


Optimal Frobenius light cone in spin chains with power-law interactions

Chi-Fang Chen^{1,*} and Andrew Lucas^{2,†}

¹*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

²*Department of Physics and Center for Theory of Quantum Matter, University of Colorado, Boulder, Colorado 80309, USA*

 (Received 28 May 2021; revised 17 October 2021; accepted 26 November 2021; published 13 December 2021)

In many-body quantum systems with spatially local interactions, quantum information propagates with a finite velocity, reminiscent of the “light cone” of relativity. In systems with long-range interactions which decay with distance r as $1/r^\alpha$, however, there are multiple light cones which control different information theoretic tasks. We show an optimal (up to logarithms) “Frobenius light cone” obeying $t \sim r^{\min(\alpha-1,1)}$ for $\alpha > 1$ in one-dimensional power-law interacting systems with finite local dimension: this controls, among other physical properties, the butterfly velocity characterizing many-body chaos and operator growth. We construct an explicit random Hamiltonian protocol that saturates the bound and settles the optimal Frobenius light cone in one dimension. We partially extend our constraints on the Frobenius light cone to a several operator p -norms, and show that Lieb-Robinson bounds can be saturated in at most an exponentially small $e^{-\Omega(r)}$ fraction of the many-body Hilbert space.

DOI: [10.1103/PhysRevA.104.062420](https://doi.org/10.1103/PhysRevA.104.062420)

I. INTRODUCTION

The celebrated Lieb-Robinson theorem proves that for arbitrary systems with nearest-neighbor interactions (with bounded strength), the speed at which quantum information can propagate is finite [1–6]. Intuitively, the time t it takes to prepare an entangled state between two qubits (in an initially unentangled state) separated by distance r is lower bounded as

$$t \geq \frac{r}{v_{\text{LR}}}, \quad (1.1)$$

with v_{LR} the so-called Lieb-Robinson velocity. In many respects, v_{LR} is analogous to the speed of light c in special relativity: no signals can be sent faster than v_{LR} . For this reason, we call (1.1) the *Lieb-Robinson light cone*. Over the past few decades, numerous unexpected and important results have been shown to follow from the Lieb-Robinson theorem, including the exponential decay of correlation functions in a gapped ground state, together with an area law for entanglement [3] and the effectiveness of matrix product state representations [7], proofs of the ability to efficiently simulate local quantum systems [8], and demonstration of the stability of topological order [9].

A. Alternative choices of norms

In quantum information dynamics, we study information propagation through the commutator between local operators (e.g., Pauli matrices) localized on lattice sites x and y . This commutator takes the form of $[A_x(t), B_y]$, where

$$A_x(t) := e^{iHt} A_x e^{-iHt} \quad (1.2)$$

denotes Heisenberg time evolution.¹ Indeed, local operators commute $[A_x(0), B_y] = 0$, and a nonvanishing commutator indicates the operator $A_x(t)$ has “grown” to a distant site y .

The traditional Lieb-Robinson bounds quantify the growth of this commutator in the infinity norm (or the operator norm): $\|[A_x(t), B_y]\|_\infty$, being the maximal singular value of the operator A . In slightly more physical terms,

$$\|A\| \equiv \|A\|_\infty = \sup_{\psi, \psi'} |\langle \psi' | A | \psi \rangle| \quad (1.3)$$

is the maximal matrix element of A between any normalized many-body states ψ and ψ' and we drop the subscript. A bound on this operator norm then rules out the possibility of signaling, under the “worst-case” choices of states ψ and ψ' . For quantum technology applications, such a bound is relevant to a device only when the *many-body* initial and final states can be controlled exactly. If the device is imperfect and prone to some local errors, then we would like to know if other states in Hilbert space can also transmit information similarly fast.

Recently, [10] emphasized that there can be *multiple different light cones* for quantum information. Mathematically, these different light cones correspond to the choices of norms, which capture different physics and constrain different quantum information theoretic tasks; the physical question would be as follows: “What quantum mechanical task are we trying to bound?”

As a concrete example, we consider whether or not one can signal between two qubits x and y *independently* of the actual quantum state. This amounts to averaging over the magnitude of $\langle \psi' | A | \psi \rangle$ over all pairs of ψ and ψ' [10]. This can be

¹For simplicity, we assume in this equation that Hamiltonian H does not depend on time. In all main results in this paper, we will eventually relax this assumption.

*chifang@caltech.edu

†andrew.j.lucas@colorado.edu

captured by the *square* of matrix element's magnitude, and (for an operator A) we then recognize that this average is nothing but the normalized 2-norm, which will be called the *Frobenius norm*

$$\|A\|_F^2 := \frac{\text{Tr}(A^\dagger A)}{\text{Tr}(I)}. \quad (1.4)$$

Here I denotes the identity matrix. If we wish to design a quantum device in which signaling or quantum state transfer is achieved regardless of the details of the many-body state, we will care not about the Lieb-Robinson bound on a commutator, but instead a ‘‘Frobenius bound’’ on $\|[A_x(t), B_y]\|_F$.

Interestingly, this Frobenius norm has a natural interpretation as an (out-of-time-ordered) correlator at infinite temperature

$$\begin{aligned} \|[A_x(t), B_y]\|_F^2 &= \frac{\text{Tr}([A_x(t), B_y]^\dagger [A_x(t), B_y])}{\text{Tr}(I)} \\ &= \text{Tr}(\rho_{T=\infty} [A_x(t), B_y]^\dagger [A_x(t), B_y]). \end{aligned} \quad (1.5)$$

Here $\rho_{T=\infty} \propto I$ denotes the infinite-temperature density matrix. Such objects have been used to diagnose many-body chaos in extensive studies over the past few years. Indeed, when the Frobenius norm becomes large for any operator B_y , information about the initial operator A_x cannot be retrieved via local measurements.

A simple way to define a ‘‘light cone’’ out of either of these norms is to define $t(r)$ as the first time that qubits separated by distance r have a large commutator, e.g.,

$$t(r) := \text{arginf}_t \left\{ 0 < t < \infty : \|[A_0(t), B_r]\|_* \geq \frac{1}{2} \right\}. \quad (1.6)$$

The notation here is simply that we look for the smallest time (arginf) at which a certain commutator norm is bigger than some $O(1)$ number, such as $\frac{1}{2}$. It is interesting to ask whether in *all* Hamiltonian dynamics consistent with various simple constraints (e.g., nearest-neighbor interactions) there are bounds on $t(r)$. For the operator norm, the answer is provided by the Lieb-Robinson theorem (1.1). But, is it possible that in typical states (Frobenius norm),² the light cone is slower:

$$v_F < v_{LR} ? \quad (1.7)$$

Preliminary work [11,12] noticed that in random unitary circuits, this was indeed the case. In other words, the choice of norm (or, more physically, the task of interest) quantitatively changed the ‘‘speed limits’’ on quantum dynamics.

With local interactions, this change is typically quantitative (but not qualitative). With long-range interactions, however, the discrepancy between Lieb-Robinson and Frobenius bounds becomes drastic. This perhaps explains why the Frobenius norm did not draw serious attention in the study of locally interacting systems.

B. Long-range interactions

We now turn to physical systems with power-law, or long-range, interactions. Many promising platforms for future

quantum technologies do not merely contain nearest-neighbor interactions: they can include trapped ion crystals [13], Rydberg atom arrays [14], polar molecules [15], etc.; each of these platforms has long-range interactions between all pairs of qubits which decay with the distance r as $V(r) \propto r^{-\alpha}$. Here α is an exponent characterizing the system: $0 < \alpha < 3$ can approximately be achieved in trapped ion crystals, while $\alpha = 3$ for polar molecules with dipolar interactions and $\alpha = 3, 6$ for Rydberg atoms. The classic techniques from Lieb-Robinson bounds were ineffective at constraining information dynamics in long-range systems. Over the past few years, increasingly sophisticated methods have been developed [2,16–23] to conclusively settle the optimality of Lieb-Robinson-type bounds. In a d -dimensional system, the time it takes to prepare an EPR state between two qubits, separated by distance r , scales as $t \sim r^{\min(\alpha-2d, 1)}$ for $\alpha > 2d$. This remarkable result means that even when $\alpha = 6$, in three dimensions it is possible to prepare highly entangled states in nearly constant time ($t \sim \ln r$).

It was first noted in [10] that the Lieb-Robinson bounds did not effectively constrain the Frobenius norm of a commutator in models with power-law interactions. There, it was argued that the bound on Frobenius norms would scale as

$$t_F^*(r) \gtrsim r^{\min(\alpha-3d/2, 1)}, \quad (1.8)$$

which was proved in $d = 1$ and also shown to be ‘‘optimal.’’ However, strictly speaking, the out-of-time-ordered correlators (OTOCs) which saturate (1.8) are *not* of the form $\|[A_x(t), B_y]\|_F$, with A_x and B_y simple Pauli matrices. Instead, they are of the form (in $d = 1$, for simplicity):

$$\left\| \left[A_x(t), \prod_{z \geq y} B'_z \right] \right\|_F,$$

with B'_z Pauli raising operators on sites $z \geq y$.³ However, (1.8) is not optimal if we demand A_x and B_y are both local matrices. It was recently proved in any dimension d that for $\alpha > d$, the Frobenius light cone for such an OTOC is of the form [24]

$$t \gtrsim r^{2(\alpha-d)/(2\alpha-d+1)}. \quad (1.9)$$

C. Our results

The main motivation of this paper is to reconcile the results of [10,24], and prove an optimal Frobenius bound of small operators in systems with long-range interactions, in one-dimensional (1D) models ($d = 1$). We find in Sec. IV that

$$t_F(r) \gtrsim r^{\min(\alpha-1, 1)} \quad (\alpha > 1), \quad (1.10)$$

where logarithms and absolute constants (that only depend on α) are suppressed. In Sec. V, we construct an explicit random dynamical protocol that saturates this scaling (up to factors of $\ln r$).

We conjecture that in d spatial dimensions, (1.10) generalizes to

$$t_F(r) \sim r^{\min(\alpha-d, 1)} \quad (1.11)$$

²Sometimes the Frobenius light cone $v_F \sim r/t(r)_F$ is called butterfly velocity in the quantum chaos literature.

³Products of raising operators have low rank and a very large singular value, unlike simple Pauli

for $\alpha > d$. In fact, we will present a random protocol which saturates such a bound. Our bounds together with a saturating protocol definitively settle the nature of the Frobenius light cone in $d = 1$ spatial dimensions. For technical reasons we explain in the body of the work, proving Frobenius light cones in higher dimensions is much more difficult and thus remains a conjecture for $d > 1$.

In our proof, we employ an alternative set of mathematical tools for deriving bounds on quantum dynamics with other norms in Sec. III. In a nutshell, the standard Lieb-Robinson bounds work by exploiting submultiplicativity, which states that

$$\|AB\| \leq \|A\| \cdot \|B\|. \quad (1.12)$$

The traditional Lieb-Robinson bound strategy proceeds by iteratively using this identity repeatedly to a nested commutator of the form $\|[H, [H, \dots \mathcal{O}]]\| t^n/n!$, reducing the problem of bounding quantum dynamics to a combinatorial problem. This property does *not* hold for the normalized Frobenius norm introduced in (1.4). As a consequence, existing proofs [10,25–27] of Frobenius light cones have typically involved other methods inspired by quantum walks and probability theory: these methods are, in turn, uniquely adapted to the Frobenius norm. A key development in this paper is to show that (1.12) can be *nearly* generalized to the *normalized* Frobenius norm (and also Schatten p -norms below), when the operators A and B obey certain spatial locality constraints:

$$\|AB\|_F \lesssim \|A\|_F \cdot \|B\|_F \ln \|B\|_F. \quad (1.13)$$

Our results therefore allow us to use traditional Lieb-Robinson-type techniques, but to obtain *parametrically stronger bounds* on dynamics using more physically relevant norms.

We are also able to obtain results for p -norms which interpolate between the Frobenius and operator norms defined above. We generalize the proof of [20] to any Schatten p -norms $[\text{Tr}([A_x(t), B_y]^p)/\text{Tr}(I)]^{1/p}$. We prove that there is a p -dependent light cone of the form

$$t(r)_p \gtrsim \frac{r^{\min(1, \alpha-3/2)}}{\sqrt{p}} \quad (1.14)$$

for any fixed value of p . Intuitively, these bounds constrain the p th moments of the singular value distribution of $[A_x(t), B_y]$, and therefore tell us how rare extreme singular values are. While we do not believe this bound is tight, it is already tighter than Lieb-Robinson bounds. In particular, we will show that Markov's inequality implies it is exponentially rare to find a state in Hilbert space that can saturate a Lieb-Robinson light cone.

D. Implications of our results and further conjectures

Our results have a number of interesting implications. In systems with power-law interactions, we find a remarkable result: in a spin chain with $\alpha = 2$ the growth of operators is hardly faster than with nearest-neighbor interactions, while existing state transfer protocol may send a single qubit in time $t \sim \ln r$ [22]! At the same time, the rapid state transfer algorithm of [22] requires an *exponentially* fine-tuned background state to implement in quantum processors. Our results

therefore have important practical implications for the wide range of possible quantum technologies, such as trapped ion crystals, which might seek to use power-law interactions to speed up the transfer of information. It will be interesting to understand the actual performance of the protocol of [22] in the presence of some errors.

We conjecture that the Frobenius light cone is an *upper bound* on the time it takes to generate a volume-law entangled state. Intuitively, generating volume-law entanglement requires taking all local information and scrambling it amongst highly nonlocal degrees of freedom; thus, a local operator must evolve into a nonlocal one. Since operator growth is captured by the Frobenius norm, the Frobenius light cone should also control entanglement. Support for this conjecture is found in the examples of [28].

While we believe this intuitive argument to be correct, we have not been able to formalize a proof of this result. Nevertheless, if this conjecture is correct, it would immediately imply a very surprising result: namely, the (asymptotically) optimal way to send volume-law entangled states using power-law interactions is to send one qubit at a time. Indeed, the protocol of [22] shows that we can send one qubit in a time $r^{\alpha-2d}$ for $2d < \alpha < 2d + 1$. However, if volume-law entanglement can only be generated in time $r^{\alpha-d}$, that means that sending one qubit at a time using the fast protocol of [22] takes a time $r^{\alpha-2d} \times r^d \sim r^{\alpha-d}$, which is the fastest possible. This observation can simplify the design of any future quantum processor performing fast quantum state transfer with power-law interactions.

Another implication of our results is that the Frobenius bound we present is essentially optimal even when the Hamiltonian is drawn randomly; in fact, our saturating protocol is designed using an ensemble of saturating Hamiltonians, which are themselves random. In contrast, typical random Hamiltonians have sharper operator norm bounds than deterministic Hamiltonian [29]. This demonstrates that the Frobenius bounds are more robust than the operator norm, in that much less fine tuning of Hamiltonians is needed to saturate a Frobenius bound. In this sense, Frobenius bounds are slightly more universal than Lieb-Robinson bounds.

Our mathematical results may help to improve existing “many-body quantum walk” techniques for bounding OTOCs in a variety of models [25–27,30,31]. A particularly challenging problem where our methods may be extremely valuable is bounding finite-temperature correlators, where the appropriate “norm” depends on the thermal density matrix [32]. Understanding how to derive specific notions of submultiplicativity, which might hold under this norm, but not in general, may help to solve longstanding conjectures about the speed limits on finite-temperature dynamics which have, in recent years, linked developments in string theory and quantum gravity to many-body physics [33–35].

Lastly, our bounds on Frobenius light cones may help to find sharper error bounds on new numerical techniques such as matrix product operator methods [36] since so long as a thermal correlator is being studied, one is not interested in the error in the numerical evaluation of an operator in the worst-case state, but only in a typical state in the ensemble.

II. PRELIMINARIES

In this section we review formalisms in operator dynamics and useful facts from noncommutative functional analysis.

A. Long-range interactions

We first provide a careful definition for long-range interacting systems. Let Λ denote the vertices of a lattice graph, consisting of some unit cell repeated periodically in d spatial dimensions.⁴ For vertices $i, j \in \Lambda$, let us define the distance $d(i, j)$ between the two vertices to be the Manhattan distance (minimal number of edges to traverse to get from i to j). For ease of presentation, let us consider a many-body Hilbert space consisting of a single qubit (two-level system) on every lattice site. The Hilbert space of operators acting on a single qubit on site x is spanned by the identity I_x , and the three Pauli matrices X_x, Y_x, Z_x . In what follows, we denote with $X^a = (X, Y, Z)$ the set of all nontrivial Paulis.

Given a 2-local Hamiltonian

$$H(t) = \sum_{x,y \in \Lambda} \sum_{a,b=1}^3 J_{xy}^{ab}(t) X_x^a X_y^b, \quad (2.1)$$

we say that it has power-law interactions of exponent α if

$$|J_{xy}^{ab}(t)| \leq \frac{1}{d(x, y)^\alpha}. \quad (2.2)$$

Without loss of generality, we set the prefactor above to be 1 by rescaling the Hamiltonian. In practice, it is useful to say a model has exponent α if it does not have exponent $\alpha' > \alpha$, in our bounds the only criterion necessary is (2.2).

B. Operator norms and operator size

In this paper, we are interested in the Heisenberg time evolution of an operator

$$\frac{dA(t)}{dt} := i[H(t), A] := \mathcal{L}(t)[A(t)]. \quad (2.3)$$

It will be helpful to interpret operator dynamics in a bra-ket notation on the ‘‘Hilbert space’’ of operators, in which case we write the above equation as

$$\frac{d}{dt}|A(t)\rangle := \mathcal{L}(t)|A(t)\rangle. \quad (2.4)$$

We define a Frobenius inner product on this space. Loosely speaking,⁵

$$(A|B) := \frac{\text{tr}(A^\dagger B)}{\text{tr}(I)}. \quad (2.5)$$

It is often helpful to define projection operators. For example, the superoperator \mathbb{P}_x , defined via

$$\mathbb{P}_x A = -\frac{1}{8} \sum_{a=1}^3 [X_x^a, [X_x^a, A]], \quad (2.6)$$

allows us to restrict to all terms in A which act nontrivially on site x . Since

$$\begin{aligned} \|[B_y, A]\|_F &= \|[B_y, \mathbb{P}_y A]\|_F \leq 2\|B_y\|_\infty \|\mathbb{P}_y A\|_F \\ &= 2\|B_y\|_\infty \sqrt{(A|\mathbb{P}_y A)}, \end{aligned} \quad (2.7)$$

it suffices to bound the (Frobenius) norm of $\mathbb{P}_y A$, in order to bound OTOCs.

In fact, without loss of generality, we may write any operator A as

$$A = \sum_{S \subset \Lambda} A_S(t), \quad \text{where } \mathbb{P}_x A_S(t) = \begin{cases} A_S(t), & x \in S \\ 0, & \text{otherwise.} \end{cases} \quad (2.8)$$

Assuming $(A|A) = 1$ is normalized, we will then refer to

$$p_S = (A_S|A_S) \quad (2.9)$$

as the probability that the operator is supported on subset S . Clearly,

$$\sum_S p_S = 1, \quad (2.10)$$

so this terminology is well defined. This probabilistic language will be highly valuable for us as we develop a fast operator growth protocol.

While the most important operator norms in this paper are the Frobenius norm (1.4), along with the operator norm (1.3), we will also find useful the Schatten p -norms⁶ that interpolate in-between the Frobenius and operator norms:

$$\|X\|_p := \text{Tr}[(X^\dagger X)^{p/2}]^{1/p}. \quad (2.11)$$

⁴Mathematically, if E_Λ is an edge set on Λ , we demand that the pair (Λ, E_Λ) has an automorphism group containing a translation subgroup \mathbb{Z}^d .

⁵On an infinite lattice, this ratio is not well defined. We may interpret it so long as the operators A and B have compact support, in which case the traces may be restricted to the support of A and B . Taking such an operator A with compact support, we may also time evolve it. With a Hamiltonian H without compact support, the operator $A(t)$ will (in general) also no longer have compact support. In this case, we may define $(A|B)$ via a limiting process: letting A_S denote the terms in A supported on compact set S , the limit of $(A_S|B_S)$ will converge as S grows; the convergence is guaranteed by Lieb-Robinson theorems.

⁶As before, one should rigorously only apply this definition on a finite-dimensional Hilbert space. We will implicitly invoke the limiting process explained above in this paper, in order to make sense out of the p -norm on an infinite-dimensional space.

We will often deal with the normalized p -norm divided by $\text{tr}(I)$ that

$$\|X\|_{\bar{p}} \leq \frac{\|X\|_p}{\|I\|_p} \leq \|X\|_{\infty} \leq \|X\|_p. \quad (2.12)$$

To show an approximate form of submultiplicativity for the Frobenius norm, we will also find useful the following standard inequality from functional analysis:

Proposition 1 (Noncommutative Hölder’s inequality [29]). If

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}, \quad (2.13)$$

then

$$\|AB\|_p \leq \|A\|_{p_1} \|B\|_{p_2}. \quad (2.14)$$

Additionally, we have the following:

Proposition 2 (Riesz-Thorin interpolation). If

$$\frac{1}{q\theta} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2}, \quad (2.15)$$

then

$$\|A\|_{q\theta} \leq \|A\|_{q_1}^{\theta} \|A\|_{q_2}^{1-\theta}. \quad (2.16)$$

III. TOWARDS SUBMULTIPLICATIVITY FOR THE FROBENIUS NORM

In this section, we develop a mathematical machinery which will be critical to proving the Frobenius light cone, and its p -norm extensions. In the main text, we will focus on the simpler case of two-body (2-local) Hamiltonians; the straightforward extension to k -local Hamiltonians is provided in Appendix A.

