Locality and Digital Quantum Simulation of Power-Law Interactions

Minh C. Tran,[1,2,3,*] Andrew Y. Guo,[1,2] Yuan Su,[1,4,5] James R. Garrison,[1,2] Zachary Eldredge,[1,2] Michael Foss-Feig,[6,1,2] Andrew M. Childs,[1,4,5] and Alexey V. Gorshkov[1,2]

[1]*Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, Maryland 20742, USA*
[2]*Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*
[3]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA*
[4]*Department of Computer Science, University of Maryland, College Park, Maryland 20742, USA*
[5]*Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742, USA*
[6]*United States Army Research Laboratory, Adelphi, Maryland 20783, USA*

(Received 25 September 2018; revised manuscript received 21 February 2019; published 10 July 2019)

The propagation of information in nonrelativistic quantum systems obeys a speed limit known as a Lieb-Robinson bound. We derive a new Lieb-Robinson bound for systems with interactions that decay with distance r as a power law, $1/r^\alpha$. The bound implies an effective light cone tighter than all previous bounds. Our approach is based on a technique for approximating the time evolution of a system, which was first introduced as part of a quantum simulation algorithm by Haah *et al.*, FOCS'18. To bound the error of the approximation, we use a known Lieb-Robinson bound that is weaker than the bound we establish. This result brings the analysis full circle, suggesting a deep connection between Lieb-Robinson bounds and digital quantum simulation. In addition to the new Lieb-Robinson bound, our analysis also gives an error bound for the Haah *et al.* quantum simulation algorithm when used to simulate power-law decaying interactions. In particular, we show that the gate count of the algorithm scales with the system size better than existing algorithms when $\alpha > 3D$ (where D is the number of dimensions).

DOI: 10.1103/PhysRevX.9.031006

Subject Areas: Atomic and Molecular Physics, Condensed Matter Physics, Quantum Information

I. INTRODUCTION

Lieb-Robinson bounds limit the rate at which information can propagate in systems that obey the laws of nonrelativistic quantum mechanics [1–10]. These bounds have found a plethora of applications [11–22], including recent results on entanglement area laws [23–25], the classical complexity of sampling bosons [26], and even a quantum algorithm for digital quantum simulation [27].

Lieb and Robinson's original proof applies only to short-range interactions, i.e., those that act over a finite range or decay at least exponentially in space. However, interactions in many physical systems, such as trapped ions [28,29], Rydberg atoms [30], ultracold atoms and molecules [31,32], nitrogen-vacancy centers [33], and superconducting circuits [34], can decay with distance r as a power law ($1/r^\alpha$) and, hence, lie outside the scope of the original Lieb-Robinson bound. Thus, understanding the fundamental limit on the speed of information propagation in these systems holds serious physical implications, including for the applications mentioned above. Despite many efforts in recent years [4–7], a tight Lieb-Robinson bound for such long-range interactions remains elusive.

In this paper, we derive a new Lieb-Robinson bound for systems with power-law decaying interactions in D dimensions. While our bound is not known to be tight, it has four main benefits compared to the best previous bound for such systems [6]. (i) It is tighter, resulting in the best effective light cone to date [Eq. (17)]. (ii) The bound applies at all times, and not just asymptotically in the large-time limit. (iii) The framework behind the proof is conceptually simpler, with an easy-to-understand interpretation based on physical intuition. (iv) Our approach is potentially applicable to studying a wider variety of quantities, including connected correlators [35,36] and higher-order correlators (for instance, the out-of-time-ordered correlator [37,38] and the full measurement statistics of boson sampling [26,39]), as we discuss in Sec. VI.

In contrast to the previous long-range Lieb-Robinson bounds [4–7], which all relied on the so-called Hastings-Koma series [4], our approach is based on a generalization of the framework Haah *et al.* (HHKL) [27] introduced as a

*Corresponding author.
minhtran@umd.edu

building block for their quantum simulation algorithm. The essence of their framework is a technique for decomposing the time evolution of a system into evolutions of subsystems, with an error bounded by the Lieb-Robinson bound for short-range interactions [1]. We extend the HHKL framework to long-range interactions and to a more general choice of subsystems. Remarkably, these modifications enable us to derive a tighter Lieb-Robinson bound for long-range interactions than the one we use in the analysis of the decomposition [5].

Additionally, we return to the original motivation of the framework of Haah *et al.*: the digital simulation of lattice-based quantum systems. We generalize the HHKL algorithm to simulate systems with power-law decaying interactions. The algorithm scales better as a function of system size than previous algorithms when $\alpha > 3D$, and the speed-up becomes more dramatic as α is increased.

The structure of the paper is as follows. In Sec. II, we state our main results and summarize the proof of the new Lieb-Robinson bound. In Sec. III, we lay out the precise mathematical framework for the proof and generalize the technique for decomposing time-evolution unitaries [27] to power-law decaying interactions and to more general choices of subsystems. After that, we present two applications of the unitary decomposition in Secs. IV and V, which can be read independently of each other. Specifically, in Sec. IV, we use the unitary decomposition to derive the improved Lieb-Robinson bound for long-range interactions. In Sec. V, we analyze the performance of the HHKL algorithm from Ref. [27] when applied to simulating long-range interacting systems. We conclude in Sec. VI with an outlook for the future.

II. SUMMARY OF RESULTS

In this section, we summarize our main results for the case of a one-dimensional lattice. Without loss of generality, we assume that the distance between neighboring sites is one. The unitary decomposition technique in Sec. III is generalized from a similar result for short-range interactions in Ref. [27]. We use it to approximate the evolution of a long-range interacting system ABC by three sequential evolutions of its subsystems AB, B, and BC (see Fig. 1). We assume that the interaction strength between any two sites in the system is bounded by $1/r^\alpha$, with r being the distance between the sites and α a non-negative constant. This restriction on the Hamiltonian norm also sets the time unit for the evolution of the system.

There are two sources of error in the approximation: one due to the truncation of the Hamiltonian of the system ABC (we ignore the interactions that connect A and C), and the other due to the Hamiltonians of the subsystems AB, B, and BC not commuting with each other. For a fixed value of α, if the distance ℓ between the two regions A and C [see Fig. 1(a)] is large enough, namely $\ell \gg \alpha$, the two error sources have the same scaling with ℓ. To estimate the error,

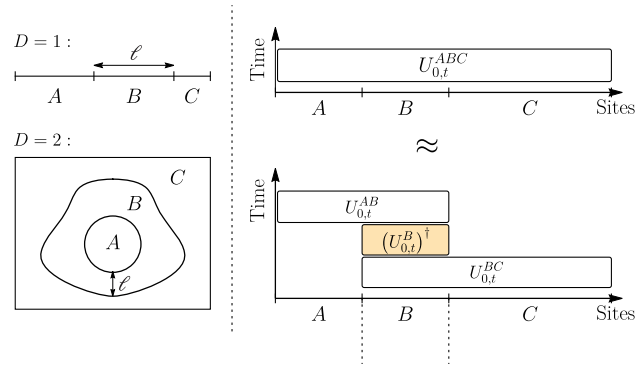

FIG. 1. A demonstration of the unitary decomposition in Lemma 1. (a) The three disjoint regions A, B, C in $D = 1$ and $D = 2$ dimensions with A convex and compact. (b) Lemma 1 allows the evolution of the whole system to be approximated by a series of three evolutions of subsystems. The horizontal axis lists the sites in each of the three sets A, B, C (not necessarily according to their geometrical arrangement, particularly in higher dimensions). Each box is an evolution for time t of a Hamiltonian supported on the sites the box covers. These evolutions can be forward (white fill) or backward (orange fill, with dagger) in time.

for example, from the truncation, we sum over interactions connecting sites in A and C, and obtain a total error of $\mathcal{O}(1/\ell^{\alpha-2})$ (in one dimension) for the approximation in the unitary decomposition (as shown in the Appendix A 1).

In Sec. IV, we use the unitary decomposition to prove a Lieb-Robinson bound for long-range interactions that is stronger than previous bounds, including the one we use in the proof of the unitary decomposition. The subject of such a bound is usually the norm of the commutator $\|[O_X(T), O_Y]\|$ between an operator $O_X(T) = U_T^\dagger O_X U_T$ evolved under a long-range Hamiltonian for time T and another operator O_Y supported on a set Y that is at least a distance R away from the support X of O_X. Here, we briefly explain the essence of the proof using a one-dimensional system with fixed α and large enough $R, T \gg \alpha$ as an example. The strategy is to use the aforementioned unitary decomposition to construct another unitary $\tilde{U}$ such that (i) $\tilde{U}^\dagger O_X \tilde{U}$ approximates $U_T^\dagger O_X U_T$ and (ii) $\tilde{U}^\dagger O_X \tilde{U}$ commutes with O_Y, so the commutator norm $\|[O_X(T), O_Y]\|$ will be approximately zero, up to the error of our approximation. For fixed α, we consider $M \propto T$ equal time slices and use the unitary decomposition to extract the relevant parts from the evolution U_T in each time slice. Each time we decompose a unitary, we choose the subsystems A, B, C so that only A overlaps with the supports of the unitaries from the previous time slices (see Fig. 2), and therefore the evolutions of B and BC can be commuted through O_X to cancel their counterparts from $U_T^\dagger$ [Figs. 2(b) and 2(d)]:

$$\begin{aligned}(U^{ABC})^\dagger O_X U^{ABC} &\approx (U^{AB})^\dagger U^B (U^{BC})^\dagger O_X U^{BC} (U^B)^\dagger U^{AB} \\ &= (U^{AB})^\dagger O_X U^{AB}. \end{aligned} \tag{1}$$

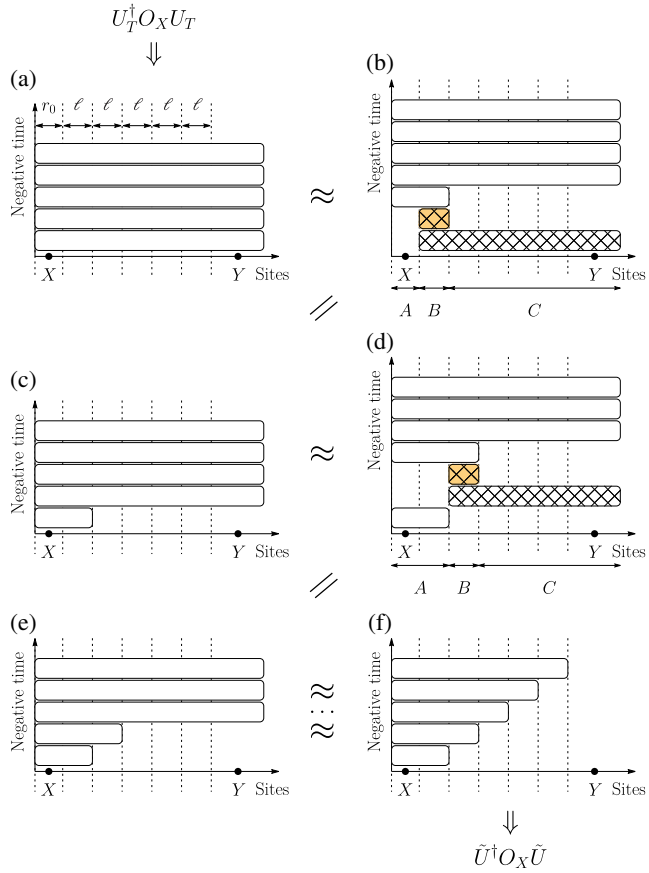

FIG. 2. A step-by-step construction of the unitary $\tilde{U}$ such that $\tilde{U}^\dagger O_X \tilde{U} \approx U_T^\dagger O_X U_T$. Each box represents an evolution of the subsystem covered by the width of the box for a fixed time. The colors of the boxes follow the same convention as in Fig. 1. In panel (a), the unitary U_T is written as a product of evolutions of the same system in $M = 5$ consecutive time slices. (b) The evolution in the last (bottom) time slice is decomposed using the method in Fig. 1, with the choice of subsystems A, B, C such that X is contained in A. The evolutions of the subsystems B and BC (hatched boxes) therefore commute with O_X and cancel out with their counterparts from $U_T^\dagger$, resulting in (c). In panel (d), we repeat the procedure for the second-from-bottom time slice, but note the different choice of A, B, C from (b). This difference is necessary to ensure that the evolutions of B and BC commute with the evolution(s) from the previously decomposed time slice(s). We then commute them through O_X again and remove them from the construction of $\tilde{U}$ in (e). Repeatedly applying the unitary decomposition for the other time slices, we obtain the unitary $\tilde{U}$ in (f), which is supported on a smaller region than the original unitary U_T. With a proper choice of the size ℓ of B, we can make sure that Y lies outside this region, and, therefore, $\tilde{U}$ commutes with O_Y.

