying to see that $H(t)$ can be represented entirely in terms of "standard" interactions, even for arbitrary $d$-state degrees of freedom. This has been demonstrated in Ref. [71] and we summarize their construction here for the sake of completeness.

The circuit diagram for a single SWAP gate is shown in Fig. 9. The central element is the add-invert gate $\widetilde{X}_{ij}$ (the subscripts indicating the two sites involved), defined as

$$\widetilde{X}_{ij}\left|q_i, q_j\right\rangle \equiv \left|q_i, -q_j - q_i\right\rangle, \tag{C1}$$

where $q_i, q_j \in \{0, 1, \ldots, d-1\}$ label single-site basis states. All arithmetic here is to be interpreted mod $d$. Then, $\mathrm{SWAP}_{ij} = \widetilde{X}_{ij}\widetilde{X}_{ji}\widetilde{X}_{ij}$:

$$\left|q_i, q_j\right\rangle \xrightarrow{\widetilde{X}_{ij}} \left|q_i, -q_j - q_i\right\rangle \xrightarrow{\widetilde{X}_{ji}} \left|q_j, -q_j - q_i\right\rangle \xrightarrow{\widetilde{X}_{ij}} \left|q_j, q_i\right\rangle. \tag{C2}$$

The construction of the add-invert gate involves the single-site Fourier transform (FT) gate,

$$\mathrm{FT}_j\left|q_j\right\rangle \equiv \frac{1}{\sqrt{d}} \sum_{p_j=0}^{d-1} e^{2\pi i q_j p_j/d}\left|p_j\right\rangle, \tag{C3}$$

and the CZ gate,

$$\mathrm{CZ}_{ij}\left|q_i, q_j\right\rangle \equiv e^{2\pi i q_i q_j/d}\left|q_i, q_j\right\rangle. \tag{C4}$$

One can directly confirm that $\widetilde{X}_{ij} = \mathrm{FT}_j \mathrm{CZ}_{ij} \mathrm{FT}_j$ [71].

We need to express this procedure in terms of a Hamiltonian $H(t) \in \mathcal{H}_J$. The only interaction required is for implementation of the CZ gate. Defining the local operator $Z_j \equiv \mathrm{diag}(0, 1, \ldots, d-1)$, it is simply a $Z_i Z_j$ interaction. In order to respect the constraint on link $l$ that $\|H_l\| \leq J_l$ (note that $\|Z_j\| = d-1$), we set

$$H_l = -\frac{J_l}{(d-1)^2} Z_{l-1} Z_l. \tag{C5}$$

The CZ gate on link $l$ then amounts to applying interaction $H_l$ for time $2\pi(d-1)^2/dJ_l$ (and three such gates are needed per SWAP gate). As for the FT gates, since arbitrary local terms are allowed in the family $\mathcal{H}_J$ (see Ref. [56]), each amounts to applying a local field for some fixed time. The total run time required to effect state transfer from site 0 to $r$ is therefore (not including the $O(r)$ run time from all FT gates)

$$T_r = \frac{6\pi(d-1)^2}{d} \sum_{l=1}^{r} \frac{1}{J_l}. \tag{C6}$$

## APPENDIX D: SOME PROBABILISTIC TOOLS

Here, we summarize some of the mathematical tools from probability theory needed in Sec. V. Far more detail can be found in textbooks on the subject, such as Refs. [73,74] (including derivations of the following, which we do not present here).

To begin, keep in mind that the fundamental objects in probability theory are *subsets* of the set of all possible outcomes, called "events," together with a function Pr that maps such subsets to the interval $[0, 1]$. Oftentimes, not all subsets can be included in the domain of Pr, and Pr must obey certain natural properties, but we do not dwell on these here. The important thing to note is simply that we can perform all of the usual set-theoretic operations on events, such as the union or intersection, and many of the basic tools in probability theory involve relating the values of Pr with respect to those operations.

Our first tool is "countable subadditivity" (also known as the "union bound"): for any countable (potentially infinite) collection of events $\{E_r\}$,

$$\Pr\left[\bigcup_r E_r\right] \leq \sum_r \Pr[E_r]. \tag{D1}$$

We primarily use Eq. (D1) after establishing that $\Pr[E_r] = 0$ for all $r$—it then follows that $\Pr[\cup_r E_r] = 0$. In words, if each event $E_r$ separately has probability zero, then the probability that any of them occur (even if there are a countably infinite number) is also zero.