The following lemma represents the key technical result of this section:

Lemma 1. Consider the 2-local Hamiltonian

$$H = \sum_{i<j} H_{ij} \quad (3.1)$$

with each H_{ij} supported on exactly sites i and j [$\mathbb{P}_k H_{ij} = \mathbb{I}(k \in \{i, j\})H_{ij}$], and $\|O\| \leq 1$, then

$$\|HO\|_F \leq 2e\|H\|_{(2)}\|O\|_F(\ln\|O\|_F + 1), \quad (3.2)$$

where the Frobenius norm is normalized as in (1.4), and

$$\|H\|_{(2)} := \sqrt{\sum_{i<j} \|H_{ij}\|_{\infty}^2}. \quad (3.3)$$

Before diving into the proof, let us put this result into some context and explain why this refinement is valuable. For simplicity, consider an operator O of diameter r . As this paper is motivated by long-range interacting systems, we consider the effect of long-range interactions with another sphere of diameter r , a distance $O(r)$ away (see Fig. 1). The leading-order type of operator growth is

$$e^{iHt} O e^{-iHt} \approx O + it[H, O] + \dots, \quad (3.4)$$

and an unconditional bound would be

$$\|HO\|_F \leq \|H\|_{\infty} \|O\|_F. \quad (3.5)$$

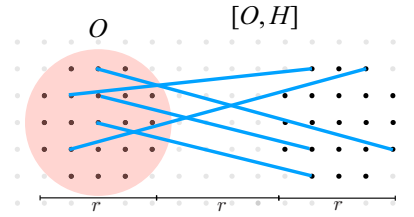


FIG. 1. Long-range interactions acting on a large operator between two balls of radius r . The interaction Hamiltonian between these two balls has norms given in (III.6). The key technical advance of this paper is to learn how to take advantage of the smaller Frobenius norm of an interaction Hamiltonian in controlling the Frobenius light cone.

If we simply wish to bound $\|O\|_F$, this triggers a recursive identity; however, it comes at the price of a large prefactor of $\|H\|_{\infty}$. Indeed, observe that between two balls of size r a distance r apart [10],

$$\|H\|_{\infty} \lesssim \sum_{x \in \text{ball } 1} \sum_{y \in \text{ball } 2} \frac{1}{d(x, y)^{\alpha}} \sim \frac{1}{r^{\alpha-2d}}, \quad (3.6a)$$

$$\begin{aligned} \|H\|_{(2)} \sim \|H\|_F &= \frac{\|H\|_2}{\|I\|_2} \lesssim \sqrt{\sum_{x \in \text{ball } 1} \sum_{y \in \text{ball } 2} \left(\frac{1}{d(x, y)^{\alpha}}\right)^2} \\ &\sim \frac{1}{r^{\alpha-d}}. \end{aligned} \quad (3.6b)$$

This equation suggests it is highly desirable to use $\|H\|_{(2)}$, not $\|H\|_{\infty}$, if at all possible. Yet (3.5) suggests we must always use the worst-case scenario, set by the operator norm.⁷

An observation in [24] was that the Frobenius norm can behave submultiplicatively when the other operator has bounded spectrum: as in (2.7),

$$\|HO\|_F \leq \|H\|_F \|O\|_{\infty}. \quad (3.7)$$

If $\|O\|_{\infty} \approx \|O\|_F$, as is the case when O is a simple Pauli matrix ($O^2 = 1$), then using this inequality gives us an approximate form of submultiplicativity, one time. Of course, clearly this does not feed in to a recursive bound; indeed in [24] this trick was only used once. But, our refined bound is written entirely in terms of the Frobenius-type norm $\|H\|_{(2)}$ (except for the global constraint $\|O\| \leq 1$ inherited from the initial operator). This is more powerful and allows us to obtain much stronger Frobenius light cone bounds. We now present the proof of the lemma.

Proof. Using Proposition 1,

$$\|HO\|_2 = \|HO\|_2 \leq \|H\|_{2/(1-\theta)} \cdot \|O\|_{2/\theta}. \quad (3.8)$$

We now bound each of these norms in turn. The idea is to choose $\theta \rightarrow 1$, without making the norm of $\|H\|_{2/(1-\theta)}$ too large. Hence, we need to prove that the p -norms $\|H\|_p/\|I\|_p$ grow sufficiently slowly (when normalized by the Hilbert

⁷Note that the normalization is absolutely crucial, otherwise the Schatten 2-norm is even looser than the operator norm, and is already submultiplicative: $\|X\|_2/\|I\|_2 \leq \|X\|_{\infty} \leq \|X\|_2$.

space dimension). In other words, we need to prove bounds on the tails of the singular distribution of H . In the mathematics and probability literature, this is called a concentration bound.

More precisely, to bound $\|H\|_q$, $q := 2/(1 - \theta)$, we will use a seemingly simple recursive moment inequality, proved in Appendix A:

Proposition 3 (Uniform smoothness for subsystem).

Consider matrices X, Y of the same dimensions that satisfy $\text{Tr}_i(Y) = 0; X = X_j \otimes I_i$. For $q \geq 2$,

$$\|X + Y\|_q^2 \leq \|X\|_q^2 + C_q \|Y\|_q^2 \quad (3.9)$$

for optimal constant $C_q = q - 1$.

Hence,

$$\begin{aligned} \left\| \sum_{i < j} H_{ij} \right\|_q^2 &\leq \left\| \sum_{1 < i < j} H_{1j} \right\|_q^2 + C_q \left\| \sum_{1 = i < j} H_{ij} \right\|_q^2 \\ &\leq \sum_i C_q \left\| \sum_{i < j} H_{ij} \right\|_q^2 \\ &\leq \sum_{i < j} (C_q)^2 \|H_{ij}\|_q^2 \end{aligned}$$

$$\|HO\|_F = \frac{1}{\|I\|_2} \|HO\|_2 \leq \frac{1}{\|I\|_2} \left[\left(\frac{2}{1 - \theta} - 1 \right) \|H\|_{(2)} \|I\|_{2/(1 - \theta)} \right] \|O\|_2^\theta \leq \left(\frac{2}{1 - \theta} \right) \|H\|_{(2)} \left(\frac{\|O\|_2}{\|I\|_2} \right)^\theta, \quad (3.12)$$

where we explicitly display the $\|I\|_2$ and used that $\|I\|_2 = (\text{Tr}[I])^{1/2}$ to normalize $\|O\|_2$. Now we choose a value of θ close to 1:

$$n = \frac{1}{1 - \theta} := \max(1, |\ln(\|O\|_F)|) \leq |\ln(\|O\|_F)| + 1. \quad (3.13)$$

(Note that $\|O\|_F \leq 1$ since $\|O\|_\infty \leq 1$.) Then,

$$\begin{aligned} \|HO\|_F &\leq \|H\|_{(2)} \|O\|_F \times n \|O\|_F^{-1/n} \\ &\leq 2 \|H\|_{(2)} \|O\|_F \times n \exp(|\ln \|O\|_F|/n). \end{aligned} \quad (3.14)$$

Combining (3.13) and (3.14), we obtain the advertised result. ■

In Appendix A, we extend Lemma 1 to both p -norms and to k -local Hamiltonians. The price of each of these extensions is a slightly worse prefactor in (A9). However, these prefactors are mild enough that we can still prove Theorem 2: a partial extension of the Frobenius light cone to any p -norm.

In Appendix B, we extend our results to an additional family of physical norms, inspired by [31,40], which allow us to straightforwardly extend the results of Appendix A to certain norms which induce physical correlation functions in thermodynamic ensembles.

These two additional sets of results demonstrate the flexibility of our methods.

$$\leq (C_q)^2 \|I\|_q^2 \sum_{i < j} \|H_{ij}\|_\infty^2 = (C_q)^2 \|I\|_q^2 \|H\|_{(2)}^2. \quad (3.10)$$

In the first line we applied Proposition 3 for $Y = \sum_{1 = i < j} H_{ij}$ and $X = \sum_{1 < i < j} H_{1j}$, i.e., “peeling off” any term traceless on the qudit $i = 1$. In the second inequality we peel off qudit $i = 2, 3, \dots$ analogously. In the third line, we recursively repeat the first and second lines for the j index, which gives a second factor of C_q . Lastly, we use Holder to bound the q -norm by the ∞ norm with a $\|I\|_q$ overhead.⁸

For the norm of O , a standard manipulation using Proposition 2, and $\|O\|_\infty \leq 1$, allows us to write

$$\|O\|_{2/\theta} \leq \|O\|_2^\theta \cdot \|O\|_\infty^{1 - \theta} \leq \|O\|_2^\theta. \quad (3.11)$$

Now combining (1.4), (3.8), (3.10), and (3.11), we plug in $q = 2/(1 - \theta)$ to get

IV. LIGHT CONES

Now we turn to our main application of the mathematics developed above: the Frobenius light cone for a one-dimensional system with power-law interactions.

A. Frobenius light cone

For simplicity, suppose that we study a spin chain with a qubit on each site. Let us define a set of projectors \mathbb{Q}_i for $i = 0, 1, \dots, R$ (here R is an integer capturing the number of sites on the chain we wish to study). On Pauli strings of the form $X_0^{a_0} X_1^{a_1} \dots$ with $a_0 = 0$ denoting I and $a_0 = 1, 2, 3$ denoting X_0, Y_0, Z_0 , we define

$$\begin{aligned} \mathbb{Q}_x \{ \{ X_i^{a_i} \} \} = & \begin{cases} \mathbb{I}(a_y = 0 \text{ if } y > 0) \{ X_i^{a_i} \}, & x = 0 \\ \mathbb{I}(a_x > 0) \mathbb{I}(a_y = 0 \text{ if } x < y) \{ X_i^{a_i} \}, & 0 < x < R \\ \mathbb{I}(a_y \neq 0 \text{ for some } y \geq R) \{ X_i^{a_i} \}, & x = R. \end{cases} \end{aligned} \quad (4.1)$$

Here \mathbb{I} denotes the indicator function which returns 1 if its argument is true, and 0 if false. The projector \mathbb{Q}_x hence projects onto operators which act nontrivially on site x , but trivially on all sites to the right of x . The exceptions are \mathbb{Q}_0 and \mathbb{Q}_R , where we simply include all operators supported on a site $\geq R$ into \mathbb{Q}_R , and all operators supported on ≤ 0 in \mathbb{Q}_0 . Note that

$$\sum_{i=0}^R \mathbb{Q}_i = 1. \quad (4.2)$$

For more general qudit systems, the extension of these definitions is straightforward (albeit notation can get clunkier).

⁸This also has the name of hypercontractivity that was proven in some special cases [37,38]. We rederive it using uniform smoothness which also has wider applications. See [39] for a further discussion.

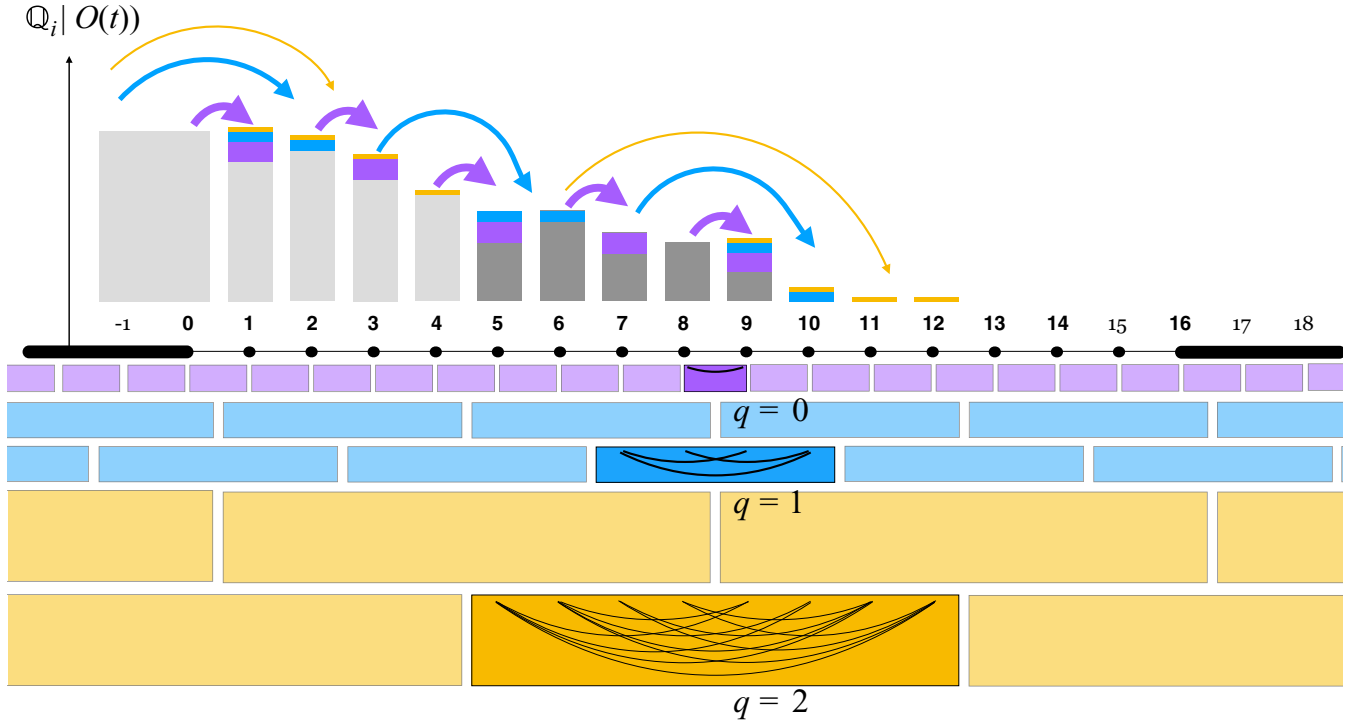


FIG. 2. We organize a growing operator $|A_0(t)\rangle$ according to the site x on which the rightmost nontrivial Pauli acts [achieved via projectors $\mathbb{Q}_x|A_0(t)\rangle$]. We then regroup the Hamiltonian into scales as in [20]. We then study the quantum walk of the weights $\mathbb{Q}_x|A_0(t)\rangle$, which are efficiently bounded one scale at a time.

With these projectors defined, we may now state our main result, which is the Frobenius light cone for spin chains with power-law interactions:

Theorem 1. Let $H(t)$ be a time-dependent Hamiltonian which has power-law exponent $\alpha > 1$, as defined in Sec. IV A; let

$$R = 2^{q_*+1} - 1 \tag{4.3}$$

for some positive integer q_* . Then for any $\delta > 0$, if $\mathbb{Q}_0|A_0\rangle = |A_0\rangle$, and

$$\|\mathbb{Q}_R|A_0(t)\rangle\|_F \geq \delta, \tag{4.4}$$

then there exists a α -dependent constant $0 < K_\alpha < \infty$ such that

$$|t| \geq \delta^2 K_\alpha \times \begin{cases} R/\ln R, & \alpha > 2 \\ R/\ln^2 R, & \alpha = 2 \\ R^{\alpha-1}, & 1 < \alpha < 2. \end{cases} \tag{4.5}$$

Note that in the statement of this theorem, we choose R to be (nearly) a power of 2 for the conceptual simplicity of Fig. 2. From the definition in (4.1), we can always just choose $r/2 < R \leq r$ if we wish to get a bound on the light cone at any distance r which is not of the form (4.3).

Proof. The method of proof is adapted from the ‘‘quantum walk’’ proofs developed in [10,25–27,30,31]. In a nutshell,

we can think of $[A_0(t)|\mathbb{Q}_x|A_0(t)]$ as the probability that the operator has made it as far right as site x . We then try to bound how quickly $[A_0(t)|\mathbb{Q}_x|A_0(t)]$ can grow with x and t . We do this using Markov’s inequality, by bounding the *expected rightmost site* of the operator, and noting that

$$(A_0(t)|\mathbb{Q}_R|A_0(t)) \leq \frac{\mathbb{E}[\text{rightmost site } x \text{ at time } t]}{R}. \tag{4.6}$$

To formalize this last equation, we define the following superoperator \mathcal{F} which acts on an operator and returns its rightmost site:

$$\mathcal{F} := \sum_{i=0}^R i\mathbb{Q}_i. \tag{4.7}$$

Our goal is to prove that at sufficiently short t ,

$$\delta^2 \leq \|\mathbb{Q}_R|A_0(t)\rangle\|_F^2 \leq \frac{1}{R}(A_0(t)|\mathcal{F}|A_0(t)). \tag{4.8}$$

We thus must bound $d/dt(A|\mathcal{F}|A)$. This is achieved by the following lemma:

Lemma 2. For constants $0 < C < \infty$,

$$\frac{d}{dt}(A_0(t)|\mathcal{F}|A_0(t)) \leq \begin{cases} C(q_* + 1)(4 + (1 + \frac{q_*}{2})\ln 2), & \alpha = 2 \\ \frac{C}{(1 - 2^{2-\alpha})^2}((4 + \ln 2)(2^{2-\alpha} - 1)(2^{(2-\alpha)q_*} - 1) + \ln 2(2^{(2-\alpha)q_*} + q_* - (1 + q_*)2^{2-\alpha})) & \text{else.} \end{cases} \tag{4.9}$$

Proof. The rigorous proof is in Appendix C; here we sketch the main ideas. We divide up the Hamiltonian into a set of scales q , with the rough intuition that couplings on scale q have lengths of order 2^q . Using some identities special to one-dimensional models, we are able to rewrite

$$\begin{aligned} \frac{d}{dt}(A_0(t)|\mathcal{F}|A_0(t)) &= (A_0(t)|[\mathcal{F}, \mathcal{L}]|A_0(t)) \\ &= \sum_{q=0}^{q_*} \sum_k (A_0(t)|[\mathcal{F}, \mathcal{L}_{q,k}]|A_0(t)) \end{aligned} \quad (4.10)$$

as a sum of similar inner products at each scale q ; here q_* denotes the maximum scale $q_* \sim \ln r$, and k denotes which “block” of couplings of size $\sim 2^q$ we are studying. These blocks are shown in Fig. 2. At each scale, we can efficiently resum up the magnitude of each $\mathcal{L}_{q,k}$ using Lemma 1. We find that each factor of $\mathcal{L}_{q,k}$ contributes $(2^q)^{1-\alpha}$, following the logic of (3.6b). When $\alpha \geq 2$, the logarithmic factors in (A9) cannot be neglected, and lead to an additional logarithmic enhancement in (4.9). ■

Using the result of this lemma, we see that there exists an α -dependent constant K_α such that

$$\frac{d}{dt}(A_0(t)|\mathcal{F}|A_0(t)) \leq \begin{cases} K_\alpha \ln R, & \alpha > 2 \\ K_2 \ln^2 R, & \alpha = 2 \\ K_\alpha R^{2-\alpha}, & 1 < \alpha < 2. \end{cases} \quad (4.11)$$

Indeed, looking at (4.9), when $\alpha = 2$, the constant scales as $q_*^2 \sim \ln^2 R$; when $\alpha > 2$, the term linear in q_* dominates and leads to $\ln R$ scaling; when $\alpha < 2$, the $2^{(2-\alpha)q_*} \sim R^{2-\alpha}$ terms dominate.

Combining (4.11) and (4.8) we obtain (4.5). ■

As we will show in the next section, the scaling of Theorem 1 is optimal. We thus arrive at the surprising result that it is possible to have a Lieb-Robinson velocity which grows exponentially with time (at $\alpha = 2$), while the butterfly velocity of the Frobenius light cone grows no faster than $\ln^2 t$.

We note that it is quite challenging to extend this result to higher dimensions. The reason is that on a one-dimensional (1D) chain, it is very simple to define the projectors \mathbb{Q}_x to classify how far to the right an operator has grown. For $d > 1$, we have tried to define analogous projectors keeping track of the “diameter” of a growing operator, yet because there are L^{d-1} sites near the edge of a ball of diameter L , the transition rates between sectors \mathbb{Q}_L and \mathbb{Q}_{L+1} will grow with L . The more sophisticated ansatz necessary to use quantum walk bounds to constrain higher-dimensional models (with power-law interactions) has not yet been found. An analogous challenge was present in the original proof of a linear light cone in 1D systems with power-law interactions [20]; a much more complicated proof [21] was necessary to prove the linear light cone in higher d . We would not be surprised if a similar lengthy extension is required to generalize Theorem 1 to $d > 1$; however, we expect that our Lemma 1 will be at the heart of any generalization of this work.

B. Lieb-Robinson light cone is rarely tight

A natural question to ask is whether the Frobenius light cone or the Lieb-Robinson light cone is more relevant for a “typical” physical state. Naively, one would expect the Frobe-

nius light cone to be more relevant. Indeed, we can take a probabilistic interpretation of the Frobenius norm:

$$\|A\|_F^2 = \text{Tr}[\mathbb{E}_\psi(|\psi\rangle\langle\psi|)A^\dagger A] = \mathbb{E}_\psi[\|A|\psi\rangle\|_{\ell_2}^2] \quad (4.12)$$

for any ensemble of pure states that average to the maximally mixed $\mathbb{E}_\psi(|\psi\rangle\langle\psi|) = I/\text{Tr}[I]$. Then, we can obtain a concentration inequality via Chebychev’s inequality:

$$\mathbb{P}(\|A|\psi\rangle\|_{\ell_2} \geq a) \leq \frac{\|A\|_F^2}{a^2}. \quad (4.13)$$

In words, a state drawn “randomly” from the ensemble will mostly like be of order $O(\|A\|_F)$. However, $1/a^2$ dependence is not the strongest concentration we can ask for.

As it turns out, a sharper concentration inequality would be possible if we further obtained bounds on the Schatten p -norms of A , for tunable values of p . Let us denote normalized p -norm by

$$\|O\|_{\bar{p}} := \frac{\|O\|_p}{\|I\|_p} \quad (4.14)$$

that avoids tedious and exponentially large normalization constants when discussing the p -norms of local operators in a many-body system. We now note the following fact:

Proposition 4 (p -norm and typical states [39]). For any operator A , when uniformly averaging over states $|\psi\rangle$ that $\mathbb{E}|\psi\rangle\langle\psi| = I/\text{Tr}[I]$,

$$\mathbb{P}(\|A|\psi\rangle\|_{\ell_2} \geq a) \leq \left(\frac{\|A\|_{\bar{p}}}{a}\right)^p. \quad (4.15)$$

And we obtain a p -norm estimate as follows.