The remaining evolutions that contribute to the construction of $\tilde{U}$ are supported entirely on a ball of radius $\sim M\ell$ around X, where ℓ is the size of B and is chosen to be the same in all time slices. By choosing $\ell \sim R/M$ and $M\ell < R$ so that Y lies outside this ball, the commutator norm $\|[O_X(T), O_Y]\|$ is at most the number of time slices multiplied by $\mathcal{O}(1/\ell^{\alpha-2})$, which is the decomposition error per time slice. Therefore, we obtain a Lieb-Robinson bound for long-range interactions in one dimension:

$$\|[O_X(T), O_Y]\| \le c_{\mathrm{lr},\alpha} \frac{T}{\ell^{\alpha-2}} = c_{\mathrm{lr},\alpha} \frac{T^{\alpha-1}}{R^{\alpha-2}}, \tag{2}$$

where $c_{\mathrm{lr},\alpha}$ is a constant that may depend on α, but not on T, R. Setting the commutator norm to a small constant yields the causal region inside the effective light cone: $T \gtrsim R^{(\alpha-2)/(\alpha-1)}$. For comparison, the previous best Lieb-Robinson bound produces a light cone $T \gtrsim R^{(\alpha-2)/\alpha}$ [6]. Our bound is therefore tighter in the asymptotic limit of large R and large T, while its proof is substantially more intuitive than in Ref. [6]. A more careful analysis (Sec. IV) shows that our light cone also becomes linear in the limit $\alpha \to \infty$, where the power-law decaying interactions are effectively short-range. Moreover, our bound works for arbitrary time T, while the bound in Ref. [6] applies only in the long-time limit. We provide a more rigorous treatment as well as a bound for D-dimensional systems in Sec. IV.

Section V then discusses the original motivation for the unitary decomposition—digital quantum simulation—in the case of long-range interactions that decay as a power law. For $\alpha > 2D$, our analysis shows that the HHKL algorithm [27] requires only $\mathcal{O}(Tn(Tn/\varepsilon)^{2D/\alpha-D} \times \log(Tn/\varepsilon))$ two-qubit gates to simulate the evolution of a system of n sites arranged in a D-dimensional lattice for time T with an error at most ε. For large α, the gate count of the algorithm scales with n significantly better than other algorithms.

III. FRAMEWORK

In this section, we present the technique for approximating the time evolution of a system by evolutions of subsystems. We later use this technique to derive a stronger Lieb-Robinson bound (Sec. IV) and an improved quantum simulation algorithm (Sec. V) for systems with long-range interactions.

We consider n sites arranged in a D-dimensional lattice $\Lambda \subset \mathbb{N}^D$ of size $L = \mathcal{O}(n^{1/D})$ and $D \ge 1$. Recall that, without loss of generality, we assume the spacing between neighboring lattice sites is one. This assumption sets the unit for distances between sites in the lattice. We shall embed the lattice Λ into the real space $\mathbb{R}^D$. The intersection $X \cap \Lambda$ therefore contains every lattice site in a subset $X \subset \mathbb{R}^D$. The system evolves under a (possibly) time-dependent Hamiltonian $H_\Lambda(t) = \sum_{\vec{i},\vec{j}} h_{\vec{i},\vec{j}}(t)$, with $h_{\vec{i},\vec{j}}(t)$ being the interaction between two sites $\vec{i}, \vec{j} \in \Lambda$. Without ambiguity, we may suppress the time dependence in the Hamiltonians. We say a system has power-law decaying interactions if $\|h_{\vec{i},\vec{j}}\| \le 1/\|\vec{i} - \vec{j}\|^\alpha$, where $\|\cdot\|$ denotes both the matrix and the vector 2-norms, for some non-negative constant α and for all $\vec{i} \ne \vec{j}$. [Note that $h_{\vec{i},\vec{i}}$

may have arbitrarily large norm.] For readability, we denote by $H_X = \sum_{\vec{i},\vec{j}\in X} h_{\vec{i},\vec{j}}$ the terms of H_Λ that are supported entirely on a subset $X \cap \Lambda$, and by $U^X_{t_1,t_2} \equiv \mathcal{T}\exp(-i\int_{t_1}^{t_2} H_X dt)$ the evolution unitary under H_X from time t_1 to t_2, where $\mathcal{T}$ is the time-ordering operator. We also denote by $\mathrm{dist}(X,Y)$ the minimum distance between any two sites in X and Y, by $X^c = \mathbb{R}^D \backslash X$ the complement of X in real space, by ∂X the boundary of a compact subset X, by $\Phi(X)$ the area of ∂X, and by XY the union $X \cup Y$. In the following, we keep track of how errors scale with time, distance, and α, while treating the dimension D as a constant.

We now describe how to approximate the evolution of the system to arbitrary precision by a series of evolutions of subsystems using a technique we generalize from Ref. [27].

Lemma 1.—Let $A, B, C \subset \mathbb{R}^D$ be three distinct regions with nonempty interiors such that $A \cup B \cup C = \mathbb{R}^D$. Let A be both compact (closed and bounded) and convex. We have

$$\|U^{ABC}_{0,t} - U^{AB}_{0,t}(U^B_{0,t})^\dagger U^{BC}_{0,t}\| \le c_0(e^{vt}-1)\Phi(A)\xi_\alpha(\ell),$$

with

$$\xi_\alpha(\ell) = \left(\frac{16}{1-\gamma}\right)^\alpha \frac{1}{\ell^{\alpha-D-1}} + e^{-\gamma\ell}, \tag{3}$$

for all $\alpha > D+1$. Here, $v, c_0 \in \mathbb{R}^+$ are positive constants, γ is a constant that can be chosen arbitrarily in the range (0,1), and $\ell = \mathrm{dist}(A,C)$ is the distance between sets A and C.

We emphasize that this lemma applies to arbitrary sets A that are both convex and compact. The sets we focus on include D balls and hyperrectangles in $\mathbb{R}^D$. The former geometry is relevant in the proof of our new Lieb-Robinson bound, the latter in the analysis of the HHKL algorithm for long-range interactions.

Lemma 1 allows us to approximate the evolution of a long-range interacting system ABC by that of subsystems AB, B, BC (Fig. 1). The features of the function $\xi_\alpha(\ell)$ are better understood by considering two limiting cases of physical interest. First, when α is finite and ℓ (the distance between A and C) is large compared to α, the function $\xi_\alpha(\ell)$ behaves like

$$\mathcal{O}\left(\frac{1}{\ell^{\alpha-D-1}}\right), \tag{4}$$

which decays only polynomially with ℓ. In the second limit, as $\alpha \to \infty$ for a large but finite ℓ, we recover from $\xi_\alpha(\ell)$ the exponentially decaying error bound $e^{-\gamma\ell}$—a trademark of finite-range interactions [1,27].

The proof of Lemma 1, while more general, bears close resemblance to the corresponding analysis for short-range interactions in Ref. [27]. However, there are two key differences. First, in order to make the approximation in Lemma 1, some interactions between sites separated by a distance greater than ℓ are truncated from the Hamiltonian. While such terms vanish in a system with short-range interactions, here they contribute $\mathcal{O}(\Phi(A)/\ell^{\alpha-D-1})$ to the error of the approximation. In addition, instead of the original Lieb-Robinson bound [1] which applies only to systems with short-range interactions, we use the generalization of the bound for long-range interactions by Gong *et al.* [5]. The result is an approximation error that decays with ℓ polynomially as $\mathcal{O}(\Phi(A)/\ell^{\alpha-D-1})$, in addition to the exponentially decaying error that exists already for short-range interactions. Nevertheless, the error can always be made arbitrarily small by choosing ℓ to be large enough.

In Sec. III A, we present the proof of Lemma 1. After that, we demonstrate the significance of Lemma 1 with two applications: a stronger Lieb-Robinson bound for long-range interacting systems (Sec. IV) and an improved error bound for simulating these systems (Sec. V). Both sections are self-contained, and readers may elect to focus on either of them.

A. Error bound on the unitary decomposition

Here, we outline the proof of Lemma 1. Similar to Ref. [27], we begin with an identity:

$$U^{ABC}_{0,t} = U^{AB}_{0,t}U^C_{0,t}\underbrace{(U^C_{0,t})^\dagger(U^{AB}_{0,t})^\dagger U^{ABC}_{0,t}}_{=W_t}. \tag{5}$$

Our aim is to approximate W_t by $(U^C_{0,t})^\dagger(U^B_{0,t})^\dagger U^{BC}_{0,t}$, from which Lemma 1 will follow. For that, we look at the generator of W_t [27], i.e., a Hamiltonian $\mathcal{G}_t$, such that

$$\frac{dW_t}{dt} = -i\mathcal{G}_t W_t, \tag{6}$$

for all time. Exact differentiation of W_t yields [40,41]

$$\mathcal{G}_t = (U^C_{0,t})^\dagger(U^{AB}_{0,t})^\dagger\underbrace{(H_{ABC} - H_{AB} - H_C)}_{=H_{A:C}+H_{B:C}}U^{AB}_{0,t}U^C_{0,t} \tag{7}$$

$$= (U^C_{0,t})^\dagger(U^{AB}_{0,t})^\dagger H_{B:C}U^{AB}_{0,t}U^C_{0,t} + \delta_{\mathrm{trunc}} \tag{8}$$

$$= (U^C_{0,t})^\dagger(U^B_{0,t})^\dagger H_{B:C}U^B_{0,t}U^C_{0,t} + \delta_{\mathrm{overlap}} + \delta_{\mathrm{trunc}}, \tag{9}$$

where $H_{X:Y} = \sum_{i\in X, j\in Y} h_{ij}(t)$ denotes the sum of terms supported across disjoint sets X and Y, and $\delta_{\mathrm{trunc}}, \delta_{\mathrm{overlap}}$ are error terms we now define and evaluate. Note that the first term in Eq. (9) is the generator of $(U^C_{0,t})^\dagger(U^B_{0,t})^\dagger U^{BC}_{0,t}$—the unitary with which we aim to approximate W_t.

In contrast to the approximation for short-range interacting systems in Ref. [27], there are two sources of error in Eq. (9). The first error term δ_{trunc} arises after we discard $H_{A:C}$ from Eq. (7). For the short-range interactions in

Ref. [27], this error vanishes when the distance ℓ between A and C is larger than the interaction range. However, in our case, there is a nontrivial truncation error associated with ignoring long-range interactions between A and C:

$$\|\delta_{\mathrm{trunc}}\| = \|H_{A:C}\| = c_{\mathrm{tr}} 2^{\alpha} \frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{10}$$

for $\alpha > D+1$, where c_{tr} is a constant [Eq. (A5)], $\ell = \mathrm{dist}(A, C)$ is the distance between A and C. The factor of $1/l^{\alpha}$ in the bound comes from the requirement that the two-body interactions decay as a power law $1/r^{\alpha}$, while the term ℓ^{D} is due to the sum over all sites in the D-dimensional set C. Another factor of $\ell\Phi(A)$ arises after summing over the volume of A, which we assume to be a compact and convex set. The detailed evaluation of the norm is presented in Appendix A 1.

The other error, which we define to be $\delta_{\mathrm{overlap}}$, is the result of the approximation used between Eqs. (8) and (9). In the former equation, the operator evolves under $H_{AB} + H_C$, whereas in the latter, it evolves under the reduced Hamiltonian $H_B + H_C$, thus incurring the error

$$\|\delta_{\mathrm{overlap}}\| = \|(U_{0,t}^{AB})^{\dagger} H_{B:C} U_{0,t}^{AB} - (U_{0,t}^{B})^{\dagger} H_{B:C} U_{0,t}^{B}\|. \tag{11}$$

To understand why $\|\delta_{\mathrm{overlap}}\|$ is small, recall that $H_{B:C}$ is the sum of terms $h_{\vec{b},\vec{c}}$ that are supported on two sites $\vec{b} \in B$ and $\vec{c} \in C$. Since the strengths of such terms decay as $1/r^{\alpha}$ (with r the distance between the sites $\vec{b}$ and $\vec{c}$), the main contribution to $H_{B:C}$—and thus to $\delta_{\mathrm{overlap}}$—comes from the terms where $\vec{b}$ and $\vec{c}$ are spatially close to each other. But since the sets A, C are separated by a large distance ℓ, if the site $\vec{b}$ is close to C, then it must be far from A. Thus, the evolution of $h_{\vec{b},\vec{c}}$ for a short time under H_{AB} can be well approximated by evolution under H_B alone. In Appendix A 2, we make this intuition rigorous using the generalization of the Lieb-Robinson bound by Gong *et al.* [5] to systems with long-range interactions.

In the end, we obtain the following bound on $\delta_{\mathrm{overlap}}$:

$$\|\delta_{\mathrm{overlap}}\| \le c_{\mathrm{ov}}(e^{vt} - 1)\Phi(A)\left[\frac{(\frac{16}{1-\gamma})^{\alpha}}{\ell^{\alpha-D-1}} + \frac{1}{e^{\gamma\ell}}\right], \tag{12}$$

where c_{ov} is a constant [Eq. (A24)] and $\gamma \in (0, 1)$ is a free parameter. The bound has contributions from two competing terms: one that decays polynomially with ℓ and another that decays exponentially. The polynomially decaying term is dominant for fixed α and large ℓ, whereas the exponentially decaying term prevails as $\alpha \to \infty$ for fixed ℓ. The errors δ_{trunc} and $\delta_{\mathrm{overlap}}$ in approximating the generator combine to give an overall error in approximating W_t with $(U_{0,t}^{C})^{\dagger}(U_{0,t}^{B})^{\dagger}U_{0,t}^{BC}$ (see Appendix B). From this, we obtain the error bound in Lemma 1, with $c_0 = \max\{c_{\mathrm{tr}}, c_{\mathrm{ov}}\}/v$.