The next tool is the "continuity" of probabilities. Suppose that $\{A_r\}$ is an "increasing" set of events, in that $A_1 \subseteq A_2 \subseteq \cdots$, and similarly that $\{B_r\}$ is a "decreasing" set of events, in that $B_1 \supseteq B_2 \supseteq \cdots$. We then have that

$$\Pr\left[\bigcup_r A_r\right] = \lim_{r\to\infty} \Pr[A_r], \quad \Pr\left[\bigcap_r B_r\right] = \lim_{r\to\infty} \Pr[B_r]. \tag{D2}$$

For our purposes, we have events either of the form $A_r =$ "the event that $E_{r'}$ occurs for all $r' > r$," or $B_r =$ "the event that $E_{r'}$ occurs for some $r' > r$." In terms of set-theoretic operations, these are given by

$$A_r \equiv \bigcap_{r'>r} E_{r'}, \qquad B_r \equiv \bigcup_{r'>r} E_{r'}. \tag{D3}$$

Note that $\{A_r\}$ is an increasing set of events and $\{B_r\}$ is a decreasing set. We specifically want to determine the probability that some $A_r$ occurs and the probability that all $B_r$ occur. The former event is denoted "$E_{r'}$ a.a.," with a.a. as an abbreviation for "almost always," and the latter event is denoted "$E_{r'}$ i.o.," with i.o. as an abbreviation for "infinitely often." These events are given by

$$E_{r'} \text{ a.a.} \equiv \bigcup_r \bigcap_{r'>r} E_{r'}, \quad E_{r'} \text{ i.o.} \equiv \bigcap_r \bigcup_{r'>r} E_{r'}. \tag{D4}$$

Note that $E_{r'}$ a.a. and $E_{r'}$ i.o. can be described in words as "there exists an $r$ past which all $E_{r'}$ occur" and "past every $r$ there is some $E_{r'}$ that occurs," i.e., "$E_{r'}$ occurs almost always" and "$E_{r'}$ occurs infinitely often," respectively, in exactly the same manner as we have used in discussing LR bounds. Continuity—Eq. (D2)—allows us to express the probabilities of these events as

$$\begin{aligned}\Pr[E_{r'} \text{ a.a.}] &= \lim_{r\to\infty} \Pr\left[\bigcap_{r'>r} E_{r'}\right], \\ \Pr[E_{r'} \text{ i.o.}] &= \lim_{r\to\infty} \Pr\left[\bigcup_{r'>r} E_{r'}\right].\end{aligned} \tag{D5}$$

Lastly, we need the "first Borel-Cantelli lemma," which says that if a collection of events $\{E_r\}$ has probabilities such that $\sum_r \Pr[E_r]$ is finite, then the probability is 1 that only a finite number of the events occur:

$$\sum_r \Pr[E_r] < \infty \quad \longrightarrow \quad \Pr[E_r \text{ i.o.}] = 0. \tag{D6}$$

In particular, suppose that we have a sequence of random variables $\{X_r\}$ and we want to prove that it converges to zero with probability 1. The definition of $\{X_r\}$ *not* converging to zero is that there exist some $\epsilon > 0$ such that for every $R$, $|X_r| > \epsilon$ for some $r > R$. In other words, $|X_r| > \epsilon$ i.o. for some $\epsilon > 0$. It suffices to consider only rational $\epsilon > 0$ and thus, by countable subadditivity, we only need to prove that $\Pr[|X_r| > \epsilon \text{ i.o.}] = 0$ for each individual $\epsilon$. By the Borel-Cantelli lemma, it further suffices (but need not be necessary) to prove that $\sum_r \Pr[|X_r| > \epsilon]$ is finite. If each individual term $\Pr[|X_r| > \epsilon]$ can be calculated or at least bounded directly, this is a very useful line of attack.


[1] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, Commun. Math. Phys. **28**, 251 (1972).

[2] M. B. Hastings, Lieb-Schultz-Mattis in higher dimensions, Phys. Rev. B **69**, 104431 (2004).

[3] B. Nachtergaele and R. Sims, Lieb-Robinson bounds and the exponential clustering theorem, Commun. Math. Phys. **265**, 119 (2006).

[4] M. B. Hastings and T. Koma, Spectral gap and exponential decay of correlations, Commun. Math. Phys. **265**, 781 (2006).