Theorem 2. Let $H(t)$ be a time-dependent Hamiltonian which has power-law exponent $\alpha > 1$, as defined in Sec. IV A. Then under the Heisenberg equation of motion generated by this $H(t)$,

$$\|\mathbb{Q}_R A_0(t)\|_p \leq c' \sqrt{p} \frac{|t|}{\mathcal{R}(r)} \quad (4.16)$$

for constant $0 < c'(\alpha) < \infty$ which only depends on α , and

$$\mathcal{R}(r) = \begin{cases} r, & \text{if } \alpha > 5/2 \\ r/\ln^{3/2}(r), & \text{if } \alpha = 5/2 \\ r^{\alpha-3/2}, & \text{if } 5/2 > \alpha > 3/2. \end{cases} \quad (4.17)$$

Alternatively, we may write that $\|\mathbb{Q}_R A_0(t)\|_p \geq \delta$ is only possible if

$$|t| \geq \delta \sqrt{p} c' \mathcal{R}(r). \quad (4.18)$$

Proof. The technical details of the proof are quite similar in spirit to the proof of [20], which itself uses the same decomposition of H into different scales shown in Fig. 2. The difference between the proof of this theorem and Theorem 1 is that we are unable to use the quantum walk bounds to tightly control the growth of $\mathbb{Q}_R|A_0(t)\rangle$ because $\|\mathbb{Q}_x|A_0(t)\rangle\|_{\bar{p}}$ does not obey a quantum walk equation. We give the proof in Appendix D. ■

We do not expect this bound to be tight; clearly it is not for $p = 2$, in which case Theorem 1 is already stronger. Nevertheless, it already suffices to guarantee a meaningful concentration bound:

Corollary 1. For power-law models with exponent $2 < \alpha < 3$, the Lieb-Robinson light cone is rarely saturated: for $0 < \epsilon < \infty$ and sufficiently large r ,

$$\mathbb{P}\left(\|[A_0(t), B_r]|\psi\rangle\|_{\ell_2} \geq \frac{\epsilon t}{r^{\alpha-2}}\right) \leq \exp(2 - \epsilon^2 C r^\beta) \quad (4.19)$$

for some constant $0 < C(\alpha) < \infty$ which does not depend on r , and with $\beta = \min(1, 6 - 2\alpha) - \delta$ for arbitrary $\delta > 0$. This result holds even if $H(t)$ is time dependent.

Note that the infinitesimal parameter δ is only due to the logarithm $\ln(r)$ at $\alpha = \frac{5}{2}$, and should not distract the reader.

Proof. Plug the p -norm bound of Theorem 2 into Proposition 4 to obtain

$$\begin{aligned} \mathbb{P}\left(\|[A_0(t), B_r]|\psi\rangle\|_{\ell_2} \geq \frac{\epsilon t}{r^{\alpha-2}}\right) &\leq \left(\frac{r^{\alpha-2}}{\epsilon t} \sqrt{p} \frac{c't}{\mathcal{R}(r)}\right)^p = \left(\sqrt{p} \frac{c'r^{\alpha-2}}{\epsilon \mathcal{R}(r)}\right)^p \\ &\leq \exp\left(2 - \epsilon^2 \frac{\mathcal{R}(r)^2 r^{4-2\alpha}}{e^2 c'^2}\right), \end{aligned} \quad (4.20)$$

where in the last inequality we used that $(\sqrt{pb})^p \leq \exp(2 - e^{-2} b^{-2})$ [set $p = \min(2, (eb)^{-2})$]. The $+2$ in the exponential is added since to use our concentration bounds we must always take $p \geq 2$. We conclude the proof by noting that for any $2 < \alpha < 3$, $r^{4-2\alpha} \mathcal{R}(r)^2 > r^\beta$ if r is sufficiently large. ■

Remarkably, even though the form of the bound in Theorem 2 is not tight, we find an exponential bound on the number of states in which the Lieb-Robinson bounds might be saturated. In the regime $2 < \alpha < \frac{5}{2}$, this bound is optimal (up to constants in the exponent): the protocol of [22] shows that in a spin chain of qubits, it is possible to saturate the Lieb-Robinson light cone in at least one state. What Corollary 1 proves is that the state identified in [22] is in fact one of the *exponentially* rare states for which the Lieb-Robinson light cone can be saturated! This strongly suggests that the rapid single-bit state transfer protocol of [22], which can transfer one qubit in time $t \sim r^{\alpha-2d}$, may require exquisite control of the background state; this contrasts with the more robust (yet slower) protocol of [10,41], which uses long-range interactions to send qubits in a “self-error-correcting” scheme. It is an interesting open question to more quantitatively compare fast quantum error-correction schemes to the optimal Frobenius and Lieb-Robinson light cones which have been developed over the past few years.

Lastly, we conjecture that the optimal bound on p -norms takes the form

$$\|[A_0(t), B_r]\|_{\tilde{p}} \stackrel{?}{\lesssim} \frac{pt}{r^{\alpha-1}}. \quad (4.21)$$

This is an educated guess: it extrapolates to the $p = 2$ case, and that the concentration holds at $e^{-\Omega(r)}$ though Corollary 1. We do not currently know how and if our quantum walk based proof of Theorem 1 can be generalized to a p -norm bound.

V. ALGEBRAICALLY OPTIMAL OPERATOR GROWTH PROTOCOL

Having established the Frobenius light cone rigorously in $d = 1$, and with a conjecture on how it generalizes to higher

dimensions d , let us now describe a protocol which we claim (and will subsequently prove) achieves these speed limits (up to subalgebraic corrections). Our approach is loosely inspired by the optimally fast single-qubit state transfer protocol developed in [22]: as in [22], we will develop our protocol via “recursive” intuition.

A. Intuitive argument

To begin, let us assume that we have a system with tunable and time-dependent power-law interactions of exponent $d < \alpha < d + 1$ on the standard hypercubic lattice \mathbb{Z}^d . We divide up this lattice into a partition of hypercubes at multiple scales $q = 0, 1, 2, \dots$. For intuitive purposes, we can say that the scale $q = 0$ corresponds to each lattice site being in its own cube; scale $q = 1$ corresponds to a partition of the lattice into hypercubes of side length m_1 in all dimensions; scale $q = 2$ corresponds to a partition into hypercubes of side length $m_1 m_2$ in all dimensions, and so on. (Note that the $q = 2$ cubes contain m_2 $q = 1$ cubes within them, etc.). In the discussion that follows, we will for simplicity set $m_1 = m_2 = \dots = m$.

At time $t = 0$, we start with a Pauli matrix X_0 on a single site. By the definitions above, that Pauli matrix occupies exactly one 0-cube, which we might as well call the origin of the lattice. Our goal is to find a quantum mechanical protocol [i.e., a unitary matrix $U(t)$, which can be generated from a power-law Hamiltonian evolving for time t , possibly with time-dependent coefficients], such that $U(t)^\dagger X_0 U(t)$ consists of Pauli strings of length L^d . If we achieve this, then we know that this operator must have support on at least one site a distance $\sim L$ away from the origin where we started. Our goal is to do this in a time $t \sim L^{\alpha-d}$ for $d < \alpha < d + 1$.

To motivate how we might achieve this task, suppose that we have a quantum protocol, a unitary matrix U_q , which is capable of taking any single-site Pauli matrix (e.g., X_0) and evolving it into an operator $U_q^\dagger W U_q$ which is supported on an $O(1)$ fraction of sites in the q -cube C_q :

$$\sum_{x \in C_q} \|\mathbb{P}_x U_q^\dagger W U_q\|_F^2 \propto R_q^d, \quad (5.1)$$

with R_q^d being the number of sites in C_q , and

$$R_q = m R_{q-1} = m^q. \quad (5.2)$$

We can certainly do this in constant time when $q = 1$ by simply using nearest-neighbor interactions.

Now, assuming that we found U_q , let us find a U_{q+1} at the next scale. First, note that any q -cube C_q lies entirely within a $(q + 1)$ -cube C_{q+1} . How can we find a unitary U_{q+1} such that (5.1) continues to hold at scale $q + 1$? One possible way to do this would be

$$U_{q+1} = U_q V_{q+1} U_q, \quad (5.3)$$

where V_{q+1} is a unitary that takes an operator supported in a single q -cube, and evolves it to have support on only a single site in each of the other $m^d - 1$ q -cubes in C_{q+1} . The three-step process is sketched in Fig. 3.

The key observation is that this “recursive” construction is quite natural to implement with power-law interactions. As the scale $q = 0 \rightarrow 1$ unitary U_1 is quite easy to implement (even using only nearest-neighbor interactions), let us focus on how

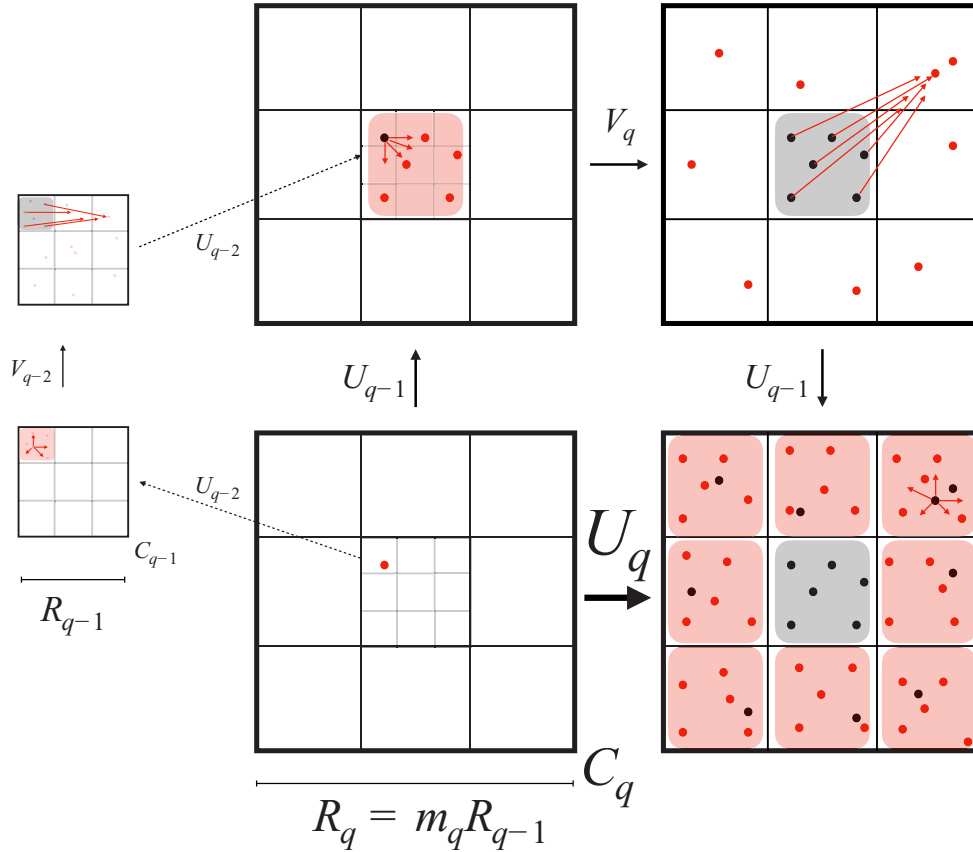


FIG. 3. The recursive step in the idealized protocol, constructing U_q from U_{q-1} . The new sites are shaded lighter. (1) Grow a single Pauli W to have support on $\sim |C_q|$ sites using the first U_{q-1} . (2) “Seed” a single Pauli matrix (e.g., Z_x) on one site x in each other q -cube $C'_q \subset C_q$ using V_q . (3) Run again the U_{q-1} protocol on each remaining q -cube C'_q to “bloom” the seeded single operator into a finite fraction of sites. On the left is the description of U_{q-1} .

V_{q+1} might effectively be implemented with power-law interactions. Suppose that as an idealistic cartoon, we found that

$$U_q^\dagger X_0 U_q = \prod_{x \in C_q} X_x. \quad (5.4)$$

Now let us consider

$$V_{q+1} = \exp \left[-i \tau_{q+1} \sum_{x,y \in C_{q+1}} \frac{Z_x Z_y}{(dR_{q+1})^\alpha} \right]. \quad (5.5)$$

Clearly, this Hamiltonian is compatible with (2.2). When τ_{q+1} is small, we may estimate that the fraction of the operator with support outside of the original C_q is given by the first-order expansion

$$\begin{aligned} & \left\| \tau_{q+1} \left[\sum_{x,y \in C_{q+1}} \frac{Z_x Z_y}{(dR_{q+1})^\alpha}, \prod_{x \in C_q} X_x \right] \right\|_{\mathbb{F}}^2 \\ &= \left\| \sum_{x \in C_q, y \in C_{q+1} - C_q} \frac{2\tau_{q+1} Y_x Z_y}{(dR_{q+1})^\alpha} \prod_{z \in C_q - x} X_z \right\|_{\mathbb{F}}^2 \\ &= \frac{4\tau_{q+1}^2}{(dR_{q+1})^\alpha} R_q^d (R_{q+1}^d - R_q^d). \end{aligned} \quad (5.6)$$

Taking

$$\tau_{q+1} \gtrsim R_{q+1}^{\alpha-d}, \quad (5.7)$$

we can estimate that our protocol may have substantial weight outside of C_q (and thus have some seeds in each new C_q). The runtime of the overall U_{q+1} is, recursively,

$$t_{q+1} = \tau_{q+1} + 2t_q. \quad (5.8)$$

And indeed,

$$t_q \propto R_q^{\alpha-d} \quad (5.9)$$

does appear consistent with (5.8). This suggests that our recursive approach will be capable of growing an operator at the provably optimal rate in $d = 1$, and conjectured optimal rate in $d > 1$.

In order to make the argument above precise, we will study a random ensemble of protocols, which proceed up to scale $q \sim \ln m \sim \sqrt{\ln r}$. This peculiar choice of m happens to speed up some technical steps in our proof, while only leading to subalgebraic factors in the overall runtime of the method. We will show that *at least half of the protocols in our random ensemble grow operators quickly*, (nearly) saturating Theorem 1. We emphasize that randomness does not play some key physical role; after all, if we just pick some random instance of our protocol, with at least 50% chance that one nonrandom protocol can grow operators quickly! Randomness is useful,

however, as a way to simplify what is already a rather tedious proof which aims to keep track of the probability that a growing operator $O(t)$ has support on a given lattice site. In a nutshell, we introduce random depolarizing unitaries (denoted as D below) which will allow us to keep track of only whether a given site has a Pauli matrix on it or not, but ignore what kind of Pauli we have. This makes our technical proof a bit more conceptually simple.

Despite this randomness, we emphasize that constructive quantum interference is absolutely essential to the success of these random protocols. Indeed, (5.6) holds as $\tau_{q+1} \rightarrow 0$; however, we also need to make sure that (5.6) holds at sufficiently late times such that the unitary V_{q+1} has spread the operator $U_q^\dagger X_0 U_q$ into many sites in C_{q+1} . The fact that V_{q+1} is the product of many $\exp[-iZZ]$ unitaries, which are all mutually commuting, ensures that the long-range interactions are interfering *coherently*.

B. Explicit protocol

Having gone through the intuitive argument, let us now present an explicit random Hamiltonian protocol that achieves this result, at the expense of a logarithmically increased runtime (which we will “justify” at the end of this section). The protocol is built using the same q -cube structure outlined above. In a nutshell, our $(q+1)$ -scale protocol takes the form

$$U_{q+1} = U'_q V_{q+1} U_q, \quad (5.10)$$

where U_q and U'_q are random q -scale unitary protocols drawn from an ensemble that we will state below, and

$$U'_q = U_q D, \quad (5.11)$$

where D is a depolarizing unitary drawn uniformly at random from a discrete ensemble (to be described below) and V_{q+1} is a random unitary, built out of power-law interactions, that will mimic (5.5). At each step, we will choose the unitaries U_q , D , etc., uniformly at random from particular distributions. Perhaps surprisingly, this randomness will actually *help* us prove that the protocol works. The depolarizer D will allow us to effectively ignore what operator is present on a given site, and just keep track of which sites our growing operator has support on. Analogous to recent work on random unitary circuits [11,12], this can make it much easier to keep track of the many-body operator growth.

Ultimately, we will prove that the ensemble-averaged Frobenius norm of a commutator is large; if the average is large, there must be one instance of a unitary time evolution operator in the ensemble (which we do not need to explicitly point out), which achieves a large Frobenius norm. Of course, it is even more interesting that *typical* unitaries in the ensemble achieve a large Frobenius norm.

Let us now carefully define the q -cube partitions of the lattice \mathbb{Z}^d . We define the q cubes

$$C_q(k_1, \dots, k_d) := \{(n_1, \dots, n_d) \in \mathbb{Z}^d : R_q k_i \leq n_i < R_q(k_i + 1)\}. \quad (5.12)$$

The set of all such q cubes will be denoted with

$$\mathcal{B}_q := \{C_q(\mathbf{k}) : \mathbf{k} \in \mathbb{Z}^d\}. \quad (5.13)$$

R_q is defined via (5.2), where

$$m := \lceil e^{\sqrt{\ln(r)}} \rceil. \quad (5.14)$$

The random depolarizing unitary D is chosen as follows:

$$D := \bigotimes_{x \in \Lambda} D_x, \quad (5.15)$$

where D_x are 2×2 unitary matrix acting on qubit x , independent and identically distributed (i.i.d.) for each x . Each D_x is chosen uniformly at random (via the discrete Haar measure) from the group

$$G = \left\{ \pm 1, \frac{\pm 1 \pm iX^a}{\sqrt{2}}, \pm iX^a, \frac{\pm iX^a \pm iX^b}{\sqrt{2}}, \frac{\pm 1 \pm iX \pm iY \pm iZ}{2} \right\}. \quad (5.16)$$

In the above equation, X^a and X^b denote *distinct* Pauli matrices (X, Y, Z). Each D_x can be generated using a single-site Hamiltonian of (operator) norm 1 in a time

$$t_D < \frac{\pi}{2}. \quad (5.17)$$

The growth unitary is

$$V_q = \exp[-i\tau_{q+1} H_q^{ZZ}], \quad (5.18)$$

where the Hamiltonian

$$H_q^{ZZ} := \frac{1}{(dR_q)^\alpha} \sum_{C \in \mathcal{B}_q} \sum_{x,y \in C} J_{xy} Z_x Z_y, \quad (5.19)$$

where J_{xy} are i.i.d. random variables uniformly distributed on the interval $[-1, 1]$. This clearly mimics what we intuitively introduced above; however, we will see that the randomness in the couplings is *beneficial* in allowing us to neglect possible quantum interference phenomena (among growing operators) that might ruin our protocol. The times τ_q will be chosen explicitly in (2.7) in Appendix E, but note for now that it obeys

$$\tau_q < 120^q m^d R_q^{\alpha-d}. \quad (5.20)$$

The protocol will stop at scale

$$q_* := \lceil \sqrt{\ln r} \rceil. \quad (5.21)$$

At this scale, at least half of the operator (as measured by the Frobenius norm) will have size $> r^{d-\epsilon}$ for any ϵ (Proposition 5). In order to demonstrate that the Frobenius light cone is saturated (up to subalgebraic prefactors), we must calculate the total runtime of the protocol U_q . Using the inductive identity (5.10), along with (5.20), we see that if t_q is the total runtime of U_q , and $t_D < \tau_q$ is the runtime of D ,⁹ then

$$\begin{aligned} t_q &= 2t_{q-1} + t_D + \tau_q < 2t_{q-1} + 2\tau_q = \sum_{q=1}^{q_*} \tau_{q_*-q} 2^q < 2^{q_*+1} \tau_{q_*} \\ &< 480(240e)^{\sqrt{\ln r}} \left(1 + e^{\sqrt{\ln r}}\right)^{(\alpha-d)(1+\sqrt{\ln r})} \\ &< 480(240e^{\alpha-d+1} 4^{\alpha-d})^{\sqrt{\ln r}} r^{\alpha-d}. \end{aligned} \quad (5.22)$$

⁹This condition will always be ensured at sufficiently large r . We can shrink the prefactor of H_q^{ZZ} to ensure this at small r

We thus conclude a *lower bound* on the Frobenius light cone in any dimension. To be more precise, given the decomposition of an operator defined in (2.8), let us define the projector

$$\mathbb{P}_{\geq L}|A := \sum_{S \subset \Lambda: \text{diam}(S \cup \{0\}) \geq L} |A_S). \quad (5.23)$$

Here $\text{diam}(S)$ denotes the maximal distance between any two elements in the set S . The discussion above immediately implies the following theorem:

Theorem 3. Let X_0 be the Pauli X matrix supported at the origin $(0, \dots, 0) \in \mathbb{Z}^d$. For any $\epsilon > 0$, define $r(L) := L^{1+\epsilon/2}$. Then there exists a sufficiently large L , such that a power-law Hamiltonian protocol drawn from the distribution (5.10) achieves

$$\|\mathbb{P}_{\geq L}|U_{q_*}^\dagger X_0 U_{q_*}\|_F \geq \frac{1}{2}, \quad (5.24)$$

with probability $\geq \frac{1}{2}$. Moreover, the shape of the Frobenius light cone is bounded by

$$L(t) \geq K_\epsilon t^{\frac{1}{d-d}-\epsilon} \quad (5.25)$$

for some constant $0 < K_\epsilon < \infty$. The asymptotic bound of Theorem 1 cannot be improved by an algebraic factor.

The last thing we need to do is to prove that our protocol in fact does grow a finite fraction of an initial small operator to be large. This result is captured by the following key technical proposition:

Proposition 5. For sufficiently large L , there exists a Hamiltonian in the random ensemble of Sec. V B in which (5.24) holds.