Before discussing applications of Lemma 1, we pause here to note that the Lieb-Robinson bound in Gong *et al.* [5] used in the above analysis is not the tightest-known bound for long-range interactions [6]. Our use of this bound, however, does not lead to a suboptimal error bound in Lemma 1. For finite α, the error bound is dominated by the polynomially decaying term $1/\ell^{\alpha-D-1}$, which arises from the truncation error δ_{trunc} rather than $\delta_{\mathrm{overlap}}$. Therefore, this error term would not benefit from a tighter Lieb-Robinson bound. In the limit $\alpha \to \infty$, on the other hand, we shall see later that the lemma already reproduces the short-range Lieb-Robinson bound, which is optimal up to a constant factor. Thus, we expect that using stronger Lieb-Robinson bounds would produce no significant improvement for the error bound in Lemma 1.

IV. STRONGER LIEB-ROBINSON BOUND

In this section, we use Lemma 1 to derive a stronger Lieb-Robinson bound for long-range interactions. The first generalization of the Lieb-Robinson bound to power-law decaying interactions was given by Hastings and Koma [4]. However, their bound diverges in the limit $\alpha \to \infty$, where the power-law decaying interactions are effectively short-range. Later, Gong *et al.* [5] derived a different bound that, in this limit, does indeed converge to the Lieb-Robinson bound for short-range interactions. While we use this bound in Sec. III to prove Lemma 1, we will also show that by using this lemma we can in turn derive a Lieb-Robinson bound for long-range interactions that is stronger than the one in Gong *et al.* In fact, our bound produces a tighter effective light cone than even the strongest Lieb-Robinson bound for long-range interactions known previously [6].

Recall that the subject of a Lieb-Robinson bound is the commutator norm

$$\mathcal{C}(T, R) \equiv \|[(U_{0,T}^{\Lambda})^{\dagger} O_X U_{0,T}^{\Lambda}, O_Y]\|, \tag{13}$$

where O_X, O_Y are two operators supported, respectively, on two sets X, Y geometrically separated by a distance R, and $U_{0,T}^{\Lambda}$ is the time-evolution unitary of the full lattice Λ under a power-law decaying Hamiltonian, as defined above.

To compare different bounds, we analyze their effective light cones, which, up to constant prefactors, predict the minimum time it takes for the correlator $\mathcal{C}(T, R)$ to reach a certain value. For example, the original Lieb-Robinson bound [1] produces a linear light cone $T \gtrsim R$ for short-range interactions. For long-range interactions, Hastings and Koma [4] first showed that $\mathcal{C}(T, R) \le c e^{vT}/R^{\alpha}$ for some (α-dependent) constants c, v. By setting $\mathcal{C}(T, R)$ equal to a constant, the bound gives an effective light cone $T \gtrsim \log R$ in the limit of large T and R. Gong *et al.* [5] later achieved a tighter light cone that is linear for short distances and becomes logarithmic only for large R. Shortly after,

Foss-Feig *et al.* [6] derived a bound with a polynomial light cone:

$$T \gtrsim R^{(\alpha-2D)/(\alpha-D+1)}. \tag{14}$$

Equation (14) was the tightest light cone known previously.

In the remainder of this section, we use Lemma 1 to derive a Lieb-Robinson bound for long-range interactions that produces an effective light cone tighter than the one in Ref. [6], while also using a much more intuitive approach. In addition, our bound works for all times, unlike the bound in Ref. [6], which applies only in the long-time limit.

Theorem 1 (Lieb-Robinson bound for long-range interactions).—Suppose O_X is supported on a fixed subset X. For $\alpha > 2D$, we have

$$\mathcal{C}(T,R) \leq \begin{cases} c_{\mathrm{lr}} e^{\alpha} T R^{D-1} \xi_\alpha(\frac{R\alpha}{vT}) & \text{if } vT \geq \alpha \\ \tilde{c}_{\mathrm{lr}}(e^{vT}-1)\xi_\alpha(R) & \text{if } vT < \alpha. \end{cases} \tag{15}$$

Here $R = \mathrm{dist}(X,Y)$ is the distance between the supports of O_X and O_Y, $c_{\mathrm{lr}}, \tilde{c}_{\mathrm{lr}}, v$ are constants that may depend only on D (defined in Appendix C), and ξ_α is given by Eq. (3).

Before we prove Theorem 1, let us analyze the features of the bound. Although the general bound in Eq. (15) looks complicated, it can be greatly simplified in some limits of interest. For example, for finite α, in the limit of large $vT > \alpha$ and large R such that $R/(vT) \gg \alpha$, the term algebraically decaying with $R/(vT)$ in $\xi_\alpha[R\alpha/(vT)]$ dominates the exponentially decaying one [see also Eqs. (3) and (C16)]. Therefore, the Lieb-Robinson bound in this limit takes the form

$$\mathcal{C}(T,R) \leq c_{\mathrm{lr},\alpha} \frac{T^{\alpha-D}}{R^{\alpha-2D}}, \tag{16}$$

where $c_{\mathrm{lr},\alpha}$ is finite and may depend on α [Eq. (C18)]. We can immediately deduce the effective light cone given by our bound for a finite α:

$$T \gtrsim R^{(\alpha-2D)/(\alpha-D)}, \tag{17}$$

which is tighter than Eq. (14) (as given by Ref. [6]). In particular, for α close to $2D$, the exponent in Eq. (17) can be almost twice that of Ref. [6] (the larger the exponent, the tighter the light cone).

On the other hand, in the limit $\alpha \to \infty$, vT is finite and therefore always less than α. Hence our bound converges to the short-range bound $\mathcal{C}(T,R) \leq 2\tilde{c}_{lr} e^{vT-\gamma R}$. We note that in this limit, the exponent of the light cone in Eq. (17) also converges to one, which corresponds to a linear light cone, at a linear convergence rate [see Eq. (C20) for details]. These behaviors are naturally expected since a power-law decaying interaction with very large α is essentially a short-range interaction.

As mentioned earlier, we derive Theorem 1 by constructing a unitary $\tilde{U}$ such that (i) $\tilde{U}^\dagger O_X \tilde{U}$ approximates $(U^\Lambda_{0,T})^\dagger O_X U^\Lambda_{0,T}$ and (ii) $\tilde{U}$ commutes with O_Y. We note that $\tilde{U}$ does not necessarily approximate $U^\Lambda_{0,T}$. It then follows from the two requirements that the commutator norm $\mathcal{C}(T,R)$, defined in Eq. (13), is upper bounded by the error of the approximation in (i).

We also note that the assumption on the norms of the interactions being bounded excludes several physical systems whose local dimensions are unbounded, e.g., bosons (see Refs. [42,43] for discussions of information propagation and Lieb-Robinson bounds in these systems).

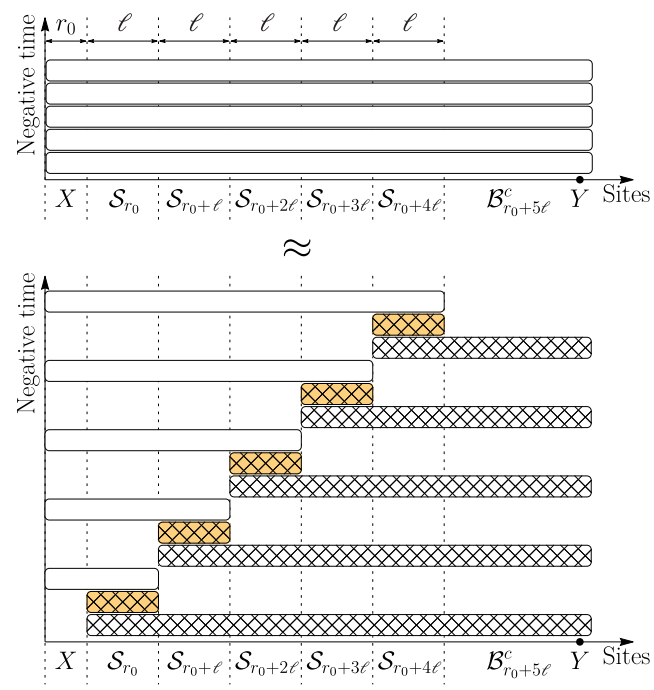

FIG. 3. A construction of the unitary $\tilde{U}$ which results in an improved Lieb-Robinson bound for long-range interactions in Theorem 1. The horizontal axes list the sites in each subset. Here $\mathcal{B}_r$ denotes a D ball of radius r centered on X, and $\mathcal{S}_r = \mathcal{B}_{r+\ell} \backslash \mathcal{B}_r$ a D-dimensional shell of inner radius r and outer radius $r+\ell$, for some parameter ℓ to be chosen later. (See Fig. 7 in Appendix C for an illustration of the sets.) The evolution unitaries are represented by boxes with the same color convention as in Fig. 1. We first divide the interval $[0,T]$ into $M = 5$ equal time slices (upper panel). Note that because we consider $O_X(T)$ in the Heisenberg picture, the vertical axis is therefore backward in time so that the bottom time slice will correspond to the first unitary applied on O_X. The evolution during each time slice is approximated by three evolutions of subsystems using Lemma 1 (lower panel). The bottom two unitaries have their supports outside X and therefore commute with O_X. They cancel with their Hermitian conjugates from $\tilde{U}^\dagger$ in $\tilde{U}^\dagger O_X \tilde{U}$. Repeating the argument for higher time slices, we can eliminate some unitaries (hatched boxes) from the construction of $\tilde{U}$. Finally, we are left with $\tilde{U}$ consisting only of unitaries (white boxes) that are supported entirely on the D ball $\mathcal{B}_{r_0+5\ell}$ of radius $r_0 + 5\ell$. Therefore, $\tilde{U}$ commutes with O_Y, whose support lies in the complement $\mathcal{B}^c_{r_0+5\ell}$ of $\mathcal{B}_{r_0+5\ell}$.

However, our Lieb-Robinson bound may still apply if the dynamics of the systems can be restricted to local Hilbert subspaces which are finite dimensional. Examples of such situations include trapped ions in the perturbative regime [29] and noninteracting bosons [26].

To construct $\tilde{U}$, we use Lemma 1 to decompose the unitary $U_{0,T}^{\Lambda}$ into unitaries supported on subsystems, each of which either contains X or is disjoint from X. The unitaries of the latter type can be commuted through O_X to cancel out with their Hermitian conjugates from $(U_{0,T}^{\Lambda})^{\dagger}$. The remaining unitaries form $\tilde{U}$, which is supported on a smaller subset than $U_{0,T}^{\Lambda}$. In particular, with a suitable decomposition, the support of $\tilde{U}$ can be made to not contain Y, and, therefore, $\tilde{U}$ commutes with O_Y. The step-by-step construction of the unitary $\tilde{U}$ has also been briefly described earlier in Sec. II and in Fig. 2, using the specific case of a one-dimensional system with a finite α. This construction immediately generalizes to higher dimensions and to arbitrary α, including the $\alpha \to \infty$ limit. The construction of $\tilde{U}$ for arbitrary D is summarized in Fig. 3.

We note that there is more than one way to decompose the unitary $U_{0,T}^{\Lambda}$ in the construction of $\tilde{U}$. Different constructions of $\tilde{U}$ result in different approximation errors, each of which provides a valid bound on the commutator norm $\mathcal{C}(T,R)$. Therefore, the goal is to find a construction of $\tilde{U}$ with the least approximation error. In Appendix C, we present the construction that results in the bound in Theorem 1. Although we have evidence suggesting that the construction is optimal, we do not rule out the existence of a better construction.

V. BETTER PERFORMANCE OF DIGITAL SIMULATION

In this section, we generalize the algorithm in Ref. [27] to simulating long-range interactions. In general, the aim of quantum simulation algorithms is to approximate the time evolution unitary $U_{0,T}^{\Lambda}$ using the fewest number of primitive, e.g., two-qubit, quantum gates. Here, we show that in addition to the stronger Lieb-Robinson bound presented in the previous section, Lemma 1 can also be used to perform error analysis for the HHKL algorithm (Ref. [27]) in the case of interactions that decay as a power law, therefore improving the theoretical gate count of digital quantum simulation for such interactions.

Using the best-known rigorous error bounds, simulations based on the first-order Suzuki-Trotter product formula [44] use $\mathcal{O}(T^2 n^6/\varepsilon)$ gates to simulate the evolution $U_{0,T}^{\Lambda}$ of a time-dependent Hamiltonian on n sites up to a fixed error ε. (In this section, the big $\mathcal{O}$ is with respect to n, T, and $1/\varepsilon$.) The generalized $(2k)$th-order product formula uses $\mathcal{O}(n^2(Tn^2)^{1+1/(2k)}/\varepsilon^{1/2k})$ quantum gates. While this scaling asymptotically approaches $\mathcal{O}(Tn^4)$ as $k \to \infty$, it suffers from an exponential prefactor of 5^{2k} [45]. More advanced algorithms, e.g., those using quantum signal processing (QSP) [46] or linear combinations of unitaries (LCU) [47], can reduce the gate complexity to $\mathcal{O}(Tn^3 \log(Tn/\varepsilon))$. Our error analysis below (Lemma 2) reveals that, when α is large, the number of quantum gates required by the HHKL algorithm to simulate long-range interactions scales better as a function of the system size than previous algorithms.