[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] T. J. Osborne, Efficient Approximation of the Dynamics of One-Dimensional Quantum Spin Systems, Phys. Rev. Lett. **97**, 157202 (2006).

[7] D. A. Abanin, W. De Roeck, and F. Huveneers, Exponentially Slow Heating in Periodically Driven Many-Body Systems, Phys. Rev. Lett. **115**, 256803 (2015).

[8] N. Shiraishi and H. Tajima, Efficiency versus speed in quantum heat engines: Rigorous constraint from Lieb-Robinson bound, Phys. Rev. E **96**, 022138 (2017).

[9] Z. Eldredge, Z.-X. Gong, J. T. Young, A. H. Moosavian, M. Foss-Feig, and A. V. Gorshkov, Fast Quantum State Transfer and Entanglement Renormalization Using Long-Range Interactions, Phys. Rev. Lett. **119**, 170503 (2017).

[10] A. Deshpande, B. Fefferman, M. C. Tran, M. Foss-Feig, and A. V. Gorshkov, Dynamical Phase Transitions in Sampling Complexity, Phys. Rev. Lett. **121**, 030501 (2018).

[11] J. Haah, M. Hastings, R. Kothari, and G. H. Low, in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE Computer Society, Washington, DC, 2018), p. 350.

[12] G. Bentsen, Y. Gu, and A. Lucas, Fast scrambling on sparse graphs, PNAS **116**, 6689 (2019).

[13] B. Nachtergaele, R. Sims, and A. Young, Quasi-locality bounds for quantum lattice systems. I. Lieb-Robinson bounds, quasi-local maps, and spectral flow automorphisms, J. Math. Phys. **60**, 061101 (2019).

[14] N. Maskara, A. Deshpande, A. Ehrenberg, M. C. Tran, B. Fefferman, and A. V. Gorshkov, Complexity phase diagram for interacting and long-range bosonic Hamiltonians, arXiv:1906.04178 (2019).

[15] I. Prémont-Schwarz, A. Hamma, I. Klich, and F. Markopoulou-Kalamara, Lieb-Robinson bounds for commutator-bounded operators, Phys. Rev. A **81**, 040102 (2010).

[16] Z. Wang and K. R. A. Hazzard, Tightening the Lieb-Robinson Bound in Locally Interacting Systems, PRX Quantum **1**, 010303 (2020).

[17] C.-F. Chen and A. Lucas, Operator growth bounds from graph theory, Commun. Math. Phys. **385**, 1273 (2021).

[18] N. Schuch, S. K. Harrison, T. J. Osborne, and J. Eisert, Information propagation for interacting-particle systems, Phys. Rev. A **84**, 032309 (2011).

[19] T. Kuwahara and K. Saito, Lieb-Robinson Bound and Almost-Linear Light Cone in Interacting Boson Systems, Phys. Rev. Lett. **127**, 070403 (2021).

[20] C. Yin and A. Lucas, Finite Speed of Quantum Information in Models of Interacting Bosons at Finite Density, Phys. Rev. X **12**, 021039 (2022).

[21] T. Kuwahara, T. V. Vu, and K. Saito, Optimal light cone and digital quantum simulation of interacting bosons, arXiv:2206.14736 (2022).

[22] M. Foss-Feig, Z.-X. Gong, C. W. Clark, and A. V. Gorshkov, Nearly Linear Light Cones in Long-Range Interacting Quantum Systems, Phys. Rev. Lett. **114**, 157201 (2015).

[23] M. C. Tran, A. Ehrenberg, A. Y. Guo, P. Titum, D. A. Abanin, and A. V. Gorshkov, Locality and heating in periodically driven, power-law-interacting systems, Phys. Rev. A **100**, 052103 (2019).

[24] C.-F. Chen and A. Lucas, Finite Speed of Quantum Scrambling with Long Range Interactions, Phys. Rev. Lett. **123**, 250605 (2019).

[25] T. Kuwahara and K. Saito, Strictly Linear Light Cones in Long-Range Interacting Systems of Arbitrary Dimensions, Phys. Rev. X **10**, 031010 (2020).

[26] M. C. Tran, A. Y. Guo, C. L. Baldwin, A. Ehrenberg, A. V. Gorshkov, and A. Lucas, Lieb-Robinson Light Cone for Power-Law Interactions, Phys. Rev. Lett. **127**, 160401 (2021).