Proof. The proof of this proposition, of course, corresponds to the overwhelming majority of the proof of Theorem 3. As it is rather technically involved, let us outline the key steps in the proof. (1) We will first develop a “superoperator density matrix” perspective for bounding the Frobenius light cone. (2) This notation will prove highly useful since we will show that the “superdepolarizing channel” (conjugation by D) destroys all (unwanted) quantum coherence, and leaves us with an effectively classical problem to analyze. (3) We will then describe the inductive hypothesis required to achieve (5.24), and reduce the quantum mechanical problem of bounding operator growth to the bounds on a classical stochastic process. (4) We will then show that (in the language of the effective stochastic process), with very high probability the V_q unitary “seeds” enough Pauli matrices in new cubes. (5) Analogously, we will show that with very high probability, these seeds in turn grow into large Pauli strings upon applying another U_{q-1} . (6) We will show that at every possible step of the protocol, the collective failure probability is small and decreases sufficiently fast that there is a finite success probability to grow a large operator. Upon converting back to the quantum mechanical language, that will imply (5.24). The technical implementation of this proof is in Appendix E. ■

The fact that *typical* protocols actually are effective at growing operators at the Frobenius light cone suggests a certain robustness of this notion of light cone. At the same time, for chaotic systems, the light cone is linear above $\alpha > d + \frac{1}{2}$ [29,42]. One way to reconcile these facts is to note that the

chaotic Hamiltonians do not choose ZZ Hamiltonians to implement operator growth in V_q . By only including ZZ terms in our long-range Hamiltonian, we ensure that there is always constructive interference in the V_q step at which the operator actually grows.¹⁰ In a Brownian circuit [29,42], in contrast, the operator is rechosen randomly at each step in time: this leads to far greater incoherence which effectively doubles the power law of the interactions from $\alpha \rightarrow 2\alpha$ (as one can only incoherently add squares).¹¹

Remarkably, our protocol *also* saturates both the Frobenius light cone of this paper (proven in $d = 1$ and conjectured for $d > 1$), and a Frobenius light cone proven for random Hamiltonians in $d = 1$ (Theorem 7, [29]). In other words, the Frobenius light cones are essentially the same, whether one fine tunes the Hamiltonian or just draws one randomly from an ensemble. We anticipate these conclusions generalize to higher dimensions, though a formal proof is not known.

VI. OUTLOOK

We have shown constraints on the dynamics of growing operators, measured by the Frobenius norm. In particular, we have proved that in one-dimensional spin chains with long-range interactions, it is possible for the Frobenius light cone to be *exponentially slower* than the Lieb-Robinson light cone ($\alpha = 2$). Such a result is based on the key insight that Frobenius norm becomes approximately submultiplicative, proven by combining standard and new functional analysis tools (uniform smoothness) with the quantum walk formalism.

Moreover, we demonstrated that our Frobenius light cone in one dimension is essentially optimal (up to subalgebraic corrections). Our protocol features the first comprehensive analysis of an explicit random Hamiltonian; in contrast, existing results relied on Brownian Hamiltonian dynamics [42]. Our usage of superdensity operator and superchannel may find further applications in studying operator growth and Frobenius light cones in other systems.

In the near future, we hope to prove our conjecture that the Frobenius light cone in higher-dimensional models with long-range interactions looks (schematically like) $t \gtrsim r^{\alpha-d}$ (for $d < \alpha < d + 1$). Beyond that obvious generalization, we anticipate that our methods will find use in a broad variety of other challenging problems, such as bounding fast scrambling and chaos in quantum simulators, including trapped ion crystals and cavity quantum electrodynamics [26,43,44]; this work may help to constrain when it is possible (or not) to mimic quantum gravity in an experiment [45–50]. A more practical possible application of the Frobenius light cone may be to constrain the generation of (volume-law) entanglement. Lastly, we hope to develop a more general toolkit (perhaps based on the quantum walk methods) to control Frobenius light cones in arbitrary many-body models.

¹⁰However, our V_q is incoherent between different interaction terms due to the random coefficients. This does not slow down the protocol because different Paulis always add up incoherently (i.e., as sum of squares) in the Frobenius norm, regardless of their phases.

¹¹These doubled interactions then obey a standard Lieb-Robinson-type light cone, which is linear when $2\alpha > 2d + 1$, or $\alpha > d + \frac{1}{2}$.

ACKNOWLEDGMENTS

We thank M. Tran and A. Gorshkov for helpful discussions and for collaboration on related work. A.L. was partially supported by a Research Fellowship from the Alfred P. Sloan Foundation, and by the Air Force Office of Scientific Research under Grant No. FA9550-21-1-0195.

APPENDIX A: EXTENSION OF LEMMA 1 TO p -NORMS AND k -LOCAL HAMILTONIANS

In this Appendix, we show that the approximate form of submultiplicativity derived in Sec. III extends to Schatten p -norms with $p > 2$, using the following version of *uniform smoothness*. We include a minimal review, and, e.g., [39,51] for further discussions.¹²

Proposition 6 (Uniform smoothness for subsystem).

Consider matrices X, Y of the same dimensions that satisfy $\text{Tr}_i(Y) = 0; X = X_j \otimes I_i$. For $p \geq 2$,

$$\|X + Y\|_p^2 \leq \|X\|_p^2 + (p-1)\|Y\|_p^2. \quad (\text{A1})$$

The proof of this proposition, which adapts from proof in [[51], Proposition 4.3], is delayed slightly. We first recall the following identity:

Proposition 7 (Uniform smoothness for Schatten classes [[51], Fact 4.1]).

$$\left[\frac{1}{2} (\|X + Y\|_p^p + \|X - Y\|_p^p) \right]^{2/p} \leq \|X\|_p^2 + C_p \|Y\|_p^2 \quad (\text{A2})$$

With these facts in hand, we are now ready to generalize Proposition 3 to the p -norm. We begin with the following simple observation:

Proposition 8. For $\text{Tr}_i(Y) = 0; X = X_{-i} \otimes I_i, p \geq 2$,

$$\|X\|_p \leq \|X + Y\|_p. \quad (\text{A3})$$

Proof. We employ a variational formulation for Schatten p -norms [52]:

$$\|X_j\|_p = \sup_{\|B_j\|_q \leq 1} \text{Tr}(X_j^\dagger B_j) \quad (\text{A4})$$

for $1/p + 1/q = 1$. Then we restrict to B which are proportional to the identity on site i : $B \propto B_{-i} \otimes I_i$:

$$\begin{aligned} \|X + Y\|_p &= \sup_{\|B\|_q \leq 1} \text{Tr}[(X + Y)^\dagger B] \\ &\geq \text{Tr} \left((X^\dagger + Y^\dagger) B_{-i} \otimes \frac{I_i}{\|I_i\|_q} \right) = \|X_{-i} \otimes I\|_p. \end{aligned} \quad (\text{A5})$$

In the last inequality we used that $\text{Tr}_i Y = 0$ and that $X_{-i} \otimes I$ has maximizer $B_{-i} \otimes I_i / \|I_i\|_q$. ■

We can now prove Proposition 6.

Proof. Observe that

$$\frac{\|X + Y\|_p^2 + \|X\|_p^2}{2} \leq \frac{\|X + Y\|_p^2 + \|X - Y\|_p^2}{2} \quad (\text{A6})$$

$$\leq \left(\frac{\|X + Y\|_p^p + \|X - Y\|_p^p}{2} \right)^{2/p} \leq \|X\|_p^2 + C_p \|Y\|_p^2. \quad (\text{A7})$$

The last line uses Lyapunov's inequality, and then Proposition 7. Rearranging terms yields a slightly worse constant $2C_p$. The advertised constant can be obtained via a more involved trick [[51], Lemma A.1]. ■

Now we can show submultiplicativity for arbitrary p -norms. We will also, for good measure, describe how to generalize our result to k -local Hamiltonians as well. As in the main text, we need to properly normalize the Schatten p -norms with the ‘‘bar norm’’ (4.14).

Proposition 9. For k -local Hamiltonian

$$H = \sum_{m \geq i_k > \dots > i_1 \geq 1} H_{i_1, \dots, i_k} \quad (\text{A8})$$

such that $\mathbb{P}_j H_{i_1, \dots, i_k} = \mathbb{I}(j \in \{i_1, \dots, i_k\}) H_{i_1, \dots, i_k}$, for any operator O obeying $\|O\| \leq 1$ and any $p \geq 2$,

$$\|HO\|_{\bar{p}} \leq e p^{k/2} \cdot \|H\|_{(2)} \cdot \|O\|_{\bar{p}} (|\ln \|O\|_{\bar{p}}| + 1)^{k/2}, \quad (\text{A9})$$

where

$$\|H\|_{(2)} := \sqrt{\sum_{m \geq i_k > \dots > i_1 \geq 1} \|H_{i_1, \dots, i_k}\|_{\infty}^2}. \quad (\text{A10})$$

Proof. Using Proposition 1,

$$\|HO\|_p = \|HO\|_{\bar{p}} \leq \|H\|_{p/(1-\theta)} \cdot \|O\|_{p/\theta}. \quad (\text{A11})$$

We now bound each of these norms in turn. We start with $\|H\|_q$ [set $q := p/(1-\theta)$ in what follows]:

$$\begin{aligned} \left\| \sum_{m \geq i_k > \dots > i_1 \geq 1} H_{i_1, \dots, i_k} \right\|_q^2 &\leq \left\| \sum_{m-1 \geq i_k > \dots > i_1 \geq 1} H_{i_1, \dots, i_k} \right\|_q^2 \\ &\quad + C_q \left\| \sum_{m=i_k > i_{k-1} > \dots > i_1 \geq 1} H_{i_1, \dots, i_k} \right\|_q^2 \\ &\leq \sum_{m \geq i_k \geq k} C_q \left\| \sum_{i_k > i_{k-1} > \dots > i_1 \geq 1} H_{i_1, \dots, i_k} \right\|_q^2 \\ &\leq \sum_{m \geq i_k > \dots > i_1 \geq 1} C_q^k \|H_{i_1, \dots, i_k}\|_q^2 \quad (\text{A12}) \\ &\leq C_q^k \|I\|_q^2 \sum_{m \geq i_k > \dots > i_1 \geq 1} \|H_{i_1, \dots, i_k}\|_{\infty}^2 \\ &= C_q^k \|I\|_q^2 \|H\|_{(2)}^2. \quad (\text{A13}) \end{aligned}$$

In the first line we applied Proposition 6 for $Y = \sum_{m=i_k > i_{k-1} > \dots > i_1 \geq 1} H_{i_1, \dots, i_k}$ and $X = \sum_{m-1 \geq i_k > \dots > i_1 \geq 1} H_{i_1, \dots, i_k}$, i.e., we are ‘‘peeling off’’ any term traceless on the qudit $i_k = m$. In the second inequality we peel off qudit $i_k = m-1$,

¹²This particular form of uniform smoothness was perceived when this work and another work [39] was developing. We include the same proof at both papers.

qudit $i_k = m - 2$ all the way to $i_k = k$.¹³ In the second line, we recursively repeat the first and second lines for i_{k-1}, \dots, i_1 , each of which gives a factor of C_q . Lastly, we use Hölder's inequality to bound the q -norm by the ∞ norm with a $\|I\|_p$ overhead.¹⁴

For the norm of O , a standard manipulation using Proposition 2, and $\|O\|_\infty \leq 1$, allows us to handle

$$\|O\|_{p/\theta} \leq \|O\|_p^\theta \|O\|_\infty^{1-\theta} \leq \|O\|_p^\theta. \quad (\text{A14})$$

Now combining (1.2), (1.4), and (1.5), we plug in $q = p/(1 - \theta)$ to get

$$\frac{1}{\|I\|_p} \|HO\|_p \leq \frac{1}{\|I\|_p} \left[\left(\frac{p}{1-\theta} - 1 \right)^{k/2} \|H\|_{(2)} \|I\|_{p/(1-\theta)} \right] \|O\|_p^\theta \leq \left(\frac{p}{1-\theta} \right)^{k/2} \|H\|_{(2)} \left(\frac{\|O\|_p}{\|I\|_p} \right)^\theta, \quad (\text{A15})$$

where we explicitly display the $\|I\|_p$ and used that $\|I\|_p = (\text{Tr}[I])^{1/p}$ to normalize $\|O\|_p$. As in the main text, we choose

$$n = \frac{1}{1-\theta} := \max(1, |\ln(\|O\|_{\bar{p}})|) \leq |\ln(\|O\|_{\bar{p}})| + 1. \quad (\text{A16})$$

Then,

$$\|HO\|_{\bar{p}} \leq p^{k/2} \|H\|_{(2)} \|O\|_{\bar{p}} \times n^{k/2} \|O\|_{\bar{p}}^{-1/n} \leq p^{k/2} \|H\|_{(2)} \|O\|_{\bar{p}} \times n^{k/2} e^{|\ln(\|O\|_{\bar{p}})|/n}. \quad (\text{A17})$$

Combining (1.7) and (1.8), we obtain the advertised result. \blacksquare

APPENDIX B: EXTENSION OF SUBMULTIPLICATIVITY TO TENSOR PRODUCT ENSEMBLES

Submultiplicativity also extends to special other choices of background density matrices ρ . We begin by recalling the following results of [53]:

Proposition 10. Let ρ be a density matrix. Define the norm

$$\|O\|_{q,\rho} := \|\rho^{1/2p} O \rho^{1/2p}\|_p = [\text{Tr}|\rho^{1/2p} O \rho^{1/2p}|^p]^{1/p}. \quad (\text{B1})$$

Then, the following properties hold:

$$\|I\|_{q,\rho} = \text{Tr}(\rho) = 1 \quad (\text{proper normalization}), \quad (\text{B2a})$$

$$\text{Tr}(\sqrt{\rho} A^\dagger \sqrt{\rho} B) \leq \|A\|_{p,\rho} \|B\|_{q,\rho} \quad (\text{Hölder inequality with } 1/p + 1/q = 1), \quad (\text{B2b})$$

$$\|A\|_{q_\theta,\rho} \leq \|A\|_{q_1,\rho}^\theta \|A\|_{q_2,\rho}^{1-\theta} \quad [\text{Riesz-Thorin interpolation with } \theta/q_1 + (1-\theta)/q_2 = 1/q_\theta]. \quad (\text{B2c})$$

Using this proposition, it is straightforward to begin to extend Proposition 3 to use this generalized norm.

Proposition 11. Suppose that $[H, \rho] = 0$. Then,

$$\|HO(t)\|_{2,\rho}^2 \leq \|H\|_{2p,\rho}^2 \|O\|_{2,\rho}^{2/q} \|\rho^{-1/4} O^\dagger \sqrt{\rho} O \rho^{-1/4}\|_\infty^{1-1/q}. \quad (\text{B3})$$

Proof. Using the identities in Proposition 10, and temporarily just writing $O(t) = O$,

$$\begin{aligned} \|HO\|_{2,\rho}^2 &= \text{Tr}(O^\dagger H^\dagger \sqrt{\rho} H O \sqrt{\rho}) = \text{Tr}[(\rho^{-1/4} H^\dagger \sqrt{\rho} H \rho^{-1/4}) \sqrt{\rho} (\rho^{-1/4} O \sqrt{\rho} O^\dagger \rho^{-1/4}) \sqrt{\rho}] \\ &\leq \|\rho^{-1/4} H^\dagger \sqrt{\rho} H \rho^{-1/4}\|_{p,\rho} \cdot \|\rho^{-1/4} O^\dagger \sqrt{\rho} O \rho^{-1/4}\|_{q,\rho} \\ &\leq \|\rho^{-1/4} H^\dagger \sqrt{\rho} H \rho^{-1/4}\|_{p,\rho} \cdot \|\rho^{-1/4} O^\dagger \sqrt{\rho} O \rho^{-1/4}\|_{1,\rho}^{1/q} \cdot \|\rho^{-1/4} O^\dagger \sqrt{\rho} O \rho^{-1/4}\|_{\infty,\rho}^{1-1/q} \\ &\leq \|\rho^{-1/4} H^\dagger \sqrt{\rho} H \rho^{-1/4}\|_{p,\rho} \cdot \text{Tr}(\sqrt{\rho} O^\dagger \sqrt{\rho} O)^{1/q} \cdot \|\rho^{-1/4} O^\dagger \sqrt{\rho} O \rho^{-1/4}\|_{\infty}^{1-1/q}. \end{aligned} \quad (\text{B4})$$

In the third line we used interpolation for $\frac{1}{q} = \frac{1/q}{1} + \frac{1-1/q}{\infty}$ and, in the fourth, we used that at ∞ it does not depend on the background anymore: $\|A\|_{\infty,\rho} = \|A\|_\infty$. Lastly, when $[H, \rho] = 0$, we note that for any a, b, c ,

$$\rho^a O(t)^\dagger \rho^b O(t) \rho^c = \rho^a e^{iHt} O^\dagger e^{-iHt} \rho^b e^{iHt} O e^{-iHt} \rho^c = e^{iHt} \rho^a O^\dagger \rho^b O \rho^c e^{-iHt}, \quad (\text{B5})$$

and moreover we note that

$$\|A(t)\|_{p,\rho} = \|A\|_{p,\rho} \quad (\text{B6})$$

¹³The $i_k = k$ term does not have the C_q coefficient, but we threw it in to simplify the expression.

¹⁴This conversion to operator norm is tight when the spectrum is suitably flat, as is the case when dealing with a Pauli string operator.

for any A . Hence, we obtain (1.3). \blacksquare

We expect that this extension is most useful in the simple scenario where ρ is a simple tensor product. This naturally arises in models with a conserved charge, where one can consider the “infinite-temperature grand canonical ensemble”

$$\rho \propto \bigotimes_i e^{-\mu Q_i}, \quad (\text{B7})$$

where Q_i represents the local charge on site i . Examples correspond to spin systems with conserved total z spin, or models like the Bose-Hubbard model in which total boson number is conserved. In these models, the infinity norm $\|\rho^{-\frac{1}{4}} O^\dagger \sqrt{\rho} O \rho^{-\frac{1}{4}}\|_\infty$ may not be too dangerous. Unfortunately, in a thermal ensemble with $\rho \propto e^{-\beta H}$, due to the nonlocality in H one must be extremely careful about factors of $e^{\beta H}$ when computing operator norms.

APPENDIX C: PROOF OF LEMMA 2

Observe that (for simplicity we will drop the explicit time dependence in \mathcal{L} and H)

$$\begin{aligned} \frac{d}{dt} [A_0(t) | \mathcal{F} | A_0(t)] &= [A_0(t) | [\mathcal{F}, \mathcal{L}] | A_0(t)] = 2 \sum_{R \geq j > i \geq 0} [A_0(t) | \mathbb{Q}_j | \mathcal{F} | \mathcal{L} | \mathbb{Q}_i | A_0(t)] \\ &= 2 \sum_{R \geq j > i \geq 0} (j - i) [A_0(t) | \mathbb{Q}_j | \mathcal{L} | \mathbb{Q}_i | A_0(t)], \end{aligned} \quad (\text{C1})$$

where the factor of 2 comes from the fact that $j > i$ is restricted in the sum, yet the terms with j and i can have either number on either side of the inner product. The antisymmetry of \mathcal{L} and symmetry of \mathcal{F} then allow us to say $(A | [\mathcal{F}, \mathcal{L}] | A) = 2(A | \mathcal{F} \mathcal{L} | A)$. Our technical observation is that we can do a discrete integration by parts mimicking $\mathbb{E}[X] = \int dt \mathbb{Q}(X > t)$ for a real-valued random variable:

$$\sum_{R \geq j > i \geq 0} (j - i) \mathbb{Q}_j | \mathcal{L} | \mathbb{Q}_i = \sum_{R \geq j > i \geq 0} \mathbb{Q}_j | \mathcal{L} \sum_{k=0}^i \mathbb{Q}_k. \quad (\text{C2})$$

In addition to combining these two results, we will want to split up the Liouvillian \mathcal{L} using the same method of [20]. Let us define the following sets:

$$S_{q,k} := \{2^q k, 2^q k + 1, \dots, 2^q(k+2) - 1\}. \quad (\text{C3})$$

We define

$$K_q = \frac{R+1}{2^q} - 1 \quad (\text{C4})$$

to be the largest possible value of k for $S_{q,k}$. As shown in Fig. 2, at each scale q , these sets form “double partitions” of the domain $[0, R]$ at every scale $0 \leq q \leq q_*$. We define

$$H_{q,k} := \sum_{i,j \in S_{q,k}} \mathbb{I}(i, j \text{ are not contained in } S_{q',m} \text{ for any } q' < q, \text{ for any } m) H_{i,j}, \quad (\text{C5})$$

and $\mathcal{L}_{q,k}$ in the obvious way, by denoting $\mathcal{L}_{ij} = i[H_{ij}, \cdot]$ to be the Liouvillian associated with sites i and j , and $\mathcal{L}_{q,k} = i[H_{q,k}, \cdot]$.