The HHKL algorithm itself uses either the QSP algorithm or the LCU algorithm as a subroutine to simulate the dynamics of a subset of the sites for one time step. Although the QSP algorithm does not handle time-dependent Hamiltonians, LCU can be applied to time-dependent Hamiltonians. Our results assume that (i) the local terms $h_{\vec{\imath},\vec{\imath}}(t)$ have bounded norms for all $\vec{\imath} \in \Lambda$, and (ii) the Hamiltonian $H_{\Lambda}(t)$ varies slowly and smoothly with time so that $h'_{|X|} \equiv \max_t \|\partial H_t^X/\partial t\|$ exists and scales at most polynomially with $|X|$ for all subsets $X \subset \Lambda$. These restrictions allow portions of the system to be faithfully simulated using LCU (or QSP, for a time-independent Hamiltonian).

A. HHKL-type algorithm for simulating long-range interactions

Although Ref. [27] focused on simulating short-range interactions, their (HHKL) algorithm can also be used to simulate long-range interactions. Here, we analyze the performance of their algorithm in simulating such systems. In the HHKL algorithm [27], the evolution of the whole system is decomposed, using Lemma 1, into elementary unitaries, each evolving a subsystem of at most $(2\ell)^D$ sites, where ℓ is again a length scale to be chosen later. For a fixed time t, the algorithm simply simulates each of these elementary unitaries using one of the existing quantum simulation algorithms. In particular, we shall use LCU or (for a time-independent Hamiltonian) QSP due to their logarithmic dependence on the accuracy.

In this section, we assume α is finite and analyze the gate count in the limit of large system size $n \gg \alpha$. As a consequence, the block size ℓ can also be taken to be much larger than α. For simplicity, we will not keep track of constants that may depend on α. Recall that in this limit, the error bound in Lemma 1 is at most

$$\mathcal{O}\left(\frac{\Phi(A)}{\ell^{\alpha-D-1}}\right), \tag{18}$$

where we have assumed $t = \mathcal{O}(1)$. Using Lemma 1, we obtain the error bound for the first step of the HHKL algorithm, which can be summarized by the following lemma.

Lemma 2 (HHKL decomposition).—There exists a circuit that approximates $U_{0,T}^{\Lambda}$ up to error $\mathcal{O}(Tn/\ell^{\alpha-D})$, where $\ell \leq n^{1/D}/2$ is a free parameter. The circuit has depth at most $3^D T$ and consists of $\mathcal{O}(Tn/\ell^D)$ elementary unitaries, each of which evolves a subsystem supported on at most $(2\ell)^D$ sites for time $t = \mathcal{O}(1)$.

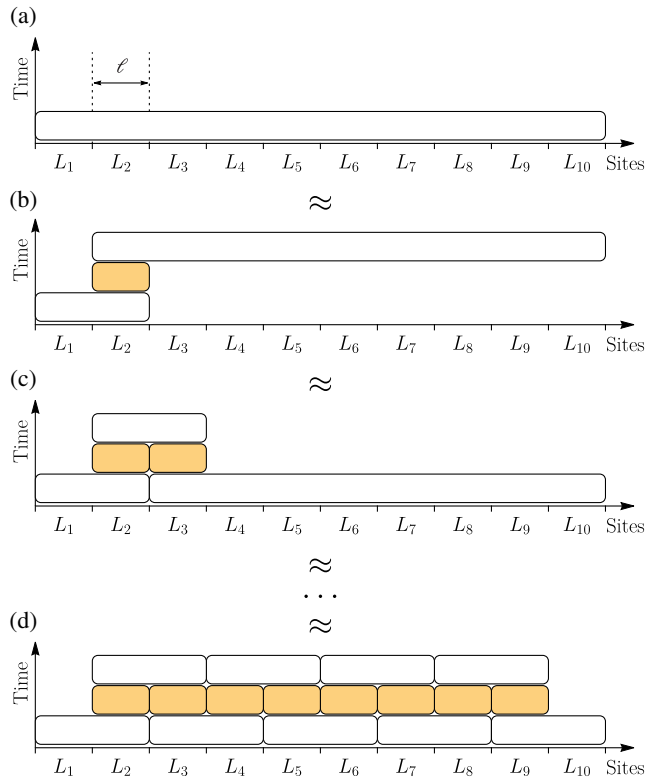

FIG. 4. A demonstration of the HHKL decomposition [27] of the evolution of a fixed time interval for a system with $m = 10$ blocks, each consisting of ℓ sites. As before, each box represents a unitary (white) or its Hermitian conjugate (orange) supported on the covered sites. Using Lemma 1, the HHKL decomposition approximates the evolution of the whole system (a) by three unitaries supported on subsystems (b). By applying Lemma 1 repeatedly [(c),(d)], the evolution of the whole system is decomposed into a series of evolutions of subsystems, each of size at most 2ℓ.

Proof.—We now demonstrate the proof by constructing the circuit for a one-dimensional lattice (Fig. 4). A generalization of the proof to arbitrary dimension follows the same lines and is presented in Appendix D.

First, we consider $M \propto T$ equal time intervals $0 = t_0 < t_1 < \ldots < t_M = T$, such that $t_{j+1} - t_j = t = T/M$ is a constant for all $j = 0, \ldots, M-1$. The simulation of $U^{\Lambda}_{0,T}$ then naturally decomposes into M consecutive simulations of $U^{\Lambda}_{t_j,t_{j+1}}$. We then divide the system into m consecutive disjoint blocks, each of size $\ell = n/m$ (Fig. 4). Denote by $L_k (k = 1, \ldots, m)$ the set of sites in the kth block. Using Lemma 1, we can approximate

$$U^{\Lambda}_{0,t} \approx U^{L_1 \cup L_2}_{0,t} (U^{L_2}_{0,t})^{\dagger} U^{L_2 \cup L_3 \cup \cdots \cup L_m}_{0,t}. \tag{19}$$

This approximation can be visualized using the top two panels of Fig. 4. Repeated application of Lemma 1 yields the desired circuit (bottom panel of Fig. 4), with each elementary unitary evolving at most 2ℓ sites for time t.

To obtain the error estimate in Lemma 2, we count the number of times Lemma 1 is used in our approximation. In each of the M time slices, we use the lemma $\mathcal{O}(m) = \mathcal{O}(n/\ell)$ times, each of which contributes an error of $\mathcal{O}(1/\ell^{\alpha-2})$ [see Eq. (18) with $\Phi = \mathcal{O}(1)$ in one dimension]. Therefore, with $M \propto T$, the error of using the constructed circuit to simulate $U^{\Lambda}_{0,T}$ is

$$\mathcal{O}\left(M \frac{n}{\ell} \frac{1}{\ell^{\alpha-2}}\right) = \mathcal{O}\left(\frac{Tn}{\ell^{\alpha-1}}\right), \tag{20}$$

as given in Lemma 2. ■

The error bound for the approximation in Lemma 2 leads to an upper bound on the gate complexity of digital quantum simulation, as stated in the following theorem.

Theorem 2.—For $\alpha > 2D$, there exists a quantum algorithm for simulating $U^{\Lambda}_{0,T}$ up to error at most ε with gate complexity

$$G_D = \mathcal{O}\left(Tn \left(\frac{Tn}{\varepsilon}\right)^{2D/(\alpha-D)} \log \frac{Tn}{\varepsilon}\right). \tag{21}$$

This gate complexity can be achieved by applying the HHKL algorithm [27] for long-range interactions, as described above. First, the evolution of the whole system $U^{\Lambda}_{0,T}$ is approximated by $\mathcal{O}(Tn/\ell^D)$ elementary unitaries as provided in Lemma 2. Each of these elementary unitaries is then simulated using one of the existing algorithms, e.g., LCU, with error that we require to be at most $\varepsilon \ell^D / Tn$. If the Hamiltonian is time independent, one can also use the QSP algorithm to simulate the elementary unitaries.

In the decomposition of the evolution, the accuracy of the approximation can be improved by increasing the block size ℓ. By Lemma 2, to achieve an overall error at most ε, we need

$$\ell \propto \left(\frac{Tn}{\varepsilon}\right)^{1/(\alpha-D)}. \tag{22}$$

When simulating the elementary unitaries, since each is an evolution of at most $(2\ell)^D$ sites for time $t = \mathcal{O}(1)$, the LCU algorithm with error at most $\varepsilon \ell^D / Tn$ uses $\mathcal{O}(\ell^{3D} \log(Tn/\varepsilon\, h'_{\ell^D}))$ two-qubit gates [45]. Recall that we assume h'_{ℓ^D} scales at most polynomially with ℓ^D. With the block size ℓ from Eq. (22), we find the total gate complexity of simulating the $\mathcal{O}(Tn/\ell^D)$ elementary unitaries is

$$G_D = \mathcal{O}\left(\frac{Tn}{\ell^D} \ell^{3D} \log\left(\frac{Tn}{\varepsilon} h'_{\ell^D}\right)\right) \tag{23}$$

$$= \mathcal{O}\left(Tn \left(\frac{Tn}{\varepsilon}\right)^{2D/(\alpha-D)} \log \frac{Tn}{\varepsilon}\right). \tag{24}$$

The scaling of G_D as a function of the system size n is significantly better than existing algorithms for large α. In particular, at $T = n$, this HHKL algorithm for long-

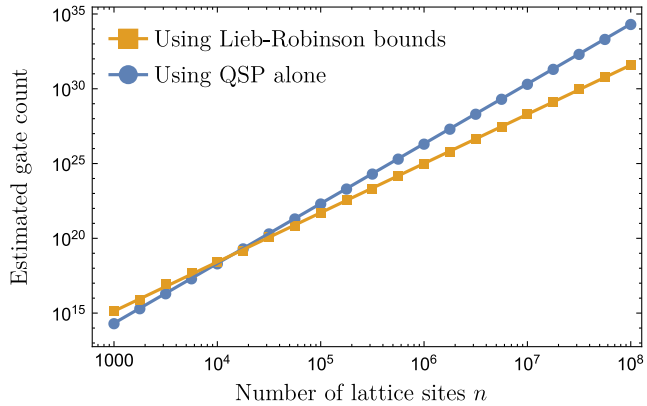

FIG. 5. The gate count for simulating the dynamics of a one-dimensional Heisenberg chain [Eq. (25)] of length n, with $\alpha = 4$, $T = n$, and $\varepsilon = 10^{-3}$. We compare the gate count of the HHKL algorithm (orange square) to the QSP algorithm (blue circle). Note that the HHKL algorithm based on Lieb-Robinson bounds also uses the QSP algorithm as a subroutine. We obtain the scatter points using the approach described in Appendix E and fit them to a power-law model (solid lines). The asymptotic scalings of the gate count obtained from the power-law fits ($n^{3.29}$ for HHKL, $n^{4.00}$ for QSP) agree well with our theoretical predictions (see Table I).

range interactions requires only $\mathcal{O}(n^{2+4D/(\alpha-D)}\log n)$ gates, while algorithms such as QSP or LCU use $\mathcal{O}(n^4 \log n)$ gates or more. Therefore, the algorithm provides an improvement for $\alpha > 3D$. However, the gate complexity of the algorithm depends polynomially on $1/\varepsilon$, in contrast to the logarithmic dependence achieved by QSP and LCU, and by the HHKL algorithm for systems with short-range interactions. While this $\mathrm{poly}(1/\varepsilon)$ scaling is undesirable, in practice, the total error of the simulation is often set to a fixed constant (for example, see Ref. [48]) and effectively the dependence of ε only contributes a prefactor to the gate complexity of the algorithm.

As an example, in Fig. 5, we estimate the actual gate count of the HHKL algorithm in simulating a Heisenberg chain [Eq. (25)] and compare it with the gate count of the QSP algorithm (up to the same error tolerance). Because of the $\mathrm{poly}(1/\varepsilon)$ overhead, the HHKL algorithm based on Lieb-Robinson bounds uses more quantum gates for simulating small systems, but eventually outperforms the QSP algorithm when the system size n is large.

It is also worth noting that, in the limit $\alpha \to \infty$, the gate complexity becomes $\mathcal{O}(Tn \log(Tn/\varepsilon))$, which coincides (up to a polylogarithmic factor) with the result for short-range interactions in Ref. [27]. This behavior is expected, given that a power-law decaying interaction with $\alpha \to \infty$ is essentially a nearest-neighbor interaction. However, we caution readers that at the beginning of this section, we have assumed that α is finite so that $n \gg \alpha$. Hence, the gate count in Eq. (24) is technically not valid in the limit $\alpha \to \infty$. Nevertheless, the error bound in Lemma 1 reproduces the estimate for short-range interactions in Ref. [27], and therefore, repeating the argument of this section in the limit $\alpha \to \infty$ should also reproduce the gate count for simulating short-range interactions in Ref. [27].

B. Numerical evidence of potential improvement

Up to now, we have seen that Lieb-Robinson bounds can improve the error bounds of quantum simulation algorithms, as demonstrated by the HHKL algorithm. We now provide numerical evidence hinting at the possibility of further improving the error bounds.

Although the HHKL algorithm outperforms previous ones when $\alpha > 3D$, it remains an open question whether there is a faster algorithm for simulating long-range interactions. We also note that the gate complexities are only theoretical upper bounds, and these algorithms may actually perform better in practice [49].