[27] D. Poulin, Lieb-Robinson Bound and Locality for General Markovian Quantum Dynamics, Phys. Rev. Lett. **104**, 190401 (2010).

[28] B. Descamps, Asymptotically decreasing Lieb-Robinson velocity for a class of dissipative quantum dynamics, J. Math. Phys. **54**, 092202 (2013).

[29] R. Sweke, J. Eisert, and M. Kastner, Lieb-Robinson bounds for open quantum systems with long-ranged interactions, J. Phys. A **52**, 424003 (2019).

[30] A. Y. Guo, S. Lieu, M. C. Tran, and A. V. Gorshkov, Clustering of steady-state correlations in open systems with long-range interactions, arXiv:2110.15368 (2021).

[31] C. K. Burrell, J. Eisert, and T. J. Osborne, Information propagation through quantum chains with fluctuating disorder, Phys. Rev. A **80**, 052319 (2009).

[32] M. Gebert, A. Moon, and B. Nachtergaele, A Lieb-Robinson bound for quantum spin chains with strong on-site impurities, Rev. Math. Phys. **34**, 2250007 (2022).

[33] C.-F. Chen, Concentration of OTOC and Lieb-Robinson velocity in random Hamiltonians, arXiv:2103.09186 (2021).

[34] Reference [33] does in fact allow for non-translation-invariant interactions but in the same manner as do the conventional bounds. It would be worthwhile to investigate whether the results of Ref. [33] can be tightened much as we do here.

[35] C. K. Burrell and T. J. Osborne, Bounds on the Speed of Information Propagation in Disordered Quantum Spin Chains, Phys. Rev. Lett. **99**, 167201 (2007).

[36] E. Hamza, R. Sims, and G. Stolz, Dynamical localization in disordered quantum spin systems, Commun. Math. Phys. **315**, 215 (2012).

[37] M. Gebert and M. Lemm, On polynomial Lieb-Robinson bounds for the XY chain in a decaying random field, J. Stat. Phys. **164**, 667 (2016).

[38] A. Elgart, A. Klein, and G. Stolz, Manifestations of dynamical localization in the disordered XXZ spin chain, Commun. Math. Phys. **361**, 1083 (2018).

[39] K. H. Fischer and J. A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, 1991).

[40] M. Mezard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).

[41] P. A. Lee and T. V. Ramakrishnan, Disordered electronic systems, Rev. Mod. Phys. **57**, 287 (1985).

[42] B. Kramer and A. MacKinnon, Localization: Theory and experiment, Rep. Prog. Phys. **56**, 1469 (1993).

[43] T. Vojta, Quantum Griffiths effects and smeared phase transitions in metals: Theory and experiment, J. Low. Temp. Phys. **161**, 299 (2010).

[44] X.-Y. Song, C.-M. Jian, and L. Balents, Strongly Correlated Metal Built from Sachdev-Ye-Kitaev Models, Phys. Rev. Lett. **119**, 216601 (2017).

[45] A. Altland, D. Bagrets, and A. Kamenev, Quantum Criticality of Granular Sachdev-Ye-Kitaev Matter, Phys. Rev. Lett. **123**, 106601 (2019).

[46] A. Altland, D. Bagrets, and A. Kamenev, Sachdev-Ye-Kitaev Non-Fermi-Liquid Correlations in Nanoscopic Quantum Transport, Phys. Rev. Lett. **123**, 226801 (2019).

[47] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states, Ann. Phys. **321**, 1126 (2006).

[48] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phenomenology of fully many-body-localized systems, Phys. Rev. B **90**, 174202 (2014).

[49] R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).

[50] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, Rev. Mod. Phys. **91**, 021001 (2019).

[51] J. Z. Imbrie, On many-body localization for quantum spin chains, J. Stat. Phys. **163**, 998 (2016).

[52] J. Suntajs, J. Bonca, T. Prosen, and L. Vidmar, Quantum chaos challenges many-body localization, Phys. Rev. E **102**, 062144 (2020).

[53] D. A. Abanin, J. H. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S. A. Parameswaran, F. Pollmann, A. C. Potter, M. Serbyn, and R. Vasseur, Distinguishing localization from chaos: Challenges in finite-size systems, Ann. Phys. **427**, 168415 (2021).

[54] D. Sels and A. Polkovnikov, Dynamical obstruction to localization in a disordered spin chain, Phys. Rev. E **104**, 054105 (2021).