Now, with (1.8),

$$\begin{aligned} \frac{d}{dt} [A_0(t) | \mathcal{F} | A_0(t)] &= 2 \sum_{i=0}^{R-1} \sum_{j=i+1}^R \sum_{k=0}^i [A_0(t) | \mathbb{Q}_j | \mathcal{L} | \mathbb{Q}_k | A_0(t)] = \left(\sum_{q=0}^{q_*} \sum_{n=0}^{K_q} \sum_{i \in S_{q,n}} \right) \sum_{j=i+1}^R \sum_{k=0}^i [A_0(t) | \mathbb{Q}_j | \mathcal{L}_{q,n} | \mathbb{Q}_k | A_0(t)] \\ &= \sum_{q=0}^{q_*} \sum_{n=0}^{K_q} \sum_{i \in S_{q,n}} \left(\sum_{j=i+1}^{2^q(n+2)} \right) \sum_{k=2^{2n}}^i [A_0(t) | \mathbb{Q}_j | \mathcal{L}_{q,n} | \mathbb{Q}_k | A_0(t)] \\ &\leq \sum_{q=0}^{q_*} \sum_{n=0}^{K_q} \sum_{i \in S_{q,n}} \left\| \sum_{j=i+1}^{2^q(n+2)} \mathbb{Q}_j | A_0(t) \right\|_{\text{F}} \left\| \mathcal{L}_{q,n} \sum_{k=2^{2n}}^i \mathbb{Q}_k | A_0(t) \right\|_{\text{F}}. \end{aligned} \quad (\text{C6})$$

In the first line we absorb the factor of 2 into the double partition of $0 \leq i \leq R-1$; in the second line, we have used the fact that $\mathcal{L}_{q,n}$ is nonvanishing only on operators with support in $S_{q,n}$, and that if it extends a Pauli string to the right, that Pauli string has rightmost site in $S_{q,n}$. In the third line, we use the Cauchy-Schwarz inequality, keeping in mind that our inner product is already normalized to the Frobenius norm. Now, we wish to further simplify this expression. To proceed, note the following:

Corollary 2. For any operator O obeying $\|O\|_\infty \leq 1$, there exists a constant $0 < C < \infty$ such that

$$\|\mathcal{L}_{q,k} O\|_{\text{F}} \leq C \|O\|_{\text{F}} (\ln \|O\|_{\text{F}} + 1) \times 2^{-q(\alpha-1)}. \quad (\text{C7})$$

Proof. Observe that for each H_{ij} contained in $H_{q,k}$, we have that (if $i < j$) $i < 2^q(k+1)$ and $j \geq 2^q(k+1)$. This means that we may apply Lemma 1:

$$\|\mathcal{L}_{q,k}|O\rangle\|_F \leq 2 \times 9 \times 2e \sqrt{\sum_{i,j} h_{ij}^2} \|O\|_F (|\ln \|O\|_F| + 1). \quad (\text{C8})$$

The first factor of 2 comes from considering either HO or OH , the factor of 9 comes from triangle inequality over two sums of Paulis $\sum_{a,b=1}^3 X^a X^b$. Now observe that

$$\sqrt{\sum_{i,j} h_{ij}^2} \leq \sqrt{\sum_{i=-\infty}^0 \sum_{j=2^{q-1}}^{\infty} \frac{1}{|i-j|^\alpha}} \leq \sqrt{\sum_{i=-\infty}^0 \sum_{j=2^{q-1}}^{\infty} \frac{1}{|i-j|^{2\alpha}}} \leq \text{const} \sqrt{\sum_{j=2^{q-1}}^{\infty} \frac{1}{j^{2\alpha-1}}} \leq \frac{\text{const}}{2^{q(\alpha-1)}}, \quad (\text{C9})$$

where C_0 and C are finite constants that depend on α . Thus, we obtain (1.4). \blacksquare

With Corollary 2,

$$\begin{aligned} \frac{d}{dt} [A_0(t)|\mathcal{F}|A_0(t)] &\leq C \sum_{q=0}^{q_*} \sum_{n=0}^{K_q} \sum_{i \in S_{q,n}} \left\| \sum_{j=i+1}^{2^q(n+2)} \mathbb{Q}_j |A_0(t)\rangle \right\|_F \left\| \sum_{k=2^{qn}}^i \mathbb{Q}_k |A_0(t)\rangle \right\|_F \left(1 - \ln \left\| \sum_{k=2^{qn}}^i \mathbb{Q}_k |A_0(t)\rangle \right\|_F \right) 2^{-q(\alpha-1)} \\ &\leq C \sum_{q=0}^{q_*} \sum_{n=0}^{K_q} \sum_{i \in S_{q,n}} \left\| \sum_{j \in S_{q,n}} \mathbb{Q}_j |A_0(t)\rangle \right\|_F \left\| \sum_{j \in S_{q,n}} \mathbb{Q}_j |A_0(t)\rangle \right\|_F \left(1 - \ln \left\| \sum_{j \in S_{q,n}} \mathbb{Q}_j |A_0(t)\rangle \right\|_F \right) 2^{-q(\alpha-1)} \\ &\leq 2C \sum_{q=0}^{q_*} \sum_{n=0}^{K_q} p_{q,n} \left(1 - \frac{1}{2} \ln p_{q,n} \right) 2^{-q(\alpha-2)}. \end{aligned} \quad (\text{C10})$$

In the second line, we used that the function $x(1 - \ln x)$ is increasing in the domain $x \in [0, 1]$, and that the Frobenius norm of a projected operator monotonically increases if we add orthogonal projectors \mathbb{Q}_i 's to the sum. In the third line, we explicitly carried out the sum over $i \in S_{q,n}$, and defined

$$p_{q,n} := \left\| \sum_{j \in S_{q,n}} \mathbb{Q}_j |A_0(t)\rangle \right\|_F^2. \quad (\text{C11})$$

Namely, $p_{q,n}$ is the probability that the operator has its rightmost support in $S_{q,n}$. Note that

$$\sum_{n=0,2,4,\dots} p_{q,n} \leq 1, \quad \sum_{n=1,3,5,\dots} p_{q,n} \leq 1. \quad (\text{C12})$$

At this point, we first use the fact that for fixed q , the maximal value of the entropy like quantity (1.7) is attained when $p_{q,n}$ is uniform: $p_{q,n} = 2/(K_q + 1)$. Hence,

$$\frac{d}{dt} [A_0(t)|\mathcal{F}|A_0(t)] \leq 2C \sum_{q=0}^{q_*} 2 \left(1 - \frac{1}{2} \ln \frac{2}{K_q + 1} \right) 2^{-q(\alpha-2)} \leq 2C \sum_{q=0}^{q_*} \left(2 + \frac{1 + q_* - q}{2} \ln 2 \right) 2^{-q(\alpha-2)}. \quad (\text{C13})$$

Evaluating this sum leads to (4.9).

APPENDIX D: PROOF OF THEOREM 2

As stated in the main text, we need to modify the technical proof given in [20] in a few places to prove Theorem 2. Let us briefly review the proof strategy of [20]. As in the proof of Theorem 2, we divide up the 1D lattice into boxes of size 2^q for $0 \leq q \lesssim \ln r$. We then observe that if (using the main text notation) $\mathbb{Q}_r |A_0(t)\rangle \neq 0$, then, if we write out

$$A_0(t) = e^{\mathcal{L}t} A_0 = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathcal{L}^n A_0 \quad (\text{D1})$$

as a Taylor series, we can look for the ‘‘irreducible’’ \mathcal{L} 's in any given sequence of n \mathcal{L} 's in \mathcal{L}^n that ‘‘move the operator forward’’: namely, move an operator from the hyperplane of

\mathbb{Q}_m to \mathbb{Q}_{m+k} for some $k > 0$. The key idea of [20], using the formalism of [4], is that (1) reducible (i.e., not irreducible) \mathcal{L} 's can be reexponentiated to form a new unitary transformation, and (2) we only need to count ‘‘irreducible’’ \mathcal{L} 's in sequences that have sufficiently many steps forward at the corresponding scale q (as if we pull down a small number of irreducible steps at scale q' , we can just reexponentiate \mathcal{L} 's at this scale).

We will want to perform each of these two steps a little differently. However, to motivate why, it will help to invert the logic and first think about what we wish to accomplish during the second step. At the longest scale q_* , we only want to count *one* irreducible path. Let \mathcal{L}_{q_*} contain all terms in \mathcal{L} corresponding to a commutator $[H_{m_1}, \dots]$ at the longest scale, $\mathcal{L}_{-q_*} = \mathcal{L} - \mathcal{L}_{q_*}$, and $\mathcal{L}_{q_*,j}$ contains all couplings in \mathcal{L}_{q_*} that

end on site j . Using the Duhamel identity, we write

$$e^{\mathcal{L}t} = e^{\mathcal{L}_{-q_*}t} + \sum_{j>r/2} \int_0^t ds e^{\mathcal{L}(t-s)} \mathcal{L}_{q_*,j} e^{\mathcal{L}_{-q_*}s}. \quad (\text{D2})$$

Now consider using the triangle inequality, which holds for all p -norms, to obtain

$$\begin{aligned} \|\mathbb{Q}_r e^{\mathcal{L}t} |\mathcal{O}_0\rangle\|_{\bar{p}} &\leq \|\mathbb{Q}_r e^{\mathcal{L}_{-q_*}t} |\mathcal{O}_0\rangle\|_{\bar{p}} \\ &+ \sum_{j>r/2} \int_0^t ds \|\mathbb{Q}_r e^{\mathcal{L}(t-s)} \mathcal{L}_{q_*,j} e^{\mathcal{L}_{-q_*}s} |\mathcal{O}_0\rangle\|_{\bar{p}}. \end{aligned} \quad (\text{D3})$$

The first term in this inequality would correspond to all paths that make it from 0 to r without ultralong couplings, and can be bounded separately (as we will describe below). The second term corresponds to all of the paths that invoked a long coupling to get to a site j more than halfway to r . We would be tempted to perform the following simplifications:

$$\begin{aligned} &\|\mathbb{Q}_r e^{\mathcal{L}(t-s)} \mathcal{L}_{q_*,j} e^{\mathcal{L}_{-q_*}s} |\mathcal{O}_0\rangle\|_{\bar{p}} \\ &\leq \|\mathcal{L}_{q_*,j} e^{\mathcal{L}_{-q_*}s} |\mathcal{O}_0\rangle\|_{\bar{p}} \\ &\leq \|H_{q_*,j}\|_{\bar{p}} \|e^{\mathcal{L}_{-q_*}s} |\mathcal{O}_0\rangle\|_{\bar{p}}. \end{aligned} \quad (\text{D4})$$

The first step is very mild, contributing at worst an $O(1)$ factor. The second step, however, can only be used *once* as we have emphasized for the Frobenius light cone (3.7): To peel off $\|H\mathcal{O}\|_{\bar{p}} \sim \|H\|_{(2)}\|\mathcal{O}\|_{\bar{p}}$, submultiplicativity (Proposition 9) requires a constraint $\|\mathcal{O}\|_{\infty} \leq 1$. For (1.4) this criterion is satisfied, but it will not be whenever there are two or more irreducible couplings in a sequence. Our remedy is the following. For illustration, consider growing an operator \mathcal{O} by \mathcal{L} , and for simplicity assume the operator is supported entirely ≤ 0 and \mathcal{L} is bipartite $j > 0 > i$. Then, uniform smoothness implies

$$\|\mathcal{L}|\mathcal{O}\|_{\bar{p}}^2 = \left\| \sum_{j>0>i} \mathcal{L}_{ij} |\mathcal{O}_{\leq 0}\rangle \right\|_{\bar{p}}^2 \quad (\text{D5})$$

$$\leq (p-1) \sum_{j>0} \left\| \sum_{0>i} \mathcal{L}_{ij} |\mathcal{O}_{\leq 0}\rangle \right\|_{\bar{p}}^2 \quad (\text{D6})$$

$$\leq (p-1) \sum_{j>0} \left\| 2 \sum_{0>i} H_{ij} \right\|_{\bar{p}}^2 \|\mathcal{O}_{\leq 0}\|_{\bar{p}}^2. \quad (\text{D7})$$

In words, the sum over the right site j leads to a sum of squares in operator weight (as in Frobenius light cones) while the sum over left site i is bounded additively as in the Lieb-Robinson light cone. What we gain from this sacrifice is that this triggers a recursion for the (normalized) p -norm. We will see that this leads to a light cone of $t \sim r^{\min(1, \alpha-3/2)}$, in-between the Frobenius light cone of $t \sim r^{\min(1, \alpha-1)}$ and Lieb-Robinson light cone $t \sim r^{\min(1, \alpha-2)}$.

Having made this observation, we can now essentially reproduce the proof used in [20], except that we need to modify a few of the constants as well as the equivalence classes used in that proof. Dissect the interactions of scale q and regroup the terms sharing the same site j for $i < j$ (i.e., j is farther from the origin). Labeled by sites $r \geq j \geq 0$, these are the building blocks we will use:

$$\begin{aligned} \sum_k H_{q,k} &= \sum_k \sum_{(i,j) \in H_{q,k}} H_{ij} = \sum_j \left[\sum_k \sum_{i<j, (i,j) \in H_{q,k}} H_{ij} \right] \\ &=: \sum_j H_{q,j}. \end{aligned} \quad (\text{D8})$$

For any sequence, we can define and read out its *forward sequence* by the following recursive algorithm. Suppose $\mathcal{L}_{\beta_n}, \dots, \mathcal{L}_{\beta_1}$ has forward sequence $f_n = (\mathcal{L}_{\beta_n}, \dots, \mathcal{L}_{\beta_1})$. Then for $\mathcal{L}_{\beta_{n+1}}, \dots, \mathcal{L}_{\beta_1}$, it has forward sequence

$$f_{n+1} = \begin{cases} (\beta_{n+1}, f_n), & \text{if } j_{n+1} \geq j_m + 1 \\ f_n, & \text{else} \end{cases} \quad (\text{D9})$$

where $\beta_n \equiv (q_n, k_n, j_n)$. In words, the $j_{n+1} \geq j_m + 1$ condition says a hop is forward if exceeds the farthest so far.

For each forward sequence, we can isolate terms of some scale q to form a q -forward subsequence. Such a sequence is *long* if it has at least

$$N_q := \left\lceil \frac{1}{2} \frac{2^{-q(\alpha-5/2)^{\frac{2}{3}}}}{\sum_{q'=0}^{q_*} 2^{-q'(\alpha-5/2)^{\frac{2}{3}}}} \frac{r}{2^{q+1}} \right\rceil \quad (\text{D10})$$

terms. Note the size of a scale- q block is 2^{q+1} . We then define characteristic function χ_q for each scale

$$\chi_q \mathcal{L}_{\beta_n} \dots \mathcal{L}_{\beta_1} := \begin{cases} \mathcal{L}_{\beta_n}, \dots, \mathcal{L}_{\beta_1}, & \text{if it contains a long } q\text{-forward subsequence} \\ 0, & \text{else} \end{cases} \quad (\text{D11})$$

and for each forward subsequence β

$$\chi_{\beta} \mathcal{L}_{\beta_n} \dots \mathcal{L}_{\beta_1} := \begin{cases} \mathcal{L}_{\beta_n}, \dots, \mathcal{L}_{\beta_1}, & \text{if it contains the forward subsequence } \beta \\ 0, & \text{else.} \end{cases} \quad (\text{D12})$$

We denote with \mathcal{F}_q the set of long q -forward (sub)sequences. The following combinatorial proposition will now prove useful (see, e.g., [54], for its context in combinatorics):

Proposition 12. Let \mathcal{F}_q denote the set of all long forward q sequences of length $\geq N_q$. Then,

$$\chi_q = \sum_{k=N_q}^r \left(\sum_{p=N_q}^k (-1)^{k-p} \binom{k}{p} \right) \sum_{\beta \in \mathcal{F}_q: |\beta|=k} \chi_\beta. \quad (\text{D13})$$

Proof. Suppose that we have a sequence $\mathcal{S} := \mathcal{L}_{\beta_n}, \dots, \mathcal{L}_{\beta_1}$. First note that χ_q clearly annihilates anything which does not have a long q -forward subsequence since $\chi_\beta \mathcal{S} = 0$ for any $\beta \in \mathcal{F}_q$. Hence, we can assume that \mathcal{S} does have a long q -forward subsequence β . Consider what happens at length $N_q + k$. For simplicity, let $\tilde{\chi}_q$ denote the right-hand side of (1.3). In general, we will have

$$\tilde{\chi}_q \mathcal{S} = \mathcal{N} \mathcal{S}, \quad (\text{D14})$$

where the integer \mathcal{N} counts the number of χ_β 's, weighted by their appropriate prefactor in (1.3), which do not annihilate \mathcal{S} . We can straightforwardly calculate

$$\mathcal{N} = \sum_{r=0}^k \binom{N_q + k}{N_q + r} \times \left[\sum_{p=0}^r (-1)^{r-p} \binom{N_q + r}{N_q + p} \right]$$

Lemma 3.

$$\chi_\beta e^{\mathcal{L}t} |A_0\rangle = \int_{\Delta^\ell(t)} dt_\ell \dots dt_1 e^{\mathcal{L}(t-t_\ell)} \mathcal{L}_{\beta_\ell} e^{\mathcal{L}_{\beta_\ell}^\beta(t_\ell-t_{\ell-1})} \mathcal{L}_{\beta_{\ell-1}} e^{\mathcal{L}_{\beta_{\ell-1}}^\beta(t_{\ell-1}-t_{\ell-2})} \dots \mathcal{L}_{\beta_1} e^{\mathcal{L}_{\beta_1}^\beta t_1} |A_0\rangle, \quad (\text{D18})$$

where $\ell = \ell(\beta)$,

$$\mathcal{L}_p^\beta := \mathcal{L} - \sum_{\lambda \in Y_p(\beta)} \mathcal{L}_\lambda \quad (\text{D19})$$

with the sets Y_p^β corresponding to the forbidden terms

$$Y_p^q(\beta) := \{(q', k, j) : j \geq j(\beta_{p-1})\}, \quad (\text{D20})$$

which would change the forwardness of the sequence, at any scale. $j(\beta_p)$ denotes the right site of coupling β_p .

The proof of this lemma is found in [20]. In words, the intermediate unitaries \mathcal{L}_p^β contain the terms (at any scale) which cannot possibly render the coupling \mathcal{L}_{β_p} a nonforward coupling. The following lemma allows us to then bound the p -norm of the resulting sequences:

Lemma 4.

$$\left\| \sum_{\beta \in \mathcal{F}_q, |\beta|=\ell} \chi_\beta e^{\mathcal{L}t} |A_0\rangle \right\|_p \leq \frac{(2\sqrt{C_p r} \|H_{q,i}\| t)^\ell}{\ell! \sqrt{\ell!}} \|A_0\|_p. \quad (\text{D21})$$

Proof. The q -forward sequence can be enumerated by the rightmost sites of the q -scale couplings: we denote these as $r \geq j_\ell > \dots > j_1 \geq 1$, using the previous lemma:

$$\begin{aligned} \sum_{\beta \in \mathcal{F}_q, |\beta|=\ell} \chi_\beta e^{\mathcal{L}t} |A_0\rangle &= \sum_{\beta \in \mathcal{F}_q, |\beta|=\ell} \int_{\Delta^\ell(t)} dt_\ell \dots dt_1 e^{\mathcal{L}(t-t_\ell)} \mathcal{L}_{\beta_\ell} e^{\mathcal{L}_{\beta_\ell}^\beta(t_\ell-t_{\ell-1})} \mathcal{L}_{\beta_{\ell-1}} e^{\mathcal{L}_{\beta_{\ell-1}}^\beta(t_{\ell-1}-t_{\ell-2})} \dots \mathcal{L}_{\beta_1} e^{\mathcal{L}_{\beta_1}^\beta t_1} |A_0\rangle \\ &= \sum_{j_\ell > \dots > j_1} \int_{\Delta^\ell(t)} dt_\ell \dots dt_1 e^{\mathcal{L}(t-t_\ell)} \mathcal{L}_{j_\ell} e^{\mathcal{L}_{(j_\ell)}^\beta(t_\ell-t_{\ell-1})} \mathcal{L}_{j_{\ell-1}} e^{\mathcal{L}_{(j_{\ell-1})}^\beta(t_{\ell-1}-t_{\ell-2})} \dots \mathcal{L}_{j_1} e^{\mathcal{L}_{(j_1)}^\beta t_1} |A_0\rangle \\ &= \int_{\Delta^\ell(t)} dt_\ell \dots dt_1 e^{\mathcal{L}(t-t_\ell)} \sum_{j_\ell} \mathcal{L}_{j_\ell} e^{\mathcal{L}_{(j_\ell)}^\beta(t_\ell-t_{\ell-1})} \sum_{j_{\ell-1} < j_\ell} \mathcal{L}_{j_{\ell-1}} e^{\mathcal{L}_{(j_{\ell-1})}^\beta(t_{\ell-1}-t_{\ell-2})} \dots \sum_{j_1 < j_2} \mathcal{L}_{j_1} e^{\mathcal{L}_{(j_1)}^\beta t_1} |A_0\rangle. \end{aligned} \quad (\text{D22})$$

We use notation $\mathcal{L}_{(i_\ell)}$ to emphasize it is independent of the previous $i_{\ell-1}, \dots, i_1$ (Lemma 3), which then allows us to move the sums inside of the integrals in the third equality. This nested sum then conveniently allows us to use the invariance of p -norms

$$\begin{aligned} &= \sum_{r=0}^k \sum_{p=0}^r (-1)^{r-p} \frac{(N_q + k)!}{(k-r)!(r-p)!(N_q + p)!} \\ &= \sum_{p=0}^k \sum_{q=0}^{k-p} (-1)^q \frac{(k-p)!}{(k-p-q)!q!} \frac{(N_q + k)!}{(k-p)!(N_q + p)!}, \end{aligned} \quad (\text{D15})$$

where in the second line, we switched variables to $r = p + q$ in the sum. Clearly, the sum over q vanishes unless it only runs over $q = 0$, namely, $p = k$. We conclude that only the $k = p$ term above is nonzero, which immediately leads to $\mathcal{N} = 1$. Hence, $\tilde{\chi}_q \mathcal{S} = \mathcal{S}$, which implies $\tilde{\chi}_q = \chi_q$ and thus (1.3). ■

We will use the following simple bound in what follows:

$$\sum_{p=q}^k (-1)^{k-p} \binom{k}{p} \leq 2^k. \quad (\text{D16})$$

Next, we let $\Delta^\ell(t)$ denote the ℓ -simplex:

$$\Delta^\ell(t) := \{(t_1, \dots, t_\ell) \in [0, t]^\ell : t_1 \leq t_2 \leq \dots \leq t_\ell\}. \quad (\text{D17})$$