As an example, we compute the empirical gate count of a Suzuki-Trotter product formula simulation of a one-dimensional long-range interacting Heisenberg model,

$$H = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{|i-j|^4} \vec{\sigma}_i \cdot \vec{\sigma}_j + \sum_{i=1}^{n} B_i \sigma_i^z, \qquad (25)$$

where $B_j \in [-1, 1]$ are chosen uniformly at random and $\vec{\sigma}_j = (\sigma_j^x, \sigma_j^y, \sigma_j^z)$ denotes the vector of Pauli matrices on the qubit j. Specifically, we consider a simulation using the fourth-order product formula (PF4). We use a classical simulation to determine the algorithm's performance for systems of size $n = 4$ to $n = 12$ for time $T = n$, and extrapolate to larger systems. For each n, we search for the minimum number of gates for which the simulation error is at most $\varepsilon = 10^{-3}$. We plot in Appendix F this empirical gate count, which appears to scale only as $\mathcal{O}(n^{3.64})$ with the system size n. We list in Table I the gate counts of several popular algorithms for comparison. The theoretical gate complexity of PF4 is $\mathcal{O}(n^{5.75})$ [44], while the QSP and

TABLE I. A comparison between the gate complexity of several quantum simulation algorithms for simulating one-dimensional power-law systems at $T = n$ and $\alpha = 4$. Our analysis shows that the HHKL algorithm performs at least as well as the empirical gate count of PF4, while having a similar $\mathrm{poly}(1/\varepsilon)$ scaling with the error ε. It is not known whether the empirical gate count of PF4 can scale with ε better than suggested by the best proven bound (third row).

Algorithm	Scaling with $n = T$	Scaling with ε
Empirical PF4	$\mathcal{O}(n^{3.64})$	…
Our HHKL bound	$\mathcal{O}(n^{3.33} \log n)$	$\mathcal{O}(\log(1/\varepsilon)/\varepsilon^{0.67})$
PF4 bound [45]	$\mathcal{O}(n^{5.75})$	$\mathcal{O}(1/\varepsilon^{0.25})$
QSP bound [46]	$\mathcal{O}(n^4 \log n)$	$\mathcal{O}(\log(1/\varepsilon))$
LCU bound [47]	$\mathcal{O}(n^4 \log n)$	$\mathcal{O}(\log(1/\varepsilon))$

LCU algorithms both have complexity $\mathcal{O}(n^4 \log n)$. These numerics show that the PF4 algorithm for simulating long-range interacting systems performs better in practice than theoretically estimated; in fact, it even performs almost as well as the HHKL algorithm based on Lieb-Robinson bounds [which scales as $\mathcal{O}(n^{3.33} \log n)$ by our earlier analysis]. Whether other quantum simulation algorithms, including the HHKL algorithm, can perform better than suggested by the existing bounds remains an important open question.

VI. CONCLUSION AND OUTLOOK

To conclude, we derived an improved bound on how quickly quantum information propagates in systems evolving under long-range interactions. The bound applies to power-law interactions with $\alpha > 2D$, such as dipole-dipole interactions in 1D (often realizable with nitrogen-vacancy centers [33] or polar molecules [32]), trapped ions in 1D [28,29], and van der Waals-type interactions between Rydberg atoms [30] in either 1D or 2D. For finite $\alpha > 2D$, our Lieb-Robinson bound gives a tighter light cone than previously known bounds—including the one used in the proof of Lemma 1. As of yet, we are not aware of any physical systems that saturate the Lieb-Robinson bounds for power-law interactions, including the new bound. In the limit $\alpha \to \infty$, our bound asymptotically approaches the exponentially decaying bound for short-range interactions. Our bound gives a linear light cone only in this limit, however, and it remains an open question whether there exists a stronger bound with a critical α_c such that the light cone is exactly linear for $\alpha \geq \alpha_c$ [50]. Currently, there are no known methods for quantum information transfer that are faster than linear for $\alpha \geq D + 1$. It is possible, therefore, that a stronger bound exists with a finite $\alpha_c \geq D + 1$. It is our hope that the present work, as well as the techniques that we use, will help motivate the search for such stronger bounds.

Our technique immediately extends the HHKL algorithm in Ref. [27] to the digital quantum simulation of the above systems. Our error bounds indicate that the gate complexity of the algorithm is better than that of other state-of-the-art simulation algorithms when α is sufficiently large ($\alpha > 3D$), and matches that of the short-range algorithm when $\alpha \to \infty$.

However, the empirical scaling of other algorithms—such as product formulas—indicates that this gate complexity may only be a loose upper bound to the true quantum complexity of the problem. While a matching lower bound for the gate complexity of the HHKL algorithm is provided in Ref. [27] for Hamiltonians with short-range interactions, we do not know of any techniques that could provide a corresponding bound for long-range interactions. In addition to improving the quantum gate complexity, our results may also aid in the design of better classical algorithms for simulating long-range interacting quantum systems. In particular, while we still expect the classical gate complexity to be exponential in the simulation time, there may be room for a polynomial improvement.

While the use of Lieb-Robinson bounds to improve the performance of quantum algorithms is a natural extension of Haah *et al.*, the opposite direction—using quantum algorithms to improve Lieb-Robinson bounds—is new. The connection from quantum simulation algorithms to Lieb-Robinson bounds that we have established opens another avenue for the condensed matter and atomic, molecular, and optical physics communities to potentially benefit from future advances in quantum algorithms. In addition to proving a stronger Lieb-Robinson bound, the tools we developed may help to answer other questions regarding both short-range and long-range interacting systems. Using the same unitary construction as Theorem 1, we can generalize the bounds on connected correlators [35,36] to long-range interacting systems. Our results can also provide a framework for proving tighter bounds on higher-order commutators, such as out-of-time-order correlators [37,38]. Previous methods used to derive Lieb-Robinson bounds—due to their use of the triangle inequality early in their proofs—have not been able to capture the nuances in the growth of such correlators. In addition to the more intuitive proof of the Lieb-Robinson bounds, our framework can be used to provide an alternative, simpler proof of the classical complexity of the boson-sampling problem [39], which generalizes the result in Ref. [26] to long-range interactions and also to more general Hamiltonians with arbitrary local interactions [51]. By taking advantage of the unitary decomposition in Lemma 1, we obtain a longer time interval within which the sampler in Ref. [26] is efficient [51]. Moreover, by generalizing from two-body to many-body interactions, our technique may find applications in systems whose Hamiltonians include interaction terms between three or more sites, e.g., many-body localized systems in the l-bit basis [52].

ACKNOWLEDGMENTS

We thank G. H. Low, A. Deshpande, P. Titum, T. Zhou, and Z.-X. Gong for helpful discussions. M. C. T., A. Y. G., Z. E., and A. V. G. acknowledge funding from the U.S. Department of Energy ASCR Quantum Testbed Pathfinder program (Award No. DE-SC0019040), ARO MURI, ARO, NSF Ideas Lab on Quantum Computing, ARL CDQI, NSF PFC at JQI, AFOSR, and the U.S. Department of Energy BES Materials and Chemical Sciences Research for Quantum Information Science program (Award No. DE-SC0019449). A. M. C. and Y. S. acknowledge funding from ARO MURI, CIFAR, NSF, and the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Quantum Algorithms Teams and Quantum Testbed Pathfinder programs (Award No. DE-SC0019040). M. C. T. is supported in part by the NSF Grant No. NSF PHY-1748958 and the

Heising-Simons Foundation. A. Y. G. is supported by the NSF Graduate Research Fellowship Program under Grant No. DGE 1322106. J. R. G. is supported by the NIST NRC Research Postdoctoral Associateship Award and performed his work in part at the Aspen Center for Physics, which is supported by National Science Foundation Grant No. PHY-1607611. Z. E. is supported in part by the ARCS Foundation.

Note added.—Recently, Else *et al.* [53] posted their work on a different Lieb-Robinson bound for power-law decaying interactions. For Hamiltonians consisting of at most two-body interactions, the bound in Ref. [53] and our bound both apply in the same regime, $\alpha > 2D$ [54]. Within this regime, our bound results in a strictly tighter light cone than Ref. [53]. However, the bound in Ref. [53] also applies to Hamiltonians consisting of k-body interactions, for any integer k. Generalizing our framework to cover such k-body interactions would be an interesting future direction.

APPENDIX A: EVALUATIONS OF THE SUM IN LEMMA 1

In this Appendix, we show how we bound δ_{trunc} from Eq. (10) (Appendix A 1) and $\delta_{\mathrm{overlap}}$ from Eq. (11) (Appendix A 2) in the proof of Lemma 1.

1. Evaluation of δ_{trunc}

Here, we provide explicit calculations of δ_{trunc} in Eq. (10). Recall that $\ell = \mathrm{dist}(A, C)$ is the shortest distance between any two points in A and C. Therefore, $\|\vec{a} - \vec{c}\|$ is always greater than ℓ. For each $\vec{a} \in A$, let $\ell_{\vec{a}} = \mathrm{dist}(\vec{a}, C)$ be the minimum distance from $\vec{a}$ to the set C and $C_{\vec{a}} = \{\vec{\imath} \in \Lambda : \mathrm{dist}(\vec{a}, \vec{\imath}) \geq \ell_{\vec{a}}\}$. Clearly, C is a subset of $C_{\vec{a}}$. Therefore,

$$\|\delta_{\mathrm{trunc}}\| = \|H_{A:C}\| \leq \sum_{\vec{a}\in A}\sum_{\vec{c}\in C}\frac{1}{\|\vec{a}-\vec{c}\|^{\alpha}} \tag{A1}$$

$$\leq \sum_{\vec{a}\in A}\sum_{\vec{c}\in C_{\vec{a}}}\frac{1}{\|\vec{a}-\vec{c}\|^{\alpha}} = \sum_{\vec{a}\in A}\sum_{\substack{\vec{r}\\ \|\vec{r}\|\geq \ell_{\vec{a}}}}\frac{1}{\|\vec{r}\|^{\alpha}} \tag{A2}$$

$$\leq \lambda_1 \sum_{\vec{a}\in A}\frac{1}{(\ell_{\vec{a}} - \sqrt{D})^{\alpha - D}}, \tag{A3}$$

where λ_1 is a constant independent of $\vec{a}$ and the sum over $\vec{r}$ is bounded using Lemma 5 in Appendix G.

Next, to evaluate the sum over $\vec{a}$, we parametrize the sites in the set A by their distance to its boundary ∂A. Note that by assumption the interior of A is nonempty, so that $A \neq \partial A$. Roughly speaking, there will be at most $\mathcal{O}(\Phi(A))$ sites whose distances to the boundary ∂A are between ℓ and $\ell + \mu$, for each $\mu = 0, 1, \ldots$, where $\Phi(A)$ is the boundary area of A. Therefore, we have (see Lemma 9 in Appendix G 2 for a rigorous proof)

$$\|\delta_{\mathrm{trunc}}\| \leq 2\eta\lambda_1\Phi(A)\sum_{\mu=0}^{\infty}\frac{1}{(\ell + \mu - \sqrt{D})^{\alpha - D}} \tag{A4}$$

$$\leq 2\eta\lambda_1\lambda_2 2^{\alpha - D}\frac{\Phi(A)}{\ell^{\alpha - D - 1}} = c_{\mathrm{tr}} 2^{\alpha}\frac{\Phi(A)}{\ell^{\alpha - D - 1}}, \tag{A5}$$

for $\ell > 2\sqrt{D}$, where λ_2 is a constant that arises after using Lemma 5 to bound the sum, and the factor $2^{\alpha - D}$ is because we lower bound $\ell - \sqrt{D} \geq \ell/2$ to simplify the expression. The constants are later absorbed into the definition of c_{tr}.

2. Evaluation of $\delta_{\mathrm{overlap}}$

In this section, we show how we bound $\delta_{\mathrm{overlap}}$ from Eq. (11) in the proof of Lemma 1. To estimate $\delta_{\mathrm{overlap}}$, we use the following lemma, which generalizes a similar lemma in Ref. [27] to arbitrary, time-dependent Hamiltonians.

Lemma 3.—Let $\Omega \subset \Lambda$ be a subset of sites. Let $H_{\Omega}(t) = \sum_{i,j\in\Omega} h_{ij}(t)$ be the terms of $H_{\Lambda}(t)$ supported entirely on Ω. Let $O_X(\tau)$ be an observable supported on a subset X at a fixed time τ. We have

$$\|(U_{0,t}^{\Lambda})^{\dagger} O_X(\tau) U_{0,t}^{\Lambda} - (U_{0,t}^{\Omega})^{\dagger} O_X(\tau) U_{0,t}^{\Omega}\| \leq \int_0^t ds \|[(U_{s,t}^{\Omega})^{\dagger} O_X(\tau) U_{s,t}^{\Omega}, H_{\Lambda}(s) - H_{\Omega}(s)]\|, \tag{A6}$$

where $U_{0,t}^{\Lambda} = \mathcal{T}\exp[-i\int_0^t H_{\Lambda}(s)ds]$.