[55] P. J. D. Crowley and A. Chandran, A constructive theory of the numerically accessible many-body localized to thermal crossover, arXiv:2012.14393 (2020).

[56] Our results actually apply to any Hamiltonian of the form $H(t) = \sum_{l\in\Lambda} H_l(t) + \sum_{i\in\Omega} h_i(t)$, for *arbitrary* local terms $h_i(t)$. The local terms can trivially be accounted for by first passing to the interaction picture with respect to them—the terms $H_l(t)$ in this new frame have both the same support and the same norm as originally; hence our analysis applies equally well using them. That said, the behavior of any *specific* system clearly can depend strongly on the local terms.

[57] C. Chamon, M. P. Kennett, H. E. Castillo, and L. F. Cugliandolo, Separation of Time Scales and Reparametrization Invariance for Aging Systems, Phys. Rev. Lett. **89**, 217201 (2002).

[58] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective, Phys. Rep. **523**, 127 (2013).

[59] C. L. Baldwin and C. R. Laumann, Quantum algorithm for energy matching in hard optimization problems, Phys. Rev. B **97**, 224201 (2018).

[60] We are certainly not the first to improve on Eq. (1)—see, e.g., Refs. [16,17] for modifications that yield smaller LR velocities, sometimes even parametrically so. We do not attempt to marry our approach with theirs but we expect that it can be done and leave this for future work.

[61] An alternative quantity is $v_{\mathrm{LR}}(r) \equiv (\mathrm{d}t(r)/\mathrm{d}r)^{-1}$ [22]. Although distance dependent, $v_{\mathrm{LR}}(r)$ has the correct units of velocity even when $z \neq 1$.

[62] A. Nahum, J. Ruhman, and D. A. Huse, Dynamics of entanglement and transport in one-dimensional systems with quenched randomness, Phys. Rev. B **98**, 035118 (2018).

[63] Strictly speaking, Eq. (6) is *not* the Heisenberg equation of motion for an arbitrary time-dependent Hamiltonian $H(t)$. The Heisenberg-picture operator $A_0^t$ is defined by the relation $\langle\Psi|A_0^t|\Psi\rangle \equiv \langle\Psi^t|A_0|\Psi^t\rangle$, where $|\Psi^t\rangle$ is the solution to the Schrodinger equation: $i\partial_t|\Psi^t\rangle = H(t)|\Psi^t\rangle$ with $|\Psi^0\rangle = |\Psi\rangle$. The solution can be written $|\Psi^t\rangle = \mathcal{T}e^{-i\int \mathrm{d}s H(s)}|\Psi\rangle$, where the operator on the right-hand side is a time-ordered exponential—earlier times appear to the right. Thus, in the expression $\langle\Psi^t|A_0|\Psi^t\rangle$, which defines $A_0^t$, earlier times appear *outside* later times. The differential equation that correctly has $A_0^t$ as its solution is $\partial_s A_0^s = i[H(t-s), A_0^s]$ with $A_0^0 = A_0$—the commutator must be applied with later times first and earlier times last. Despite all of this, however, our analysis applies equally well with $H(t-s)$ in place of $H(s)$. Since the latter is less burdensome notationwise, we use Eq. (6) without further comment.

[64] M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, and A. Lucas, Hierarchy of Linear Light Cones with Long-Range Interactions, Phys. Rev. X **10**, 031009 (2020).

[65] Strictly speaking, past works have derived bounds that hold for every single site (i.e., $R = 0$). Here, we do not differentiate between "almost-always" and such "always" bounds (although one arguably could), since our focus is on asymptotic behavior. Given an a.a. bound with $R > 0$, one simple way to extend it to all sites is to use the trivial bound $\|[A_0^t, B_r]\| \leq 2\|A_0\|\|B_r\|$ for $r \leq R$. If the couplings $\{J_l\}$ are bounded by some $J_{\max}$, another way is to use the conventional LR bound corresponding to $J_{\max}$ for $r \leq R$ [as in Eq. (31)]. Since the number of sites that violate the a.a. bound is by definition finite, neither prescription changes any of the asymptotic behavior.