The following lemma helps us to use the indicator functions above to group all sequences in $e^{\mathcal{L}t}$ into “irreducible” sequences:

under unitary evolution to remove all but the “irreducible” steps \mathcal{L}_{β_p} from our bound:

$$\begin{aligned} & \left\| e^{\mathcal{L}(t-t_\ell)} \sum_{j_\ell} \mathcal{L}_{j_\ell} e^{\mathcal{L}_{(j_\ell)}(t_\ell-t_{\ell-1})} \sum_{j_{\ell-1} < j_\ell} \mathcal{L}_{j_{\ell-1}} e^{\mathcal{L}_{(j_{\ell-1})}(t_{\ell-1}-t_{\ell-2})} \dots \sum_{j_1 < j_2} \mathcal{L}_{j_1} e^{\mathcal{L}_{(j_1)t_1}} |A_0\rangle \right\|_p^2 \\ & \leq C_p \sum_{j_\ell} \|2H_{j_\ell}\|_\infty^2 \left\| e^{\mathcal{L}_{(j_\ell)}(t_\ell-t_{\ell-1})} \sum_{j_{\ell-1} < j_\ell} \mathcal{L}_{j_{\ell-1}} e^{\mathcal{L}_{(j_{\ell-1})}(t_{\ell-1}-t_{\ell-2})} \dots \sum_{j_1 < j_2} \mathcal{L}_{j_1} e^{\mathcal{L}_{(j_1)t_1}} |A_0\rangle \right\|_p^2 \\ & \leq (4C_p)^\ell \sum_{j_\ell > \dots > j_1} \|H_{j_\ell}\|^2 \dots \|H_{j_1}\|^2 \|A_0\|_p^2 \leq \frac{(4C_p r \|H_{q,j}\|^2)^\ell}{\ell!} \|A_0\|_p^2. \end{aligned} \tag{D23}$$

The first inequality uses uniform smoothness for each term $\mathcal{L}_{j_\ell} \dots$, which is nontrivial on site j_ℓ and trivial beyond. Indeed, if we start with the rightmost possible j_ℓ , we know that all other terms in the sum are trivial on site j_ℓ , so uniform smoothness applies; we simply repeat this argument until we have summed over all j_ℓ . The second line repeats this for $i_{\ell-1} \dots i_1$. In the last line, we used the combinatorial bound

$$\sum_{r \geq i_\ell > \dots > i_1 \geq 1} \leq \frac{r^\ell}{\ell!}. \tag{D24}$$

Finally, we use one final triangle inequality to state that

$$\left\| \sum_{\beta \in \mathcal{F}_q, |\beta|=\ell} \chi_\beta e^{\mathcal{L}t} |A_0\rangle \right\|_p \leq \sqrt{\frac{(4C_p r \|H_{q,j}\|^2)^\ell}{\ell!} \|A_0\|_p^2} \int_{\Delta^\ell(t)} dt_\ell \dots dt_1 1 = \frac{(2\sqrt{C_p r} \|H_{q,j}\|)^\ell}{\ell!^{3/2}} \tag{D25}$$

which completes the proof. ■

We now generalize this result to sequences that have q -forward subsequences at multiple scales q , which we index by set $Z \subset (0, \dots, q^*)$. As before, we label with \mathcal{F}_Z the set of all such (sub)sequences (keep in mind that the ordering of terms at different scales is important to keep track of), and we define

$$\chi_Z := \prod_{q \in Z} \chi_q. \tag{D26}$$

Lemma 5.

$$\left\| \sum_{\beta \in \mathcal{F}_Z, |\beta_q|=\ell_q} \chi_\beta e^{\mathcal{L}t} |A_0\rangle \right\|_p \leq \|A_0\|_p \prod_{q \in Z} \frac{(2\sqrt{C_p r} \|H_{q,j}\| t)^{\ell_q}}{\ell_q! \sqrt{\ell_q!}}. \tag{D27}$$

Proof. The proof is nearly identical, but is slightly more tedious due to the multiple scales. Let us first bound the number of ways that different scales can weave through each other, which is bounded by the multinomial coefficient

$$\sum_{\beta \in \mathcal{F}_Z, |\beta_q|=\ell_q} 1 \leq \ell! \prod_{q \in Z} \frac{1}{\ell_q!}, \tag{D28}$$

where

$$\ell := \sum_{q \in Z} \ell_q. \tag{D29}$$

For each sequence in Z , we obtain a factor of

$$\left\| \chi_\beta e^{\mathcal{L}t} |A_0\rangle \right\|_p^2 \leq (4C_p)^\ell \sum_{j_\ell > \dots > j_1} \|H_{j_\ell}\|^2 \dots \|H_{j_1}\|^2 \|A_0\|_p^2 \leq \frac{(4C_p r)^\ell}{\ell!} \|A_0\|_p^2 \prod_{j=1}^\ell \|H_{q,j}\|^2. \tag{D30}$$

Following the last steps of the prior lemma, we find that

$$\left\| \sum_{\beta \in \mathcal{F}_Z, |\beta_q|=\ell_q} \chi_\beta e^{\mathcal{L}t} |A_0\rangle \right\|_p \leq \|A_0\|_p \sum_{\beta \in \mathcal{F}_Z, |\beta_q|=\ell_q} \frac{t^\ell}{\ell!} \leq \frac{1}{\sqrt{\ell!}} \|A_0\|_p \prod_{q \in Z} \frac{(2\sqrt{C_p r} \|H_{q,j}\| t)^{\ell_q}}{\ell_q!} \sqrt{\frac{(4C_p r)^\ell}{\ell!}} \prod_{j=1}^\ell \|H_{q,j}\|. \tag{D31}$$

To finish the proof, we simply use the loose bound

$$\ell! \geq \prod_{q \in Z} \ell_q!, \tag{D32}$$

which leads to (1.7). ■

At this point, we invoke the inclusion-exclusion of different scales found in [20]:

$$\|\mathbb{P}_r e^{\mathcal{L}t} |A_0\rangle\|_p = \left\| \mathbb{P}_r \sum_{Z \neq \emptyset} (-1)^{|Z|-1} \chi_Z e^{\mathcal{L}t} |A_0\rangle \right\|_p \leq \sum_{Z \neq \emptyset} \|\mathbb{P}_r \chi_Z e^{\mathcal{L}t} |A_0\rangle\|_p. \tag{D33}$$

Together with Proposition 12 (which immediately generalizes to multiscale Z), with

$$N_Z := \sum_{q \in Z} N_q, \tag{D34}$$

we find that

$$\begin{aligned} \|\mathbb{P}_r \chi_Z e^{\mathcal{L}t} |A_0\rangle\|_p &= \left\| \sum_{\ell=N_Z}^r \left[\sum_{p=N_Z}^{\ell} (-1)^{\ell-p} \binom{\ell}{p} \right] \sum_{\beta \in \mathcal{F}_Z: |\beta|=\ell} \chi_{\beta} e^{\mathcal{L}t} |A_0\rangle \right\|_p \\ &\leq \sum_{\ell=N_Z}^r 2^{\ell} \sum_{\beta \in \mathcal{F}_Z: |\beta|=\ell} \|\chi_{\beta} e^{\mathcal{L}t} |A_0\rangle\|_p \leq \sum_Z \prod_{q \in Z} \sum_{\ell_q=N_q}^{\infty} \frac{(4\sqrt{C_p r} \|H_{q,j}\| t)^{\ell_q}}{\ell_q!^{3/2}}. \end{aligned} \tag{D35}$$

In the second line we used (1.6), followed by Lemma 5.

Now, let us suppose that

$$1 \geq \frac{4\sqrt{C_p r} \|H_{q,i}\| t}{N_q^{3/2}} \tag{D36}$$

for every scale q . Assuming this inequality (which will fix the values of t for which our bound is valid), then (1.5) becomes

$$\|\mathbb{P}_r e^{\mathcal{L}t} |A_0\rangle\|_p \leq 2 \sum_Z \prod_{q \in Z} \frac{(4\sqrt{C_p r} \|H_{q,j}\| t)^{N_q}}{N_q!^{3/2}} \leq -1 + \exp \left(2 \sum_q \frac{(4\sqrt{C_p r} \|H_{q,j}\| t)^{\ell_q}}{\ell_q!^{3/2}} \right). \tag{D37}$$

It is useful to determine the first value q_1 at which a long q -forward path has a single coupling: $N_q = 1$ for $q \geq q_1$. [Recall N_q 's definition in (D10).] This occurs when

$$\frac{M}{r} \geq \frac{1}{4} \frac{1}{(2q_1)^{(\alpha-1)\frac{2}{3}}}, \tag{D38}$$

where we defined

$$M = \sum_{q=0}^{q_*} 2^{-q(\alpha-5/2)\frac{2}{3}}. \tag{D39}$$

Then, noting that for any j , there exists a constant $0 < c'_1 < \infty$ such that

$$\|H_{q,j}\| \leq \sum_{i=0}^{j-1} \mathbb{I}(H_{ij} \in H_{q,k}) \|H_{ij}\| \leq \frac{c'_1}{4(2q)^{\alpha-1}}, \tag{D40}$$

we find that there exist constants $0 < c'_2, c_2 < \infty$ such that

$$\begin{aligned} \sum_q 2 \frac{(4\sqrt{C_p r} \|H_{q,i}\| t)^{N_q}}{N_q!^{3/2}} &\leq \sum_q \left(\frac{c'_2 \sqrt{p r t}}{2^{q(\alpha-1)} N_q^{3/2}} \right)^{N_q} \\ &\leq \sum_{q=0}^{q_1-1} \left(\frac{c_2 \sqrt{p t} M^{3/2}}{r} \right)^{N_q} + c_3 \sqrt{p r t} \sum_{q=q_1}^{q_*} \frac{1}{2^{q(\alpha-1)}}. \end{aligned} \tag{D41}$$

We now analyze this sum for different ranges of α .

Case: $\alpha > \frac{5}{2}$. First, we note that M does not depend on r since even as $r \rightarrow \infty$ the sum in (1.8) is convergent. Second, we observe that $N_1 > N_2 > \dots > N_{q_*}$, which follows from (D10) and the fact that N_q (before the floor function) changes by a factor of at least $2^{\frac{2}{3}(\alpha-1)} > 2$ each time. These two inequalities imply that there exist constants $0 < c_{4,5,6} < \infty$ which do not depend

on r such that

$$\sum_{q=0}^{q_1-1} \left(\frac{c_2 \sqrt{pt} M^{3/2}}{r} \right)^{N_q} + c_3 \sqrt{prt} \sum_{q=q_1}^{q_*} \frac{1}{2^{q(\alpha-1)}} \leq \sum_{n=2}^{\infty} \left(\frac{c_2 \sqrt{pt} M^{3/2}}{r} \right)^n + \frac{c_3 \sqrt{prt}}{2^{q_*(\alpha-1)}} \leq c_4 \frac{\sqrt{pt}}{r - c_5 \sqrt{pt}} + \frac{c_6 \sqrt{pt}}{r}. \tag{D42}$$

Observe that (1.6) holds so long as

$$t \leq \frac{N_q^{3/2}}{4\sqrt{pr}} 2^{q(\alpha-1)} \leq k'_1 \frac{r^{3/2}}{2^{q(\alpha-1)} \sqrt{pr} (4M)^{3/2}} 2^{q(\alpha-1)} \leq \frac{k_1 r}{\sqrt{p}}, \tag{D43}$$

for constants $0 < k'_1, k_1 < \infty$ that do not depend on r .

Case: $\frac{3}{2} \leq \alpha < \frac{5}{2}$. Now we find that (for constants $0 < c_7 < \infty$ and an integer m independent of r)

$$M = \sum_{q=0}^{q_*} 2^{-q(\alpha-5/2)\frac{2}{3}} = c_7 r^{(5/2-\alpha)\frac{2}{3}} \tag{D44}$$

and

$$q_1 = q_* - m. \tag{D45}$$

Then, note that $N_1 > N_4 > N_7 \dots$ because the argument of (D10) now only varies by $2^{(\alpha-1)2/3} \geq 2^{1/3}$ each time q varies by 1. Hence, we obtain (for constants $0 < c_{8,9,10} < \infty$ independent of r)

$$\sum_{q=0}^{q_1-1} \left(\frac{c_2 \sqrt{pt} M^{3/2}}{r} \right)^{N_q} + c_3 \sqrt{prt} \sum_{q=q_1}^{q_*} \frac{1}{2^{q(\alpha-1)}} \leq 3 \sum_{n=2}^{\infty} \left(\frac{c_2 \sqrt{pt} M^{3/2}}{r} \right)^n + c_3 \sqrt{prt} \sum_{q=q_1}^{q_*} \frac{1}{2^{q(\alpha-1)}} \leq \frac{c_8 \sqrt{pt}}{r^{\alpha-3/2} - c_9 \sqrt{pt}} + \frac{c_{10} \sqrt{pt}}{r^{\alpha-3/2}}. \tag{D46}$$

Observe that (1.6) holds so long as

$$t \leq \frac{N_q^{3/2}}{4\sqrt{pr}} 2^{q(\alpha-3/2)} \leq \frac{r^{3/2} 2^{q(5/2-\alpha)} 2^{-3q/2}}{\sqrt{pr} M^{3/2}} 2^{q(\alpha-1)} = k_2 \frac{r}{\sqrt{p}} r^{5/2-\alpha} = k_2 \frac{r^{\alpha-3/2}}{\sqrt{p}} \tag{D47}$$

for $0 < k_2 < \infty$.

Case : $\alpha = \frac{5}{2}$. We obtain

$$M = q_* + 1. \tag{D48}$$

For r -independent constant $0 < c_{11} < \infty$,

$$2^{q_1} = c_{11} \frac{r}{\ln_2 r}. \tag{D49}$$

Again we have $N_1 > N_2 > \dots > N_{q_1-1} > 1$, for the same reason as when $\alpha > \frac{5}{2}$. Hence for $0 < c_{12,13,14} < \infty$,

$$\sum_{q=0}^{q_1-1} \left(\frac{c_2 \sqrt{pt} M^{3/2}}{r} \right)^{N_q} + c_3 \sqrt{prt} \sum_{q=q_1}^{q_*} \frac{1}{2^{q(\alpha-1)}} \leq \sum_{n=2}^{\infty} \left(\frac{c_2 \sqrt{pt} \ln(r)^{3/2}}{r} \right)^n + c_3 \sqrt{prt} \sum_{q=q_1}^{q_*} \frac{1}{2^{q(\alpha-1)}} \tag{D50}$$

$$\leq \frac{c_{12} \sqrt{pt}}{r \ln^{-3/2} r - c_{13} \sqrt{pt}} + \frac{c_{14} \sqrt{pt} \ln r}{r}. \tag{D51}$$

Lastly, observe that (1.6) is satisfied so long as

$$t \leq \frac{N_q^{3/2}}{4\sqrt{pr}} 2^q \leq k_3 \frac{r^{3/2}}{\ln^{3/2} r} \frac{2^{q-q}}{\sqrt{pr}} = \frac{k_3}{\sqrt{p}} \frac{r}{\ln^{3/2} r} \tag{D52}$$

∞ , for times $t < k_4 \mathcal{R}(r)$,

$$\frac{\|\mathbb{P}_r e^{\mathcal{L}t} |A_0\rangle\|_p}{\|A_0\|_p} \leq \frac{k_5 t}{\mathcal{R}(r)}. \tag{D53}$$

for some constant $0 < k_3 < \infty$.

Now let us summarize our results. Recall the definition of $\mathcal{R}(r)$ in (4.17). We have shown that for constants $0 < k_{4,5} < \infty$ This completes the proof.

APPENDIX E: PROOF OF PROPOSITION 5

Step 1. We begin by developing the “superdensity matrix” picture of operator growth, following [55]. If \mathcal{H} is the quantum mechanical Hilbert space, and (with a slight abuse of notation) $\mathcal{H} \otimes \mathcal{H}$ is a Hilbert space of all normalizable operators on \mathcal{H} (in what follows, we will restrict to Hermitian operators), the superdensity matrix is a normalizable element of $(\mathcal{H} \otimes \mathcal{H}) \otimes (\mathcal{H} \otimes \mathcal{H})$. It is easier to visualize with bra-ket notation: using the operator ket $|A\rangle$ introduced above, the pure superdensity matrix corresponding to A becomes $|A\rangle\langle A|$. It is straightforward to build a good basis for superoperators. Using the fact that the space of Hermitian operators acting on a single qubit is spanned by the orthonormal basis

$$|X^a\rangle \quad (a = 0, 1, 2, 3) := |I\rangle, |X\rangle, |Y\rangle, |Z\rangle \in \mathbb{R}^{1+3} \quad (\text{E1})$$

endowed with the canonical inner product in \mathbb{R}^{1+3} , a basis for superoperators on a single qubit is evidently $|X^a\rangle\langle X^b|$ for $a, b = 0, 1, 2, 3$. The standard tensor product between lattice sites then allows us to build up a good basis for our superoperator space.

Observe that the Frobenius norm is now simply (up to square) the superoperator trace

$$\text{Tr}[|A\rangle\langle A|] := \langle A|A\rangle = \|A\|_{\mathbb{F}}^2. \quad (\text{E2})$$

The latter equality follows from (2.5). As a consequence, evaluating the Frobenius norms of projected operators in this superoperator language will be particularly simple: we simply pick out the basis operators which we wish to keep, and sum the coefficients of the diagonal elements of the pure superdensity matrix. In particular, given the growing operator $X_0(t)$, we can define the probabilities $p_S(t)$ (for subsets $S \subset \Lambda$) as [recall (2.9)]

$$\begin{aligned} |X_0(t)\rangle\langle X_0(t)| &:= \sum_{S \subset \Lambda} p_S(t) |X_0(t)_S\rangle\langle X_0(t)_S| \\ &+ \text{off-diagonal terms.} \end{aligned} \quad (\text{E3})$$

Much of the proof that follows will amount to bounding (sums of) $p_S(t)$ throughout the protocol. To avoid confusing the superoperator spaces from operators, we here define the adjoint operation acting on operators $\mathcal{H} \otimes \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{H}$,

$$\text{Adj}_i(X) := |i\rangle\langle X| \quad (\text{E4})$$

and the conjugation operation

$$\text{Conj}_B|X\rangle := |B^\dagger X B\rangle. \quad (\text{E5})$$

We define the superdepolarizing channel, acting on superoperators $(\mathcal{H} \otimes \mathcal{H}) \otimes (\mathcal{H} \otimes \mathcal{H}) \rightarrow (\mathcal{H} \otimes \mathcal{H}) \otimes (\mathcal{H} \otimes \mathcal{H})$ as

$$\mathcal{D} := \mathbb{E}_D[(\text{Conj}_D)(\cdot)(\text{Conj}_D)^\dagger]. \quad (\text{E6})$$

Here and below, the (\cdot) in an expression is defined such that $\mathcal{V}_q \mathcal{O}$ is given by the above expression with (\cdot) replaced by the object \mathcal{O} on which the linear transformation \mathcal{V}_q acts.

Recall the depolarizer D is a random tensor product unitary defined in the main text. To give a concrete example,

$$\mathcal{D}[|X\rangle\langle Y|] = \mathbb{E}_D |D^\dagger X D\rangle\langle D^\dagger Y D|. \quad (\text{E7})$$

We define the q -scale growth channel as

$$\mathcal{V}_q := \mathbb{E}_J[(\text{Conj}_{\mathcal{V}_q})(\cdot)(\text{Conj}_{\mathcal{V}_q})^\dagger]. \quad (\text{E8})$$

Expectation values are taken over the random D_x in (5.15), and random J_{xy} in (5.19), respectively. We thus see that our (averaged) protocol as a “quantum superchannel,” defined via the analog of (5.10):

$$\mathcal{M}_q := \mathcal{M}_{q-1} \mathcal{D} \mathcal{V}_q \mathcal{M}_{q-1}. \quad (\text{E9})$$

Step 2. We now turn to the analysis of the superdepolarizing channel \mathcal{D} . Let us define the “maximally mixed nontrivial operator” (on a single site)

$$\mu := \frac{|X\rangle\langle X| + |Y\rangle\langle Y| + |Z\rangle\langle Z|}{3}. \quad (\text{E10})$$

In a nutshell, the essence of this proof is that all we need to keep track of is whether or not each site has a μ or the trivial operator

$$\mathcal{I} = |I\rangle\langle I| \quad (\text{E11})$$

on it: namely, we will show that after each step of the random protocol,

$$\mathcal{D} \mathcal{V}_q |X_0\rangle\langle X_0| = \sum_{S \subset \Lambda} \rho_S \bigotimes_{i \in S} \mu_i \bigotimes_{j \in S^c} I_j. \quad (\text{E12})$$

The depolarizing channel is what makes this simplification possible, as formalized via the following:

Proposition 13. For any $a, b = 1, 2, 3$ (and on every site independently),

$$\mathcal{D}|I\rangle\langle I| = |I\rangle\langle I|, \quad (\text{E13a})$$

$$\mathcal{D}|I\rangle\langle X^a| = \mathcal{D}|X^a\rangle\langle I| = 0, \quad (\text{E13b})$$

$$\mathcal{D}|W^a\rangle\langle W^b| = \delta^{ab} \mu. \quad (\text{E13c})$$

Proof. The key idea behind this proof is to use the group-theoretic structure of the random unitaries D_x [for simplicity in what follows, we drop the x subscript, since \mathcal{D} is simply a tensor product of the channel defined via (E.13) on every site anyway]. By definition in our protocol,

$$\mathcal{D} := \frac{1}{48} \sum_{D \in G} [(\text{Conj}_D)(\cdot)(\text{Conj}_D)^\dagger], \quad (\text{E14})$$

as 48 is the number of group elements in G , defined in (5.16). By construction, \mathcal{D} takes any 4×4 matrix M acting on the space of operators on a qubit, defined in (1.4), and projects it onto G -invariant maps.