Proof.—To prove the lemma, we shall move into the interaction picture of $H_{\Omega}(t)$ and treat $V(t) \equiv H_{\Lambda}(t) - H_{\Omega}(t)$ as a perturbation. Let $V_I(t) = (U_{0,t}^{\Omega})^{\dagger} V(t) U_{0,t}^{\Omega}$ and $U_I(t) = \mathcal{T}\exp(-i\int_0^t V_I(s)ds)$ be, respectively, the Hamiltonian and the evolution operator in the interaction picture. We have

$$\|(U_{0,t}^{\Lambda})^{\dagger} O_X(\tau) U_{0,t}^{\Lambda} - (U_{0,t}^{\Omega})^{\dagger} O_X(\tau) U_{0,t}^{\Omega}\| = \left\| \int_0^t ds \frac{d}{ds}[U_I^{\dagger}(s)(U_{0,t}^{\Omega})^{\dagger} O_X(\tau) U_{0,t}^{\Omega} U_I(s)] \right\| \tag{A7}$$

$$= \left\| \int_0^t ds U_I^{\dagger}(s)[(U_{0,t}^{\Omega})^{\dagger} O_X(\tau) U_{0,t}^{\Omega}, V_I(s)] U_I(s) \right\| \tag{A8}$$

$$\leq \int_0^t ds \|[(U_{0,t}^{\Omega})^{\dagger} O_X(\tau) U_{0,t}^{\Omega}, V_I(s)]\| \tag{A9}$$

$$\leq \int_0^t ds \|[(U_{s,t}^{\Omega})^{\dagger} O_X(\tau) U_{s,t}^{\Omega}, V(s)]\|. \tag{A10}$$

Thus, Lemma 3 follows. ■

By substituting $\Lambda \to AB$, $\Omega \to B$, $O_X \to H_{B:C}$, $\tau \to t$, and noting that operators supported on disjoint subsets commute, we can show using Lemma 3 that

$$\begin{aligned}\|\delta_{\text{overlap}}\| &= \|(U_{0,t}^{AB})^\dagger H_{B:C}(t) U_{0,t}^{AB} - (U_{0,t}^{B})^\dagger H_{B:C}(t) U_{0,t}^{B}\| \\ &\le \int_0^t ds \|[(U_{s,t}^{B})^\dagger H_{B:C}(t) U_{s,t}^{B}, H_{A:B}(s)]\| \\ &\le \sum_{\vec{a}\in A}\sum_{\vec{b},\vec{b}'\in B}\sum_{\vec{c}\in C}\int_0^t ds \|[(U_{s,t}^{B})^\dagger h_{\vec{b},\vec{c}}(t) \\ &\quad \times U_{s,t}^{B}, h_{\vec{a},\vec{b}'}(s)]\|, \end{aligned} \tag{A11}$$

where the observables are, in general, evaluated at different times $s \le t$. We note that while it is necessary to keep track of s, t for completeness, one should pay more attention to the supports of the operators, as they carry useful information about the locality of the system.

In fact, let us pause for a moment to discuss why the right-hand side of Eq. (A11) should be small when ℓ, the distance between A and C, is large. Whenever the supports of $h_{\vec{a},\vec{b}'}(s)$ and $h_{\vec{b},\vec{c}}(t)$ are far from each other, we can bound their commutator norm using a Lieb-Robinson bound for long-range interactions. We use the bound by Gong *et al.* [5]:

$$\begin{aligned}&\|[(U_{s,t}^{B})^\dagger h_{\vec{b},\vec{c}}(t) U_{s,t}^{B}, h_{\vec{a},\vec{b}'}(s)]\| \\ &\le c e^{v(t-s)} \|h_{\vec{b},\vec{c}}(t)\| \|h_{\vec{a},\vec{b}'}(s)\| \left(\frac{1}{(1-\gamma)^\alpha}\frac{1}{r^\alpha} + \frac{1}{e^{\gamma r}}\right), \end{aligned} \tag{A12}$$

where $r = \text{dist}(\{\vec{a},\vec{b}'\},\{\vec{b},\vec{c}\})$ is the distance between the supports, $\gamma \in (0,1)$ is a constant that can be made arbitrarily close to 1, and c, v are constants that depend only on D.

However, in contrast to short-range interacting systems, here $\vec{b}, \vec{b}'$ run over all possible sites in B, so in principle the distance between the supports of $h_{\vec{a},\vec{b}'}(s)$ and $h_{\vec{b},\vec{c}}(t)$ can be small (Fig. 6). Fortunately, if that is indeed the case, then

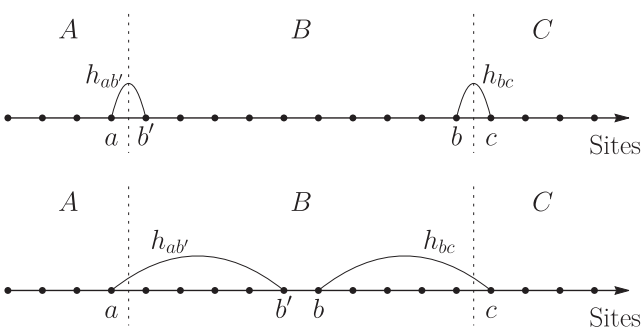

FIG. 6. An illustration of $h_{ab'}$ and h_{bc} in a one-dimensional lattice. For short-range interactions, the sets $\{a,b'\}$ and $\{b,c\}$ are separated by a distance of the same order as the size of B (upper panel). The contributions from these terms to δ_{overlap} are bounded using a Lieb-Robinson bound. However, for long-range interactions, $\{a,b'\}$ and $\{b,c\}$ can be geometrically close to each other (lower panel). In such cases, the norms of $h_{ab'}$ and h_{bc} decay as $|b'-a|^{-\alpha}$ and $|c-b|^{-\alpha}$ and, therefore, their contributions to δ_{overlap} are small.

although the Lieb-Robinson bound is trivial, the assumption that $\|h_{\vec{a},\vec{b}'}\|$ and $\|h_{\vec{b},\vec{c}}\|$ fall off as $\|\vec{b}'-\vec{a}\|^{-\alpha}$ and $\|\vec{c}-\vec{b}\|^{-\alpha}$, respectively, makes the summand in Eq. (A11) small.

Let us now evaluate the sum in Eq. (A11). In the following, we shall consider $\vec{b} \ne \vec{b}'$, since the estimation for the case $\vec{b} = \vec{b}'$ follows a similar but less complicated argument. Using the Lieb-Robinson bound for long-range interactions of Gong *et al.* [5],

$$\begin{aligned}&\int_0^t ds \|[(U_{s,t}^{B})^\dagger h_{\vec{b},\vec{c}}(t) U_{s,t}^{B}, h_{\vec{a},\vec{b}'}(s)]\| \\ &\le c\int_0^t ds \|h_{\vec{b},\vec{c}}(t)\| \|h_{\vec{a},\vec{b}'}(s)\| \left(\frac{e^{v(t-s)}}{((1-\gamma)r)^\alpha} + \frac{e^{v(t-s)}}{e^{\gamma r}}\right) \\ &\le \frac{c}{v}\frac{(e^{vt}-1)}{\|\vec{b}-\vec{c}\|^\alpha \|\vec{a}-\vec{b}'\|^\alpha}\left(\frac{1}{(1-\gamma)^\alpha}\frac{1}{r^\alpha} + \frac{1}{e^{\gamma r}}\right), \end{aligned} \tag{A13}$$

where $\gamma \in (0,1)$ is a constant that can be chosen arbitrarily close to 1, while c, v are finite and bounded constants for all α, and

$$\begin{aligned} r &= \text{dist}(\{\vec{a},\vec{b}'\},\{\vec{b},\vec{c}\}) \\ &= \min\{\|\vec{b}'-\vec{b}\|, \|\vec{b}'-\vec{c}\|, \|\vec{a}-\vec{b}\|, \|\vec{a}-\vec{c}\|\} \end{aligned} \tag{A14}$$

is the distance between the supports of $h_{\vec{b},\vec{c}}(t)$ and $h_{\vec{a},\vec{b}'}(s)$ (see Fig. 6). Since each term of δ_{overlap} contributes a sum of an algebraically decaying as $1/r^\alpha$ and an exponential decaying as $e^{-\gamma r}$ terms, it is convenient to evaluate their contributions separately.

First, let us find the contribution from the algebraically decaying part. It is straightforward to find out their contributions to δ_{overlap} when r takes one of the four allowed values. Depending on which value r takes, we use either Lemma 5 or Lemma 7 in Appendix G to evaluate the sums over $\vec{b}$ and $\vec{b}'$. For example, the contribution from the terms where $r = \|\vec{b}'-\vec{b}\|$ is at most

$$\begin{aligned}&\sum_{\vec{a}\in A}\sum_{\vec{b}\ne\vec{b}'\in B}\sum_{\vec{c}\in C}\frac{c(e^{vt}-1)(\frac{1}{1-\gamma})^\alpha}{v\|\vec{b}-\vec{c}\|^\alpha\|\vec{a}-\vec{b}'\|^\alpha\|\vec{b}'-\vec{b}\|^\alpha} \\ &\le \frac{c}{v}\lambda_3 \sum_{\vec{a}\in A}\sum_{\vec{b}\in B}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)(\frac{2}{1-\gamma})^\alpha}{\|\vec{b}-\vec{c}\|^\alpha\|\vec{a}-\vec{b}\|^\alpha} \end{aligned} \tag{A15}$$

$$\le \frac{c}{v}\lambda_3\lambda_4 \sum_{\vec{a}\in A}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)(\frac{4}{1-\gamma})^\alpha}{\|\vec{a}-\vec{c}\|^\alpha} \tag{A16}$$

$$\le \lambda_5(e^{vt}-1)\left(\frac{8}{1-\gamma}\right)^\alpha \frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{A17}$$

where λ_3, λ_4 are constants that arise after we use Lemma 7 in Appendix G twice to evaluate the sums over $\vec{b}'$ and $\vec{b}$

consecutively, and the sums over $\vec{a}$, $\vec{c}$ have been bounded in the previous section [see Eq. (A2)]. The constant λ_5 absorbs both λ_3, λ_4 and the constants from the sums over $\vec{a}$, $\vec{c}$.

On the other hand, if $r = \|\vec{b}' - \vec{c}\|$, we use Lemma 7 to evaluate the sum over $\vec{b}'$ and Lemma 5 for the sum over $\vec{b}$:

$$\sum_{\vec{a}\in A}\sum_{\vec{b}\neq\vec{b}'\in B}\sum_{\vec{c}\in C}\frac{1}{\|\vec{b}-\vec{c}\|^{\alpha}}\frac{1}{\|\vec{a}-\vec{b}'\|^{\alpha}}\frac{c(e^{vt}-1)(\frac{1}{1-\gamma})^{\alpha}}{v\|\vec{b}'-\vec{c}\|^{\alpha}}$$

$$\leq \frac{c}{v}\lambda_6\sum_{\vec{a}\in A}\sum_{\vec{b}\in B}\sum_{\vec{c}\in C}\frac{1}{\|\vec{b}-\vec{c}\|^{\alpha}}\frac{(e^{vt}-1)(\frac{2}{1-\gamma})^{\alpha}}{\|\vec{a}-\vec{c}\|^{\alpha}} \tag{A18}$$

$$\leq \frac{c}{v}\lambda_6\lambda_7\sum_{\vec{a}\in A}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)(\frac{2}{1-\gamma})^{\alpha}}{\|\vec{a}-\vec{c}\|^{\alpha}} \tag{A19}$$

$$\leq \lambda_8(e^{vt}-1)\left(\frac{4}{1-\gamma}\right)^{\alpha}\frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{A20}$$

where λ_6, λ_7 come from the uses of Lemma 7 and Lemma 5, respectively. The constant λ_8 absorbs both λ_6, λ_7 and the constants from the sums over $\vec{a}$, $\vec{c}$. Repeating for the other values of r, we find that the contribution from the algebraically decaying terms in Eq. (A13) to δ_{overlap} is at most

$$\lambda_9(e^{vt}-1)\left(\frac{8}{1-\gamma}\right)^{\alpha}\frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{A21}$$

for some constant λ_9.

Next, let us find the contribution from the exponentially decaying term in Eq. (A13). If $r = \|\vec{b}' - \vec{b}\|$, we have

$$\sum_{\vec{a}\in A}\sum_{\vec{b}\neq\vec{b}'\in B}\sum_{\vec{c}\in C}\frac{c(e^{vt}-1)}{v\|\vec{b}-\vec{c}\|^{\alpha}\|\vec{a}-\vec{b}'\|^{\alpha}e^{\gamma\|\vec{b}'-\vec{b}\|}}$$

$$\leq \frac{c}{v}\lambda_{10}\sum_{\vec{a}\in A}\sum_{\vec{b}\in B}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)}{\|\vec{b}-\vec{c}\|^{\alpha}}\left(\frac{(\frac{4}{1-\gamma})^{\alpha}}{\|\vec{a}-\vec{b}\|^{\alpha}}+\frac{\|\vec{a}-\vec{b}\|^{D-1}}{e^{\gamma\|\vec{a}-\vec{b}\|}}\right)$$

$$\leq \frac{c}{v}\lambda_{10}\lambda_{11}\sum_{\vec{a}\in A}\sum_{\vec{c}\in C}(e^{vt}-1)\left(\frac{(\frac{8}{1-\gamma})^{\alpha}}{\|\vec{a}-\vec{c}\|^{\alpha}}+\frac{\|\vec{a}-\vec{c}\|^{2D-2}}{e^{\gamma\|\vec{a}-\vec{c}\|}}\right)$$

$$\leq \lambda_{12}(e^{vt}-1)\Phi(A)\left(\frac{(\frac{16}{1-\gamma})^{\alpha}}{\ell^{\alpha-D-1}}+\frac{\ell^{3D-3}}{e^{\gamma\ell}}\right), \tag{A22}$$

where we have applied Lemma 8 in Appendix G to obtain the first inequality, Lemma 8 twice again and Lemma 7 to get the second inequality, and then Lemma 5 and Lemma 6 for the sums over $\vec{a}$, $\vec{c}$ similarly to Appendix A 1. The constants λ_{10}, λ_{11} arise from the applications of the lemmas and are absorbed into a constant λ_{12}. We note that the constant γ in the last three lines are different from the one in the first line (see Lemma 8 for details). However, they both are constants that can be chosen arbitrarily between 0 and 1. Therefore, we denote them by the same constant γ for convenience.