[66] It is interesting to note that this same technique has been used to derive optimal LR bounds for power-law-interacting systems, e.g., $H(t) = \sum_{ij} H_{ij}(t)$ with $\|H_{ij}(t)\| \leq |i-j|^{-\alpha}$. To improve on the conventional bound, which would predict a front growing exponentially in time, group the interactions into short-range ($|i-j| \leq r^*$ for some $r^*$) and long-range ($|i-j| > r^*$) and then pass to the interaction picture with respect to short-range terms [22]. It has recently been shown that this procedure can be repeated for a hierarchy of length scales to obtain LR bounds that are provably tight [26].

[67] Note that this technique only works for *real*-time evolution—even though one can perfectly well define interaction-picture transformations in imaginary time—since it is essential that the transformation preserve the norms of the remaining terms (the same statement in fact applies when deriving the conventional LR bound in the first place).

[68] Recall the precise definition of the limit of a sequence: the statement $\lim_{r\to\infty} x_r = x$ is the statement that for any $\eta > 0$, there exists $R(\eta) \in \mathbb{N}$ such that $|x_r - x| < \eta$ for all $r > R(\eta)$.

[69] Throughout this paper, a statement such as "$f(x) \sim g(x)$ as $x \to 0$" means that $\lim_{x\to 0} f(x)/g(x) = 1$, i.e., for any $\eta > 0$, there exists $\delta > 0$ such that $(1-\eta)g(x) < f(x) < (1+\eta)g(x)$ for all $|x| < \delta$.

[70] One might be concerned by us taking $r \to \infty$, since we are considering a finite chain of $N$ sites. Yet note that all of the bounds derived in Secs. IV and V are independent of $N$. Thus, to be precise, we are supposing that we have an *infinite* sequence of couplings $\{J_l\}$ that obeys Eqs. (42) and (43) and we are considering the restriction of an infinite lattice to $N$ sites. Any statement involving $r$ should be interpreted as applying when $N > r$.

[71] J. C. Garcia-Escartin and P. Chamorro-Posada, A SWAP gate for qudits, Quantum Inf. Process. **12**, 3625 (2013).

[72] As pointed out to us by P. J. D. Crowley and C. R. Laumann, it is not necessary to give an explicit construction of the SWAP gate. From the definition of SWAP, the quantity $H_l \equiv i\pi^{-1} J_l \log \mathrm{SWAP}_l$ is a Hermitian operator supported on link $l$ with norm $J_l$. The Hamiltonian that applies $H_l$ sequentially as in Fig. 3 for time $\pi J_l^{-1}$, although not given in a completely constructive form, is a legitimate member of $\mathcal{H}_J$.

[73] J. S. Rosenthal, *A First Look at Rigorous Probability Theory* (World Scientific, Singapore, 2006).

[74] R. Durrett, *Probability: Theory and Examples* (Cambridge University Press, Cambridge, 2019).

[75] https://mathoverflow.net/questions/418111/what-is-the-liminf-of-a-sum-of-i-i-d-random-variables-with-heavy-tails.

[76] M. C. Tran, A. Y. Guo, A. Deshpande, A. Lucas, and A. V. Gorshkov, Optimal State Transfer and Entanglement Generation in Power-Law Interacting Systems, Phys. Rev. X **11**, 031016 (2021).

[77] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995).

[78] T. Kuwahara, T. Mori, and K. Saito, Floquet-Magnus theory and generic transient dynamics in periodically driven many-body quantum systems, Ann. Phys. **367**, 96 2016).

[79] Dmitry A. Abanin, Wojciech De Roeck, Wen Wei Ho, and Francois Huveneers, Effective Hamiltonians, prethermalization, and slow energy absorption in periodically driven many-body systems, Phys. Rev. B **95**, 014112 (2017).

[80] The factor of $1/|t|$ in the integrand is annoying but inconsequential. First, the integrand is nonsingular as $|t| \to 0$ because $\|[A_0^t, B_r]\| \le (Ct)^r/r!$. One sees this in Eq. (30) but note that it does *not* hold in the weaker Eq. (31), which is usually taken as the LR bound. Thus divide the integral over $t$ into $|t| < 1$ and $|t| \ge 1$. For $|t| < 1$, bound the integral by $C^r/r!$, which decays faster than any exponential regardless of $C$. For $|t| \ge 1$, upper bound $1/|t|$ by 1 and then proceed as in the main text.

[81] R. Movassagh, Generic Local Hamiltonians Are Gapless, Phys. Rev. Lett. **119**, 220504 (2017).

[82] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, Provably efficient machine learning for quantum many-body problems, arXiv:2106.12627 (2021).