To find these irreducible representations, we first observe that G is isomorphic to the “double cover” of the group of rotations which leave invariant a three-dimensional cube, which is in turn isomorphic to $S_4 \times \mathbb{Z}_2$ [a subgroup of $SU(2)$]. By standard representation-theoretic computation, we find that \mathbb{R}^{1+3} decomposes into two irreducible representations of G :

$$\text{span}(I, X, Y, Z) = \text{span}(I) \oplus \text{span}(X, Y, Z). \quad (\text{E15})$$

By Schur’s Lemma, the only G -invariant matrices obeying $\mathcal{D}M = M$ are therefore of the form

$$M = a|I\rangle\langle I| + b|X\rangle\langle X| + b|Y\rangle\langle Y| + b|Z\rangle\langle Z| = a\mathcal{I} + 3b\mu. \quad (\text{E16})$$

We then obtain (E.13) by using the fact that \mathcal{D} is a probability-weighted linear sum of unitary operators, and

therefore is completely positive and trace preserving. This condition fixes $a = 1$ and $b = \frac{1}{3}$. ■

Combining this proposition with (1.2), we immediately find the following:

Corollary 3. For any superoperator ρ ,

$$\mathcal{D}\rho = \sum_{S \subset \Lambda} \rho_S \bigotimes_{i \in S} \mu_i \bigotimes_{j \in S^c} I_j. \quad (\text{E17})$$

In our random protocol, (1.5) holds, even in intermediate protocol steps, after any application of \mathcal{D} .

Step 3. Corollary 3 shows that the only information we need to keep track of, after depolarizing, is the probability p_S that our growing operator is supported on the subset S . We define the number superoperators

$$\hat{N} = \sum_{x \in \Lambda} \hat{N}_x, \quad (\text{E18a})$$

$$\hat{N}_x := (|X\rangle\langle X| + |Y\rangle\langle Y| + |Z\rangle\langle Z|)_x = 3\mu_x. \quad (\text{E18b})$$

However, due to Corollary 3, we can actually think of each of these as *classical random variables*. In particular, after applying (averaged) superchannel $\mathcal{D}\mathcal{U}$ (for any \mathcal{U}), the probability distribution $\hat{\mathbb{P}}$ on the variables $\hat{N}_x \in \{0, 1\}$ can be defined (with a slight abuse of notation) via

$$\begin{aligned} \hat{\mathbb{P}}[\hat{N}_{x_1} = \hat{N}_{x_2} = \dots = \hat{N}_{x_m} = 1, \hat{N}_{y_1} = \dots = \hat{N}_{y_n} = 0] \\ := \text{Tr}[\hat{N}_{x_1} \dots \hat{N}_{x_m} (1 - \hat{N}_{y_1}) \dots (1 - \hat{N}_{y_n}) \\ \times \mathcal{D}\mathcal{U}|X_0\rangle\langle X_0|]. \end{aligned} \quad (\text{E19})$$

In words, we consider the basis which all \hat{N} are simultaneously diagonalized. The classical probability distribution $\hat{\mathbb{P}}$ counts the weight of $\mathcal{D}\mathcal{U}|X_0\rangle\langle X_0|$ that satisfies the Boolean quantifiers represented by the product of \hat{N} 's and $(1 - \hat{N})$'s above, namely, the sum of all diagonal elements of $\mathcal{D}\mathcal{U}|X_0\rangle\langle X_0|$ compatible with $\hat{N}_{x_1} = 1$, etc. The well posedness of $\hat{\mathbb{P}}$ is guaranteed by the fact that $\mathcal{D}\mathcal{U}$ is completely positive and trace preserving, and the observables N are commuting. Knowing the classical probabilities p_S (corresponding to the probability that $\hat{N}_x = 1$ if and only if $x \in S$) is enough to know $\mathcal{D}\mathcal{U}|X_0\rangle\langle X_0|$.

Armed with this knowledge, we are ready to lay out the foundations for the remainder of our inductive proof. The induction hypothesis we begin with is that for any set $S \subset C_0 \in \mathcal{B}_{q-1}$, with C_0 the $(q-1)$ -scale cube containing the origin

$$\hat{\mathbb{P}}\left(\hat{N} \mathcal{M}_{q-1} \bigotimes_{i \in S} \mu_i \geq \lambda_{1,q-1} R_{q-1}^d\right) \geq \eta_{1,q-1}. \quad (\text{E20})$$

Here $0 < \lambda_{1,q-1}, \eta_{1,q-1} < 1$ are constants that we will obtain a little later; in particular though, $\eta_{1,q-1}$ will be interpreted as a lower bound on the *success probability* of \mathcal{M}_{q-1} , namely,

the probability that it grows an operator to have support on $S \subset C_0$ with S containing fraction $\geq \lambda_{1,q-1}$ of all sites in C_0 .

Next, we condition on the assumption that at least $s := \lambda_{1,q-1} R_{q-1}^d$ sites are occupied. This throws away a fraction of at most $1 - \eta_{1,q-1}$ of the operator (an amount that we will see is small). We will then show that if we run the unitary V_q for sufficiently long time, we will seed more than 3.3% of the $(q-1)$ cubes in the q -cube containing the origin, with probability $\eta_{2,q-1}$ (see Lemma 6). We note that $1 - \eta_{2,q-1}$ decays exponentially with m^d and s .

Lastly, we apply \mathcal{M}_{q-1} , which will attempt to grow the operators in $(q-1)$ cubes $C_{q-1}(\mathbf{k})$ that we seeded above into large Pauli strings. Conditioned on there are at least n_s seeded blocks C_{q-1} , we will show in Proposition 15 that with probability $\eta_{3,q-1}$, a rather large number of sites (to be quantified later) will be occupied [when considering all $(q-1)$ cubes together]. This proves the inductive hypothesis. We summarize the way that operators grow throughout this process in Fig. 4.

Step 4. We now analyze the V_q operator growth step of our protocol in detail. The key observation is that, upon averaging over all possible V_q , the superchannel \mathcal{V}_q takes an arbitrary operator of sufficient size in any $(q-1)$ cube $C_{q-1} \subset C_q$, and “seeds” new μ 's in a finite fraction of the remaining $(q-1)$ cubes $C'_{q-1} \subset C_q$ which are contained within the q cube C_q (see Fig. 4). More precisely, we have the following:

Lemma 6. Let $C_q \in \mathcal{B}_q$. For any subset $S \subseteq C_q$ of size $|S| \geq s$, set the time

$$\tau_q := \frac{R_q^\alpha}{\sqrt{2sR_{q-1}^d}}. \quad (\text{E21})$$

Define the superoperator

$$\rho' = \mathcal{V}_q(\mu^{\otimes S} \mathcal{I}^{\otimes (C_q - S)}). \quad (\text{E22})$$

Let $C \in \mathcal{B}_{q-1}$ denote one of the $(q-1)$ cubes contained within C_q , and define the classical random variables

$$\hat{y}_C := \mathbb{I}\left(\sum_{x \in C} \hat{N}_x > 0\right). \quad (\text{E23})$$

Then,

$$\begin{aligned} 1 - \eta_{2,q-1} &:= \hat{\mathbb{P}}_{\rho'} \left[\sum_{C=1}^{m^d} \hat{y}_C \leq \frac{m^d}{30} \right] \\ &\leq \exp\left(-\frac{m^d}{120}\right) + \exp\left(-\frac{s}{72}\right), \end{aligned} \quad (\text{E24})$$

where $\hat{\mathbb{P}}_{\rho'}$ denotes the probability distribution obtained from (2.8). For notational simplicity, we have labeled the m^d $(q-1)$ cubes in C_q as $C = 1, \dots, m^d$.

Proof. First, we rewrite \mathcal{V}_q in a more elegant way, using the fact that all ZZ terms in H_q^{ZZ} commute with each other:

$$\begin{aligned} \mathcal{V}_q &= \mathbb{E}_J \left\{ \exp \left[\frac{\tau_q}{R_q^\alpha} \sum_{\{i,j\} \subset C_q} J_{ij} \text{Adj}_{iZ_j Z_j} \right] \right\} (\cdot) \left\{ \exp \left[\frac{\tau_q}{R_q^\alpha} \sum_{\{i,j\} \subset C_q} J_{ij} \text{Adj}_{jZ_i Z_i} \right] \right\}^\dagger \\ &= \mathbb{E}_J \prod_{\{i,j\} \subset C_q} \left\{ \exp \left[\frac{\tau_q}{R_q^\alpha} J_{ij} \text{Adj}_{iZ_j Z_j} \right] \right\} (\cdot) \left\{ \exp \left[\frac{\tau_q}{R_q^\alpha} J_{ij} \text{Adj}_{jZ_i Z_i} \right] \right\}^\dagger. \end{aligned} \quad (\text{E25})$$

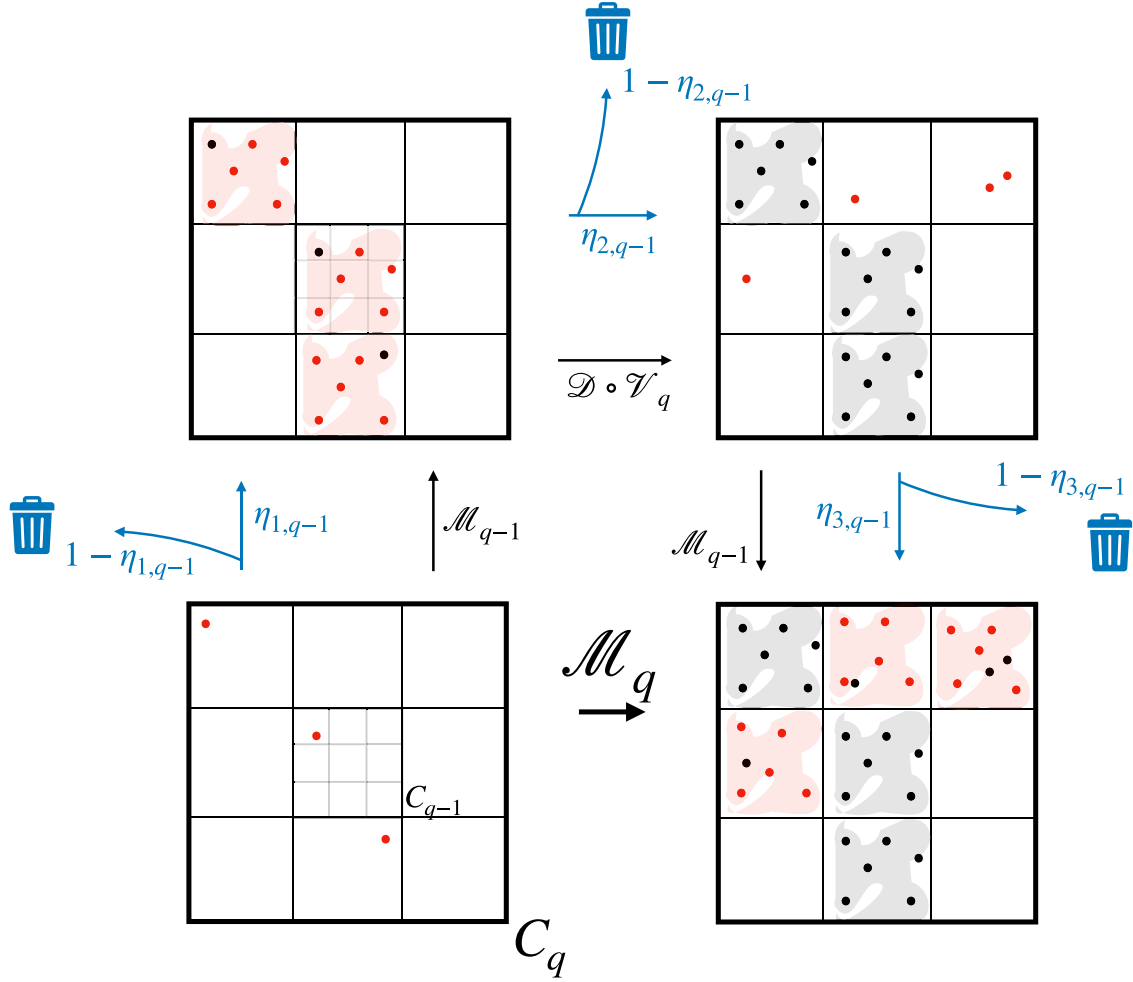


FIG. 4. The rigorous operator growth protocol, in which we no longer guarantee full occupancy of all sites. Instead, we only get to keep a fraction of occupancy at each recursion, albeit with high probability. Using the central limit theorem, the failure probabilities $1 - \eta$ are suppressed by rate of expansion $e^{-O(m)}$. The choice of $m \sim \exp(\sqrt{\ln r})$ is large enough to render these probabilities mild, and leave us with (5.24).

Since Z terms commute with Z 's, but not with X 's or Y 's, let us define

$$v = \frac{|X\rangle\langle X| + |Y\rangle\langle Y|}{2}, \quad \mathcal{Z} = |Z\rangle\langle Z| \quad (\text{E26})$$

such that

$$\mu = \frac{2}{3}v + \frac{1}{3}\mathcal{Z}. \quad (\text{E27})$$

From (2.1), we can first calculate what a single $Z_i Z_j$ coupling does to μ_i or $\mu_i \mu_j$; the application of \mathcal{V}_q will then correspond to repeating this procedure on all pairs of sites in the cube C_q . It is straightforward to show that (temporarily denoting $\theta_{ij} := J_{ij} \tau_q R_q^{-\alpha}$)

$$\exp[\theta \text{Adj}_{iZ_j}] \mathcal{I}_i \mathcal{I}_j \exp[\theta \text{Adj}_{iZ_j}]^\dagger = \mathcal{I}_i \mathcal{I}_j, \quad (\text{E28a})$$

$$\exp[\theta \text{Adj}_{iZ_j}] v_i v_j \exp[\theta \text{Adj}_{iZ_j}]^\dagger = v_i v_j. \quad (\text{E28b})$$

Moreover, since

$$\begin{aligned} & \exp[\theta \text{Adj}_{iZ_j}] (|X_i I_j\rangle\langle\langle Y_i Z_j|) \\ &= \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ -\sin(2\theta) & \cos(2\theta) \end{pmatrix} (|X_i I_j\rangle\langle\langle Y_i Z_j|), \end{aligned} \quad (\text{E29})$$

we conclude that

$$\begin{aligned} & \mathbb{E}[\exp[\theta \text{Adj}_{iZ_j}] (|X_i I_j\rangle\langle\langle X_i I_j| | Y_i Z_j\rangle\langle\langle Y_i Z_j|) \exp[\theta \text{Adj}_{iZ_j}]^\dagger] \\ &= \begin{pmatrix} 1 - p_1 & p_1 \\ p_1 & 1 - p_1 \end{pmatrix} (|X_i I_j\rangle\langle\langle X_i I_j| | Y_i Z_j\rangle\langle\langle Y_i Z_j|), \end{aligned} \quad (\text{E30})$$

where

$$\begin{aligned} p_1 &= \mathbb{E}[\cos^2(2\theta_{ij})] = \frac{1}{2} - \frac{1}{2} \mathbb{E}[e^{4iJ_{ij}\tau_q/R_q^\alpha}] \\ &= \frac{1}{2} \left[1 - F\left(\frac{4\tau_q}{R_q^\alpha}\right) \right], \end{aligned} \quad (\text{E31})$$

where

$$F(x) := \frac{\sin x}{x}. \tag{E32}$$

The following Taylor expansion estimate for $F(x)$ will prove useful:

$$1 - \frac{x^2}{6} \leq F(x) \leq 1 - \frac{x^2}{6} + \frac{x^4}{120}. \tag{E33}$$

Combining (2.7) and (E33), we obtain

$$\begin{aligned} & \mathbb{E} \left[\exp[\theta \text{Adj}_{i\mathcal{Z}, \mathcal{Z}_j}] \begin{pmatrix} v_i \mathcal{I}_j \\ v_i \mathcal{Z}_j \end{pmatrix} \exp[\theta \text{Adj}_{i\mathcal{Z}, \mathcal{Z}_j}]^\dagger \right] \\ &= \begin{pmatrix} 1 - p_1 & p_1 \\ p_1 & 1 - p_1 \end{pmatrix} \begin{pmatrix} v_i \mathcal{I}_j \\ v_i \mathcal{Z}_j \end{pmatrix}, \end{aligned} \tag{E34}$$

where, upon plugging in for θ_{ij} , we obtain

$$\frac{4}{3} \frac{\tau_q^2}{R_q^{2\alpha}} > p_1 > \frac{4}{3} \frac{\tau_q^2}{R_q^{2\alpha}} - \frac{16}{15} \frac{\tau_q^4}{R_q^{4\alpha}}. \tag{E35}$$

An important consequence of all of these identities is that if we write out an initial operator density in the form

$$\hat{\rho} = \prod_{i \in S_v} v_i \prod_{j \in S_z} \mathcal{Z}_j \prod_{k \in C_q - S} \mathcal{I}_k, \tag{E36}$$

where S_v and S_z form a partition of S , then this tensor product structure is preserved under the average of the time evolution. This is because v_i 's are invariant, while \mathcal{Z} 's and \mathcal{I} 's convert between each other in a Markovian fashion independently on each site. Using the independence of the J_{ij} 's, we see that if $\ell = |S_v|$, then

$$\begin{aligned} \mathcal{V}_q \hat{\rho} &= \prod_{i \in S_v} v_i \prod_{j \in S_z} [(1 - p_\ell) \mathcal{Z}_j + p_\ell \mathcal{I}_j] \\ &\times \prod_{k \in C_q - S} [(1 - p_\ell) \mathcal{I}_k + p_\ell \mathcal{Z}_k], \end{aligned} \tag{E37}$$

where we define p_ℓ via

$$\begin{pmatrix} 1 - p_1 & p_1 \\ p_1 & 1 - p_1 \end{pmatrix}^\ell = \begin{pmatrix} 1 - p_\ell & p_\ell \\ p_\ell & 1 - p_\ell \end{pmatrix}. \tag{E38}$$

More concretely, we may bound p_ℓ as follows:

$$\begin{aligned} p_\ell &= \sum_{j=1,3,5,\dots}^\ell \binom{\ell}{j} p_1^j (1 - p_1)^{\ell-j} \\ &= \frac{1}{2} [(1 - p_1 + p_1)^\ell - (1 - p_1 - p_1)^\ell] \geq \frac{1 - e^{-2p_1\ell}}{2}. \end{aligned} \tag{E39}$$

Now, of the m^d cubes $C_{q-1}(\mathbf{k}) \subset C_q$, we must count how many of them contain either a v or a \mathcal{Z} on at least one site, after applying $\mathbb{E}[\mathcal{V}_q] \hat{\rho}$. Consider Bernoulli random variables $\hat{x}_j \in \{0, 1\}$ on sites $j \in C_q$, with

$$\hat{\mathbb{P}}[\hat{x}_j = 1] := \begin{cases} 1, & j \in S_v \\ 1 - p_\ell, & j \in S_z \\ p_\ell, & \text{otherwise.} \end{cases} \tag{E40}$$

The interpretation of \hat{x}_j is the probability that we find a non-identity operator on site j in $\mathbb{E}[\mathcal{V}_q] \hat{\rho}$. Since $p_\ell \leq \frac{1}{2}$, we can

easily see that for any $(q - 1)$ cube $C \subset C_q$,

$$\hat{\mathbb{P}} \left[\sum_{k \in C} \hat{x}_k = 0 \right] = \hat{\mathbb{P}}[\hat{y}_C = 0] \leq (1 - p_\ell)^{|Q|} := 1 - p_*, \tag{E41}$$

where Q is any subset of C_q , which will be taken as $C_{q-1}(\mathbf{k}) \subset C_q$. This inequality follows from the independence of \hat{x}_k , regardless of the set Q . We can now bound the probability that at least $\lambda_{2,q-1} m_q^d$ of the cubes C_{q-1} have at least one operator in them using the standard Chernoff bounds:

Proposition 14 (Chernoff bounds). Let A be a discrete set. For $i \in A$, let $x_i \in \{0, 1\}$ be independent Bernoulli random variables. If

$$S := \sum_{i \in A} x_i, \tag{E42}$$

then

$$\mathbb{P}[S \leq (1 - \delta)\mathbb{E}[S]] \leq \exp\left(-\frac{1}{2}\delta^2\mathbb{E}[S]\right). \tag{E43}$$

In our calculation, the Bernoulli random variables of interest are \hat{y}_C , which obey $\mathbb{E}[\hat{y}_C] \geq p_*$. Choosing

$$\lambda_{2,q-1} := \frac{p_*}{2}, \tag{E44}$$

we arrive at

$$\hat{\mathbb{P}} \left[\sum_C \hat{y}_C \leq \frac{p_* m^d}{2} \right] \leq \exp\left(-\frac{p_* m^d}{8}\right). \tag{E45}$$

To conclude the proof, note that the above calculation was based on the number of v 's. By our initial assumption, the initial operator will have $\geq s$ μ 's, which can be anywhere in cube C_q . Breaking up μ into v and \mathcal{Z} (2.3), we see that our initial operator is binomial distributed: on each of (at least) s sites, the probability of v is $\frac{2}{3}$. Using Chernoff bounds on an initial operator of s μ , the probability of having $\ell \geq s/2$ is extremely high when s is large:

$$\hat{\mathbb{P}}\left(\ell \geq \frac{s}{2}\right) \geq 1 - \exp\left(-\frac{s}{72}\right). \tag{E46}$$

This inequality holds regardless of the number of sites in the set S on which our operator is initially supported, so long as $|S| \geq s$.