Repeating the argument for other choices of r in Eq. (A14), we find that the contribution from the exponentially decaying terms to δ_{overlap} is still at most the right-hand side of Eq. (A22).

Combining Eqs. (A21) and (A22), we have

$$\|\delta_{\text{overlap}}\| \leq \lambda_{13}(e^{vt}-1)\Phi(A)\left(\frac{(\frac{16}{1-\gamma})^{\alpha}}{\ell^{\alpha-D-1}}+\frac{\ell^{3D-3}}{e^{\gamma\ell}}\right), \tag{A23}$$

for a constant λ_{13}. Since $\ell^{D-1} \leq [(D-1)!/\varepsilon^{D-1}]e^{\varepsilon\ell}$ for any arbitrary small positive constant ε, we can upper bound

$$\|\delta_{\text{overlap}}\| \leq c_{\text{ov}}(e^{vt}-1)\Phi(A)\left(\frac{(\frac{16}{1-\gamma})^{\alpha}}{\ell^{\alpha-D-1}}+\frac{1}{e^{\gamma\ell}}\right), \tag{A24}$$

where we have absorbed ε into the definition of γ and c_{ov}. This completes the estimation of δ_{overlap}.

APPENDIX B: ERROR PROPAGATION FROM GENERATING FUNCTION

In this Appendix, we reproduce a lemma in Ref. [27] which shows how the error in approximating the generating function $\mathcal{G}_W$ propagates to an error of the unitary W_t in Eq. (6). Suppose we approximate $\mathcal{G}_W$ by $\mathcal{G}'_W$ such that

$$\|\mathcal{G}_W - \mathcal{G}'_W\| \leq f(t)\delta, \tag{B1}$$

for some function of time $f(t)$, and δ is time independent. We shall prove that the unitary W'_t generated by $\mathcal{G}'_W$ approximates W_t with error

$$\|W'_t - W_t\| \leq \delta\int_0^t ds f(s). \tag{B2}$$

Proof.—By simple differentiation, we have

$$\|W_t^\dagger W'_t - \mathbb{I}\| = \left\|\int_0^t ds\frac{d}{ds}(W_s^\dagger W'_s)\right\| \tag{B3}$$

$$= \left\|\int_0^t ds W_s^\dagger(G_W - G'_W)W'_s\right\| \tag{B4}$$

$$\leq \int_0^t ds\|W_s^\dagger(G_W - G'_W)W'_s\| \tag{B5}$$

$$= \int_0^t ds\|G_W - G'_W\| \tag{B6}$$

$$\leq \delta\int_0^t ds f(s). \tag{B7}$$

■

APPENDIX C: PROOF OF THE LIEB-ROBINSON BOUND FOR LONG-RANGE INTERACTIONS

We present a more detailed proof of Theorem 1 in this Appendix. The key ingredient in the proof of Theorem 1 is the following lemma.

Lemma 4.—Denote by $\mathcal{B}_r = \{\vec{\imath} \in \Lambda : \|\vec{\imath}\| \le r\}$ a D ball of radius r centered around the origin. Let O_X be an observable supported on $X = \mathcal{B}_{r_0}$ with r_0 being finite. For each $U^{\Lambda}_{0,T}$ and a positive integer M, there exists a unitary $\tilde{U}$ supported on a D ball $\mathcal{B}_r$ with $r = r_0 + M\ell$, such that

$$\|(U^{\Lambda}_{0,T})^{\dagger} O_X U^{\Lambda}_{0,T} - \tilde{U}^{\dagger} O_X \tilde{U}\| \le b_1 M e^{vt} (r-\ell)^{D-1} \xi_{\alpha}(\ell), \tag{C1}$$

where b_1 is a constant, $t = T/M$, and $\ell \in (0, R)$ is a free parameter.

Proof.—We shall prove the lemma by constructing the unitary $\tilde{U}$. In addition to $\mathcal{B}_r$ above, we define

$$\mathcal{S}_r = \mathcal{B}_{r+\ell} \backslash \mathcal{B}_r \tag{C2}$$

to be a shell consisting of sites between r and $r + l$ away from the origin (Fig. 7).

We divide $[0, T]$ into M equal time intervals, namely, $[(M-k-1)t, (M-k)t]$ for $k = 0, \ldots, M-1$, where $t = T/M$. The unitary $U^{\Lambda}_{0,T}$ then naturally decomposes into a product of unitaries $U^{\Lambda}_k \equiv U^{\Lambda}_{(M-k-1)t,(M-k)t}$:

$$U^{\Lambda}_{0,T} = U^{\Lambda}_0 U^{\Lambda}_1 \ldots U^{\Lambda}_{M-1}. \tag{C3}$$

We now use Lemma 1 to further decompose each U^{Λ}_k into evolutions of subsystems. We start with $k = 0$ and use

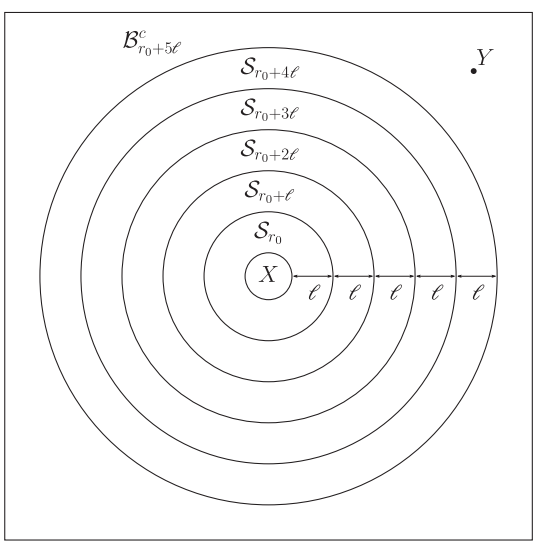

FIG. 7. An example of the subset $X = \mathcal{B}_{r_0}$ and five shells $\mathcal{S}_r$ for $r = r_0, r_0 + \ell, \ldots, r_0 + 4\ell$. The operator O_Y is supported on Y, which lies on $\mathcal{B}^c_{r_0+5\ell}$, the complement of the ball $\mathcal{B}_{r_0+5\ell}$.

Lemma 1 with $A \to X = \mathcal{B}_{r_0}, B \to \mathcal{S}_{r_0}$, and $C \to \mathcal{B}^c_{r_0+\ell}$ (Fig. 3) to decompose $(U^{\Lambda}_0)^{\dagger}$ (instead of U^{Λ}_0):

$$\begin{aligned} &\|(U^{\Lambda}_0)^{\dagger} - (U_0^{\mathcal{B}_{r_0+\ell}})^{\dagger} U_0^{\mathcal{S}_{r_0}} (U_0^{\mathcal{B}^c_{r_0}})^{\dagger}\| \\ &\quad \le c_0 e^{vt} \Phi(\mathcal{B}_{r_0}) \xi_{\alpha}(\ell), \end{aligned} \tag{C4}$$

where again $\mathcal{B}^c_{r_0}$ denotes the complement subset $\mathbb{R}^D \backslash \mathcal{B}_{r_0}$, and $\Phi(\mathcal{B}_{r_0})$ is the boundary area of $\mathcal{B}_{r_0}$. This choice of decomposition allows us to eliminate the contribution to the evolution from the terms of the Hamiltonian that commute with X, i.e., those supported entirely on $\mathcal{B}^c_{r_0}$. Explicitly, we have

$$\begin{aligned} &(U^{\Lambda}_0)^{\dagger} O_X U^{\Lambda}_0 \\ &\quad \approx \left(U_0^{\mathcal{B}_{r_0+\ell}}\right)^{\dagger} U_0^{\mathcal{S}_{r_0}} \left(U_0^{\mathcal{B}^c_{r_0}}\right)^{\dagger} O_X U_0^{\mathcal{B}^c_{r_0}} \left(U_0^{\mathcal{S}_{r_0}}\right)^{\dagger} U_0^{\mathcal{B}_{r_0+\ell}} \qquad \text{(C5)} \\ &\quad = \left(U_0^{\mathcal{B}_{r_0+\ell}}\right)^{\dagger} O_X U_0^{\mathcal{B}_{r_0+\ell}} = \tilde{U}_0^{\dagger} O_X \tilde{U}_0, \qquad \text{(C6)} \end{aligned}$$

where $\tilde{U}_0 \equiv U_0^{\mathcal{B}_{r_0+\ell}}$ is supported entirely on $\mathcal{B}_{r_0+\ell}$.

If we repeat the above argument for U^{Λ}_1 but with O_X replaced by $\tilde{U}_0^{\dagger} O_X \tilde{U}_0$, we can approximate

$$(U^{\Lambda}_1)^{\dagger} \tilde{U}_0^{\dagger} O_X \tilde{U}_0 U^{\Lambda}_1 \approx \tilde{U}_1^{\dagger} \tilde{U}_0^{\dagger} O_X \tilde{U}_0 \tilde{U}_1, \tag{C7}$$

for some $\tilde{U}_1$ supported entirely on $\mathcal{B}_{r_0+2\ell}$. The error of this approximation is at most $c_0 e^{vt} \Phi(\mathcal{B}_{r_0+\ell}) \xi_{\alpha}(\ell)$.

By induction to all $k = 2, \ldots, M-1$, we can construct $\tilde{U} = \tilde{U}_0 \tilde{U}_1 \ldots \tilde{U}_{T-1}$ such that

$$(U^{\Lambda}_{0,T})^{\dagger} O_X U^{\Lambda}_{0,T} \approx \tilde{U}^{\dagger} O_X \tilde{U}, \tag{C8}$$

where the overall error is at most

$$\begin{aligned} &\sum_{k=0}^{M-1} c_0 e^{vt} \Phi(\mathcal{B}_{r_0+k\ell}) \xi_{\alpha}(\ell) \\ &\quad \le M c_0 e^{vt} \Phi(\mathcal{B}_{r_0+(M-1)\ell}) \xi_{\alpha}(\ell) \qquad \text{(C9)} \\ &\quad \le \underbrace{c_0 \frac{2\pi^{D/2}}{\Gamma(\frac{D}{2})}}_{\equiv b_1} M e^{vt} [r_0 + (M-1)\ell]^{D-1} \xi_{\alpha}(\ell), \qquad \text{(C10)} \end{aligned}$$

and where we have replaced the surface area $\Phi(\mathcal{B}_r)$ of a D ball $\mathcal{B}_r$ by $\{2\pi^{D/2}/[\Gamma(D/2)]\} r^{D-1}$ and M by T/t. Also by induction, the unitary $\tilde{U}$ is supported entirely on $\mathcal{B}_{r_0+M\ell}$. Therefore, the lemma follows. ■

We are now ready to prove our Lieb-Robinson bound in Theorem 1. Without loss of generality, we assume the origin is in X. Since $\|X\| = \mathcal{O}(1)$, there exists $r_0 = \mathcal{O}(1)$ such that X is a subset of $\mathcal{B}_{r_0}$. By Lemma 4, there exists a unitary $\tilde{U}$ supported entirely on a D ball $\mathcal{B}_r$ with $r = r_0 + M\ell$ such that

$$\varepsilon = \|(U_{0,T}^{\Lambda})^{\dagger} O_X U_{0,T}^{\Lambda} - \tilde{U}^{\dagger} O_X \tilde{U}\| \leq b_1 M e^{vt} (r-\ell)^{D-1} \xi_\alpha(\ell). \tag{C11}$$

If we choose the number of time slices (M) and the block size (ℓ) such that $r \leq R + r_0$, the set Y will lie outside the support $\mathcal{B}_r$ of $\tilde{U}^{\dagger} O_X \tilde{U}$, and therefore $\tilde{U}^{\dagger} O_X \tilde{U}$ will commute with O_Y. Note that for a fixed value of M, the error should decrease with a larger value of ℓ. Therefore, to prove the strongest bound, we should choose ℓ as large as possible, i.e., $\ell = R/M$, and hence $M = R/\ell$. Substituting the value of M and $t = T/M$ into Eq. (C11), we obtain the error bound in terms of ℓ alone:

$$\varepsilon \leq b_2 \frac{R}{\ell} (e^{vT\ell/R} - 1)(1 + R - \ell)^{D-1} \xi_\alpha(\ell), \tag{C12}$$

where $b_2 = b_1 r_0^{D-1}$ is a finite constant. We note that the above bound is valid for all values of $\ell \leq R$. The tightest bound can therefore be obtained by choosing a value for ℓ that minimizes the above expression. Our intuition and numerical evidence suggest that this happens when $\ell \sim R\alpha/(vT)$, so in the below analysis, we aim to choose ℓ as close to this value as possible.