Now, we proceed as follows. Using the simple fact that in classical probability theory, for any two events A and B , $\mathbb{P}(A) \geq \mathbb{P}(A \text{ and } B)$, we will lower bound the probability $\eta_{2,q-1}$ by calculating

$$\begin{aligned} \hat{\mathbb{P}}_{\rho'} \left[\sum_{C=1}^{m^d} \hat{y}_C \leq \frac{m^d}{30} \right] &\geq \hat{\mathbb{P}}_{\rho'} \left[\sum_{C=1}^{m^d} \hat{y}_C \leq \frac{m^d}{30} \text{ and } \ell \geq \frac{s}{2} \right] \\ &\geq \hat{\mathbb{P}}_{\rho'} \left[\sum_{C=1}^{m^d} \hat{y}_C \leq \frac{m^d}{30} \mid \ell \geq \frac{s}{2} \right] \\ &\times \hat{\mathbb{P}}_{\rho'} \left[\ell \geq \frac{s}{2} \right]. \end{aligned} \tag{E47}$$

To bound the conditional probability above, we start with (2.6) and (2.8):

$$\begin{aligned} p_* &= 1 - (1 - p_\ell)^{R_{q-1}^d} \geq 1 - \exp(-p_\ell R_{q-1}^d) \\ &\geq 1 - \exp\left(-\frac{1 - e^{-p_1 s}}{2} R_{q-1}^d\right). \end{aligned} \quad (\text{E48})$$

Note that we have used that $\ell \geq s/2$. Now, observe that for $x < 1$, $e^{-x} \leq 1 - \frac{x}{2}$. Since

$$p_1 s \leq \frac{4s}{3} \frac{\tau_q^2}{R_q^{2\alpha}} \leq \frac{2}{3R_{q-1}^d} \leq 1, \quad (\text{E49})$$

we may further simplify (2.5) to

$$p_* \geq 1 - \exp\left(-\frac{p_1 s R_{q-1}^d}{4}\right) \geq \frac{p_1 s R_{q-1}^d}{6}, \quad (\text{E50})$$

noting that (2.6) implies the argument of exp above is $\leq \frac{1}{6}$, in which case $1 - e^{-x} \geq \frac{2}{3}x$. Now using (2.7) and (2.2), we find

$$\begin{aligned} p_* &\geq \frac{s R_{q-1}^d}{6} \left(\frac{2}{3s R_{q-1}^d} - \frac{4}{15s^2 R_q^{2(d-1)}} \right) \\ &\geq \frac{s R_{q-1}^d}{6} \frac{2}{3s R_{q-1}^d} \left(1 - \frac{2}{5s R_{q-1}^d} \right) \geq \frac{1}{15}. \end{aligned} \quad (\text{E51})$$

Combining (2.2) and (2.8), we find that

$$\hat{\mathbb{P}}_{\rho'} \left[\sum_{C=1}^{m^d} \hat{y}_C \leq \frac{m^d}{30} \mid \ell \geq \frac{s}{2} \right] \geq 1 - \exp\left(-\frac{m^d}{120}\right). \quad (\text{E52})$$

Combining (2.3), (2.4), and (2.9), we obtain (E24). ■

Step 5. Now that we know at least $\frac{1}{30}$ th of the $(q-1)$ cubes in a q -cube are seeded after the first application of \mathcal{V}_q , we must now ask what happens after applying \mathcal{M}_{q-1} again (recall Fig. 4). The answer is provided by the following proposition:

Proposition 15. Consider a tensor product superoperator of the form

$$\rho = \mu^{\otimes S} \mathcal{I}^{\otimes S^c}, \quad (\text{E53})$$

where $S \subset \Lambda$ is finite. Then,

$$\begin{aligned} 1 - \eta_{3,q-1} &:= \hat{\mathbb{P}}_{\rho} \left[\hat{N} \leq \frac{w \eta_{1,q-1} \lambda_{1,q-1} R_{q-1}^d}{2} \mid \sum_{C \in \mathcal{B}_{q-1}} \hat{y}_C = w \right] \\ &\leq \exp\left(-\frac{w \eta_{1,q-1}}{8}\right). \end{aligned} \quad (\text{E54})$$

Proof. By our inductive hypothesis (2.6), for any $C \in \mathcal{B}_{q-1}$,

$$\hat{\mathbb{P}} \left[\sum_{x \in C} \hat{N}_x \mathcal{M}_{q-1} \rho \geq \lambda_{1,q-1} R_{q-1}^d \mid \text{Tr}(\hat{y}_C \rho) = 1 \right] \geq \eta_{1,q-1}. \quad (\text{E55})$$

So, letting \hat{x}_C denote a Bernoulli random variable denoting whether or not the criterion above is satisfied, we observe that $\mathbb{E}[\hat{x}_C] \geq \eta_{1,q-1}$. Moreover, \hat{x}_C form independent Bernoulli random variables for each cube C . So again, we may use the

Chernoff bounds to show that

$$\hat{\mathbb{P}} \left[\sum_{C \in \mathcal{B}_{q-1}} \hat{x}_C \leq \frac{w \eta_{1,q-1}}{2} \right] \leq \exp\left(-\frac{w \eta_{1,q-1}}{8}\right). \quad (\text{E56})$$

Since we are guaranteed that the event in (2.1) does *not* occur if the event in (2.3) does *not* occur (since for each $\hat{x}_C \geq 1$, we get a contribution of at least $\lambda_{1,q-1} R_{q-1}^d$ to \hat{N}), (2.3) implies (2.1). ■

Step 6. We now must combine the results from the previous two steps to prove that we may choose $\eta_{1,q}$ such that (2.6) continues to hold at scale q . As in Fig. 4, we can conclude from our discussion above that if we choose

$$\lambda_{1,q} := \frac{\eta_{1,q-1} \lambda_{1,q-1}}{60}, \quad (\text{E57})$$

then the success probability at scale q is given by

$$\begin{aligned} \eta_{1,q} &:= \hat{\mathbb{P}} \left[\text{Tr} \left(\hat{N} \mathcal{M} \left[\bigotimes_{i \in S \subset C_q} \mu_i \right] \right) \geq \lambda_{1,q} R_q^d \right] \\ &\geq \eta_{1,q-1} \eta_{2,q-1} \eta_{3,q-1}, \end{aligned} \quad (\text{E58})$$

where $\eta_{1,q-1}$ is the probability that the first \mathcal{M}_{q-1} is successful [namely, the induction event in (2.6) has occurred], $\eta_{2,q-1}$ is the probability that the first $\mathcal{D}\mathcal{V}_q$ is successful (given in Lemma 6), and $\eta_{3,q-1}$ is the probability that the second \mathcal{M}_{q-1} is successful (given in Proposition 15).

To start off this recursive relation, we first discuss what happens at scale $q=1$. This scale is somewhat special since the operator starts off with probability 1 being a Pauli X at the origin. Because it is a Pauli X , the bound in Lemma 6 can be simplified (*only* for this very first application of \mathcal{V}_1 , before any depolarizing channel!). It is in fact simplest to convert $|X_0\rangle\langle X_0|$, the initial superoperator, into ν_0 (this simplifies some equations, but does not change the protocol's performance). Using (2.4), we see that

$$\mathcal{V}_1 |X_0\rangle\langle X_0| = \nu_0 \prod_{x \in C_1 - 0} [(1 - p_1) \mathcal{I}_x + p_1 \mathcal{Z}_x], \quad (\text{E59})$$

where C_1 denotes the 1-cube containing the origin, and, using (2.2),

$$p_1 > \frac{2}{3} - \frac{4}{15} = \frac{6}{15}. \quad (\text{E60})$$

Using the Chernoff bounds, we can easily see that

$$1 - \eta_{2,0} = \hat{\mathbb{P}} \left[\text{Tr}(\hat{N} \mathcal{M}_1 \nu_0) \leq \frac{m^d}{30} \right] \leq \exp\left(-\frac{m^d}{120}\right). \quad (\text{E61})$$

Note also that $\eta_{1,0} = \eta_{3,0} = 1$ since there is no \mathcal{M}_0 , so

$$\eta_{1,1} = \eta_{2,0} \geq 1 - \exp\left(-\frac{m^d}{120}\right). \quad (\text{E62})$$

This corresponds to the base case of our inductive proof.

Given this base case, the following lemma demonstrates that $\eta_{1,q}$ is indeed (for all $q \leq q_*$) finite.

Lemma 7. For large enough r , the recursion relation

$$\eta_{1,q} = \eta_{1,q-1}\eta_{2,q-1}\eta_{3,q-1} \tag{E63}$$

admits a solution obeying, for all $1 \leq q \leq q_*$,

$$\eta_{1,q} \geq \frac{1}{2}. \tag{E64}$$

Proof. The key idea of this proof is that m is (perhaps surprisingly) sufficiently large so that $\eta_{2,k}$ and $\eta_{3,k}$ are so close to 1 that the repeated multiplication of η probabilities above converges to a nonzero result (for $q \leq q_*$). To see this concretely, let us use Lemma 6 and Proposition 15 to note that

$$\alpha_{q-1} := \eta_{2,q-1}\eta_{3,q-1} \geq \left[1 - \exp\left(-\frac{\lambda_{1,q-1}R_{q-1}^d}{72}\right) - \exp\left(-\frac{m^d}{120}\right) \right] \left[1 - \exp\left(-\frac{m^d\eta_{1,q-1}}{240}\right) \right]. \tag{E65}$$

We wish to analyze the nonlinear recursion relation

$$\eta_{1,q} = \eta_{1,q-1}\alpha_{q-1}. \tag{E66}$$

Happily, to demonstrate the lemma, we can assume that (2.1) holds when evaluating α_{q-1} . The reason for this is that α_{q-1} monotonically increases with $\eta_{1,q-1}$; hence, setting $\eta_{1,q-1} = \frac{1}{2}$ when evaluating α_{q-1} gives us a lower bound on α_{q-1} , and in turn on $\eta_{1,q}$. From (2.4) and (2.8), we can bound

$$\lambda_{1,q} \geq \lambda_{1,1}120^{1-q} = \frac{4}{120^q}. \tag{E67}$$

We thus find that for $q > 1$,

$$\begin{aligned} \alpha_{q-1} &\geq 1 - \exp\left(-\frac{m^d}{480}\right) - \exp\left(-\frac{m^d}{120}\right) - \exp\left[-\frac{1}{18}\left(\frac{m^d}{120}\right)^{q-1}\right] \\ &\geq 1 - \exp\left(-\frac{m^d}{480}\right) - \exp\left(-\frac{m^d}{120}\right) - \exp\left(-\frac{m^d}{2160}\right). \end{aligned} \tag{E68}$$

We conclude that for all $q \leq q_*$,

$$\eta_{1,q} \geq \left[1 - \exp\left(-\frac{m^d}{480}\right) - \exp\left(-\frac{m^d}{120}\right) - \exp\left(-\frac{m^d}{2160}\right) \right]^{q_*}. \tag{E69}$$

Let us now show that (2.6) is compatible with (2.1). To do this, we simply recall the definitions of $m \sim \exp(\sqrt{\ln r})$ in (5.14), and $q_* \sim \sqrt{\ln r}$ in (5.21). In fact, since for sufficiently (and not very) large r , $\sqrt{\ln r} > \ln(\ln r)$,

$$\exp(-m^d) \leq \exp(-e^{d\sqrt{\ln r}}) \leq \exp(-e^{d \ln(\ln r)}) \leq \exp[-(\ln r)^d] \leq \frac{1}{r}. \tag{E70}$$

Thus, we observe that for sufficiently large m (and hence r),

$$\begin{aligned} \left[1 - \exp\left(-\frac{m^d}{480}\right) - \exp\left(-\frac{m^d}{120}\right) - \exp\left(-\frac{m^d}{2160}\right) \right]^{q_*} &> 1 - 3q_* \exp\left(-\frac{m^d}{2160}\right) \\ &> 1 - \frac{3(1 + \sqrt{\ln r})}{r^{1/2160}}. \end{aligned} \tag{E71}$$

For sufficiently large r , this is larger than $\frac{1}{2}$. This ensures (2.1). ■

The final step in the proof of Proposition 5 is very simple. We have shown that, on average, we can take a single Pauli X_0 supported on one site, and have half of the operator weight supported on a fraction of sites

$$\hat{N} \geq \lambda_{1,q_*}R_{q_*}^d \geq \left(\frac{m^d}{120}\right)^{q_*} \geq \left(\frac{e^{d\sqrt{\ln r}}}{120}\right)^{\sqrt{\ln r}} \geq \frac{r^d}{120^{\sqrt{\ln r}}}. \tag{E72}$$

Setting $r = L^{1+\epsilon/2}$, we find

$$\hat{N} \geq L^d \frac{L^{d\epsilon/2}}{120^{\sqrt{(1+\epsilon/2)\ln L}}} > L^d \tag{E73}$$

for sufficiently large r and L . Combining (2.1) with (E73), we obtain that (5.24) holds on average in our ensemble of random unitaries:

$$\mathbb{E}[\|\mathbb{P}_{\geq L}|X_0(t)\rangle\|_F] \geq \frac{1}{2}. \tag{E74}$$

Since $\|\mathbb{P}_{\geq L}|X_0(t)\rangle\|_F \leq 1$, with at least 50% chance a unitary drawn from the ensemble has grown an operator to be large:

$$\mathbb{P}\left[\|\mathbb{P}_{\geq L}|X_0(t)\rangle\|_F \geq \frac{1}{2}\right] \geq \frac{1}{2}. \tag{E75}$$

- [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 and T. Koma, Spectral gap and exponential decay of correlations, *Commun. Math. Phys.* **265**, 781 (2006).
- [3] 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).
- [4] C.-F. Chen and A. Lucas, Operator growth bounds from graph theory, *Commun. Math. Phys.* **385**, 1273 (2021).
- [5] G. Bentsen, Y. Gu, and A. Lucas, Fast scrambling on sparse graphs, *Proc. Natl. Acad. Sci. USA* **116**, 6689 (2019).
- [6] Z. Wang and K. R.A. Hazzard, Tightening the lieb-robinson bound in locally interacting systems, *PRX Quantum* **1**, 010303 (2020).
- [7] M. B. Hastings, Solving gapped hamiltonians locally, *Phys. Rev. B* **73**, 085115 (2006).
- [8] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low, Quantum algorithm for simulating real time evolution of lattice hamiltonians, [arXiv:1801.03922](https://arxiv.org/abs/1801.03922).
- [9] S. Bravyi, M. B. Hastings, and S. Michalakis, Topological quantum order: Stability under local perturbations, *J. Math. Phys.* **51**, 093512 (2010).
- [10] 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).
- [11] A. Nahum, S. Vijay, and J. Haah, Operator Spreading in Random Unitary Circuits, *Phys. Rev. X* **8**, 021014 (2018).
- [12] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws, *Phys. Rev. X* **8**, 021013 (2018).
- [13] J. W. Britton, B. C. Sawyer, A. C. Keith, C.-C. Joseph Wang, J. K. Freericks, H. Uys, M. J. Biercuk, and J. J. Bollinger, Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins, *Nature (London)* **484**, 489 (2012).
- [14] M. Saffman, T. G. Walker, and K. Mølmer, Quantum information with Rydberg atoms, *Rev. Mod. Phys.* **82**, 2313 (2010).
- [15] Bo Yan, S. A. Moses, B. Gadway, J. P. Covey, K. R. A. Hazzard, A. M. Rey, D. S. Jin, and J. Ye, Observation of dipolar spin-exchange interactions with lattice-confined polar molecules, *Nature (London)* **501**, 521 (2013).
- [16] Z.-X. Gong, M. Foss-Feig, S. Michalakis, and A. V. Gorshkov, Persistence of Locality in Systems With Power-Law Interactions, *Phys. Rev. Lett.* **113**, 030602 (2014).
- [17] 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).
- [18] D. V. Else, F. Machado, C. Nayak, and N. Y. Yao, Improved lieb-robinson bound for many-body hamiltonians with power-law interactions, *Phys. Rev. A* **101**, 022333 (2020).
- [19] M. C. Tran, A. Y. Guo, Y. Su, J. R. Garrison, Z. Eldredge, M. Foss-Feig, A. M. Childs, and A. V. Gorshkov, Locality and Digital Quantum Simulation of Power-Law Interactions, *Phys. Rev. X* **9**, 031006 (2019).
- [20] C.-F. Chen and A. Lucas, Finite Speed of Quantum Scrambling with Long Range Interactions, *Phys. Rev. Lett.* **123**, 250605 (2019).
- [21] T. Kuwahara and K. Saito, Strictly Linear Light Cones in Long-Range Interacting Systems of Arbitrary Dimensions, *Phys. Rev. X* **10**, 031010 (2020).
- [22] 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).
- [23] 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).
- [24] T. Kuwahara and K. Saito, Absence of Fast Scrambling in Thermodynamically Stable Long-Range Interacting Systems, *Phys. Rev. Lett.* **126**, 030604 (2021).
- [25] A. Lucas, Non-perturbative dynamics of the operator size distribution in the Sachdev-Ye-Kitaev model, *J. Math. Phys.* **61**, 081901 (2020).
- [26] C. Yin and A. Lucas, Bound on quantum scrambling with all-to-all interactions, *Phys. Rev. A* **102**, 022402 (2020).
- [27] C. Yin and A. Lucas, Quantum operator growth bounds for kicked tops and semiclassical spin chains, *Phys. Rev. A* **103**, 042414 (2021).
- [28] A. W. Harrow, L. Kong, Zi-Wen Liu, S. Mehraban, and P. W. Shor, Separation of out-of-time-ordered correlation and entanglement, *PRX Quantum* **2**, 020339 (2021).
- [29] C.-F. Chen, Concentration of otoc and lieb-robinson velocity in random Hamiltonians, [arXiv:2103.09186](https://arxiv.org/abs/2103.09186).
- [30] A. Lucas and A. Osborne, Operator growth bounds in a cartoon matrix model, *J. Math. Phys.* **61**, 122301 (2020).
- [31] X. Chen, Y. Gu, and A. Lucas, Many-body quantum dynamics slows down at low density, *SciPost Phys.* **9**, 071 (2020).
- [32] A. Lucas, Operator Size at Finite Temperature and Planckian Bounds on Quantum Dynamics, *Phys. Rev. Lett.* **122**, 216601 (2019).
- [33] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, *J. High Energy Phys.* **08** (2016) 106.
- [34] M. Blake, Universal Charge Diffusion and the Butterfly Effect in Holographic Theories, *Phys. Rev. Lett.* **117**, 091601 (2016).
- [35] D. A. Roberts and B. Swingle, Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories, *Phys. Rev. Lett.* **117**, 091602 (2016).
- [36] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, Matrix product operator representations, *New J. Phys.* **12**, 025012 (2010).
- [37] A. Montanaro and T. J. Osborne, Quantum Boolean functions, [arXiv:0810.2435](https://arxiv.org/abs/0810.2435).
- [38] A. Montanaro, Some applications of hypercontractive inequalities in quantum information theory, *J. Math. Phys.* **53**, 122206 (2012).
- [39] C.-F. Chen, M. Berta, and F. Brandao (unpublished).
- [40] C. Yin and A. Lucas, Finite speed of quantum information in models of interacting bosons at finite density, [arXiv:2106.09726](https://arxiv.org/abs/2106.09726).
- [41] Y. Hong and A. Lucas, Fast high-fidelity multiqubit state transfer with long-range interactions, *Phys. Rev. A* **103**, 042425 (2021).
- [42] T. Zhou, S. Xu, X. Chen, A. Guo, and B. Swingle, Operator lévy Flight: Light Cones in Chaotic Long-Range Interacting Systems, *Phys. Rev. Lett.* **124**, 180601 (2020).
- [43] Z. Li, S. Choudhury, and W. V. Liu, Fast scrambling without appealing to holographic duality, *Phys. Rev. Research* **2**, 043399 (2020).

- [44] R. Belyansky, P. Bienias, Y. A. Kharkov, A. V. Gorshkov, and B. Swingle, Minimal Model for Fast Scrambling, *Phys. Rev. Lett.* **125**, 130601 (2020).
- [45] A. Chew, A. Essin, and J. Alicea, Approximating the Sachdev-Ye-Kitaev model with Majorana wires, *Phys. Rev. B* **96**, 121119(R) (2017).
- [46] A. Chen, R. Ilan, F. de Juan, D. I. Pikulin, and M. Franz, Quantum Holography in a Graphene Flake with an Irregular Boundary, *Phys. Rev. Lett.* **121**, 036403 (2018).
- [47] J. Marino and A. M. Rey, Cavity-QED simulator of slow and fast scrambling, *Phys. Rev. A* **99**, 051803(R) (2019).
- [48] R. J. Lewis-Swan, A. Safavi-Naini, J. J. Bollinger, and A. M. Rey, Unifying scrambling, thermalization and entanglement through measurement of fidelity out-of-time-order correlators in the Dicke model, *Nat. Commun.* **10**, 1581 (2019).
- [49] Y. Alavirad and A. Lavasani, Scrambling in the Dicke model, *Phys. Rev. A* **99**, 043602 (2019).
- [50] G. Bentsen, T. Hashizume, A. S. Buyskikh, E. J. Davis, A. J. Daley, S. S. Gubser, and M. Schleier-Smith, Treelike Interactions and Fast Scrambling with Cold Atoms, *Phys. Rev. Lett.* **123**, 130601 (2019).
- [51] De Huang, J. Niles-Weed, J. A. Tropp, and R. Ward, Matrix concentration for products, [arXiv:2003.05437](https://arxiv.org/abs/2003.05437).
- [52] M. M. Wilde, Preface to the second edition, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2017), pp. xi–xii.
- [53] S. Beigi, Sandwiched rényi divergence satisfies data processing inequality, *J. Math. Phys.* **54**, 122202 (2013).
- [54] M. Aigner, *A Course in Enumeration* (Springer, Berlin, 2010).
- [55] A. Lucas, Quantum many-body dynamics on the star graph, [arXiv:1903.01468](https://arxiv.org/abs/1903.01468).