To proceed, we consider two regimes of time T, when $vT \geq \alpha$ and when $vT < \alpha$. In the former regime, we choose $\ell = R\alpha/(vT) \leq R$ and substitute into Eq. (C12) to get

$$\begin{aligned} \varepsilon &\leq b_2 \frac{vT}{\alpha} (e^{\alpha} - 1) \left[1 + R\left(1 - \frac{\alpha}{vT}\right)\right]^{D-1} \xi_\alpha\left(\frac{R\alpha}{vT}\right) \\ &\leq \underbrace{b_2 v 2^{D-1}}_{\equiv c_{\mathrm{lr}}} T R^{D-1} \xi_\alpha\left(\frac{R\alpha}{vT}\right), \end{aligned} \tag{C13}$$

where we have used $e^{\alpha} - 1/\alpha \leq 1$ for all $\alpha \geq 1$, $1 - \alpha/(vT) \leq 1$, and $1 + R \leq 2R$. In particular, if $R\alpha/(vT) > x_0$, where x_0 is the larger solution of $x^{\alpha-D-1} = e^{\gamma x}$, the algebraically decaying term in ξ_α dominates the exponentially decaying one, and therefore

$$\xi_\alpha\left(\frac{R\alpha}{vT}\right) = \left(\frac{16}{1-\gamma}\right)^{\alpha} \frac{1}{(\frac{R\alpha}{vT})^{\alpha-D-1}} + e^{-\gamma(R\alpha/vT)} \tag{C14}$$

$$\leq 2\left(\frac{16}{1-\gamma}\right)^{\alpha} \frac{1}{(\frac{R\alpha}{vT})^{\alpha-D-1}} \tag{C15}$$

$$= 2\left(\frac{16}{1-\gamma}\right)^{\alpha} \left(\frac{v}{\alpha}\right)^{\alpha-D-1} \left(\frac{T}{R}\right)^{\alpha-D-1}. \tag{C16}$$

Combining Eq. (C13) and (C16), we obtain a bound on the commutator norm:

$$\mathcal{C}(T,R) \leq \varepsilon \leq c_{\mathrm{lr},\alpha} \frac{T^{\alpha-D}}{R^{\alpha-2D}}, \tag{C17}$$

where

$$c_{\mathrm{lr},\alpha} := 2c_{\mathrm{lr}} \left(\frac{16}{1-\gamma}\right)^{\alpha} \left(\frac{v}{\alpha}\right)^{\alpha-D-1}. \tag{C18}$$

The light cone implied by the bound is

$$T \gtrsim R^{(\alpha-2D)/(\alpha-D)}. \tag{C19}$$

In the limit $\alpha \to \infty$, the exponent of the light cone converges to one at a rate given by

$$\mu = \lim_{\alpha\to\infty} \frac{|\frac{\alpha+1-2D}{\alpha+1-D} - 1|}{|\frac{\alpha-2D}{\alpha-D} - 1|} = 1. \tag{C20}$$

On the other hand, if $vT < \alpha$, we simply choose $\ell = R$. Equation (C12) then becomes

$$\mathcal{C}(T,R) \leq \varepsilon \leq b_2 (e^{vT} - 1) \xi_\alpha(R). \tag{C21}$$

Therefore, we arrive at the Lieb-Robinson bound in Theorem 1 with $\tilde{c}_{\mathrm{lr}} = b_2$.

APPENDIX D: PROOF OF LEMMA 2 IN HIGHER DIMENSIONS

In this Appendix, we discuss the construction of the circuit in Lemma 2 that generalizes the lemma to higher dimensions. Similar to the $D = 1$ case, we first break the unitary into $\mathcal{O}(T)$ unitaries $\exp(-iHt)$ for some $t = \mathcal{O}(1)$. We then use an algorithm consisting of D steps to break the simulation of $\exp(-iHt)$ into simulations of Hamiltonians on smaller hypercubes of size at most 2ℓ. In the first of the D steps, we cut the D-dimensional lattice into L/ℓ layers, each with the same thickness ℓ, a parameter to be chosen later. In this step, the cross section of the cut is L^{D-1}. Therefore, by Lemma 1, each time a new layer is generated, we accumulate an error of

$$\mathcal{O}\left(\frac{L^{D-1}}{\ell^{\alpha-D-1}}\right).$$

For TL/ℓ layers of the first step, the accumulated error will be $\varepsilon^{(1)} = \mathcal{O}(TL^D/\ell^{\alpha-D})$.

Next, for each of the $\mathcal{O}(TL/\ell)$ layers of $D-1$ dimensions, we break them again into L/ℓ layers of $D-2$ dimensions. Using Lemma 1 with a cross section L^{D-2}, we find the error of the second step,

$$\varepsilon^{(2)} = \frac{TL}{\ell}\frac{L}{\ell}\mathcal{O}\left(\frac{L^{D-2}}{\ell^{\alpha-(D-1)-1}}\right) = \mathcal{O}\left(\frac{TL^D}{\ell^{\alpha-D+2}}\right), \tag{D1}$$

which decreases with ℓ faster than the error of the first step.

More explicitly, in the kth of the D steps, the error is $\varepsilon^{(k)} = \mathcal{O}(L^D/\ell^{\alpha-D-2k})$, which is dominated by the error in

the first step for all $k > 1$. Therefore, the error of cutting the D-dimensional lattice of size L into L^D/ℓ^D subsystems is still $\mathcal{O}(TL^D/\ell^{\alpha-D})$. To meet a fixed total error ε, we need to choose $\ell \propto (TL^D/\varepsilon)^{1/(\alpha-D)}$. The geometrical constraint $\ell < L$ requires $\alpha > 2D$. Finally, simulating each of the $\mathcal{O}(TL^D/\ell^D)$ subsystems using the LCU algorithm up to $\varepsilon\ell^D/(TL^D)$ accuracy requires $\mathcal{O}(\ell^{3D}\log(TL^D/\varepsilon\ell^D))$ quantum gates. Therefore, the overall gate complexity of the algorithm is

$$G_D = \mathcal{O}\left(\frac{(Tn)^{1+2D/(\alpha-D)}}{\varepsilon^{2D/(\alpha-D)}}\log\frac{Tn}{\varepsilon}\right). \tag{D2}$$

APPENDIX E: ESTIMATION OF THE ACTUAL GATE COUNT

In this Appendix, we describe how we estimate the actual gate count of the HHKL algorithm and the QSP algorithm in simulating one-dimensional power-law systems.

The direct implementation of the QSP algorithm requires computing a sequence of rotation angles on a classical computer, which is prohibitive for large-size Hamiltonian simulation. Instead, we use a suboptimal approach described in Ref. [49]. To simulate $H = \sum_{j=1}^{L}\beta_j H_j$ for time t and accuracy ε, where L is the number of terms in the Hamiltonian, $\beta_j \geq 0$ and H_j are both unitary and Hermitian, we divide the entire evolution into r segments. We choose r sufficiently large so that each segment is short enough for the classical preprocessing. Specifically, we choose

$$r = \left\lceil \frac{\sum_j \beta_j t}{\tau_{\max}} \right\rceil \tag{E1}$$

and $\tau_{\max} = 1000$ [55]. Within each segment, we choose q to be the smallest positive integer satisfying

$$\frac{4(\sum_j \beta_j t/r)^q}{2^q q!} \leq \frac{\varepsilon}{8r}, \tag{E2}$$

so that the overall error is at most ε. This gives $M = 2(q-1)$ phased iterates within each segment [49].

The number of elementary operations of each phased iterate is $\log(L) + 4L + 8L$. Here, the first term corresponds to the reflection along an L-dimensional state $|0\rangle$, the second term costs the preparation or unpreparation of an L-dimensional state, and the third term is the cost of selecting L two-body operators. We thus estimate the gate complexity of the QSP algorithm as $[\log(L) + 12L]rM$.

Next, in order to determine the gate count of the HHKL algorithm, we need an estimate for the error of the unitary decomposition in Lemma 1. Recall that for $D = 1$, the error given by our analysis is $b/\ell^{\alpha-2}$, where b is a constant that can be estimated numerically by computing the actual error for small values of ℓ and extrapolating for larger ℓ.

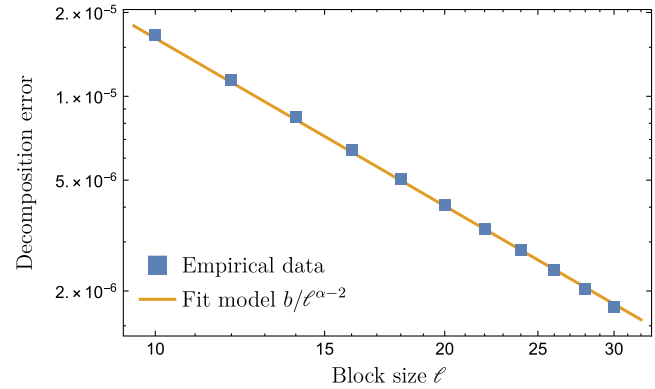

FIG. 8. The empirical error of the unitary decomposition in Lemma 1, computed for the single-excitation one-dimensional Heisenberg chain ($\alpha = 4$) in Eq. (25) at different values of ℓ. The system size is fixed at $n = 300$ and the evolution time at $t = 0.01$. We fit the data (blue square) to the theoretical model $b/\ell^{\alpha-2}$ and obtain $b = 1.62 \times 10^{-3}$.

Since simulating the evolution of a generic system is classically intractable even for a moderate system size, we study only the one-dimensional Heisenberg model given in Eq. (25) and restrict our calculation to the single-excitation subspace. In Fig. 8, we plot the error of the unitary decomposition in Lemma 1 at several different values of ℓ (for system size $n = 300$ and evolution time $t = 0.01$). The scaling of the error agrees well with our prediction. By fitting the data to $b/\ell^{\alpha-2}$, we obtain an estimate $b = 1.62 \times 10^{-3}$.

Recall that there are T/t time slices in the HHKL algorithm. In each time slice, there are n/ℓ blocks of size ℓ and $2n/(2\ell)$ blocks of size 2ℓ. To meet the total error at most ε, we need to choose [see also Eq. (20)]

$$\ell = \left(\frac{T}{t}\frac{2nb}{\varepsilon}\right)^{1/(\alpha-1)}. \tag{E3}$$

By multiplying the number of blocks by the gate count for using QSP to simulate a single block, we arrive at the total gate count presented in Fig. 5.

APPENDIX F: NUMERICAL PERFORMANCE OF THE PRODUCT FORMULA

This Appendix includes the numerical performance of the fourth-order product formula (PF4) used to simulate the evolution of the system given in Eq. (25) for time $T = n$. We plot this numerical performance as well as the theoretical estimates for the gate counts of the PF4, QSP, LCU, and HHKL algorithms in Fig. 9.

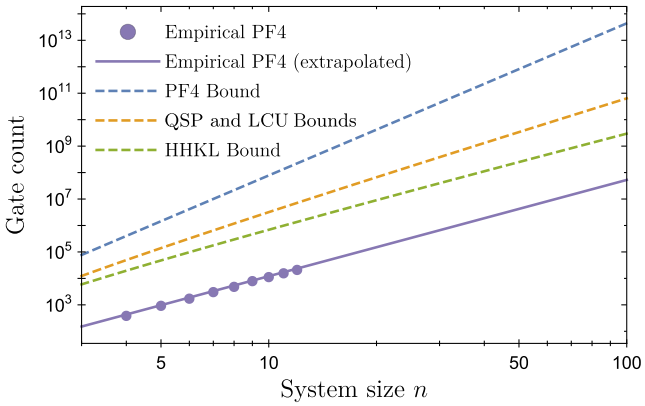

FIG. 9. The empirical gate count of PF4 (purple dots) from $n = 4$ to $n = 12$, extrapolated to larger system sizes (solid, purple), for simulating the dynamics of the Hamiltonian in Eq. (25) for time $T = n$ at a fixed error tolerance. The error bars are smaller than the size of the markers and hence not visible in the plot. Also shown in dashed lines are the slopes of the gate counts of several advanced algorithms for comparison. These slopes represent the scaling of the gate counts as functions of n. Their y-intercepts, which represent a constant multiplicative factor, should be ignored.

APPENDIX G: MATHEMATICAL TOOLS

This Appendix contains a collection of mathematical results omitted from the previous sections. In Appendix G 1, we present the upper bounds on standard sums we use in the proof of Lemma 1 in Appendix A. In Appendix G 2, we show how we estimate the sum over the convex set A in Eq. (A3) by parametrizing the elements of the set by their distance to the boundary of A. We also note that we use the same notation "λ" for constants that appear in different lemmas.

1. Standard sums

Here, we present upper bounds on a a few standard sums used in the previous sections. Specifically, we use Lemma 5 to bound Eq. (A2), Eq. (A18), Lemma 6 to bound Eq. (A15), Eq. (A22), Lemma 7 to bound Eq. (18), Eq. (A22), and Lemma 8 to bound Eq. (A22).

Lemma 5.—Let Λ be a D-dimensional lattice and $\vec{r}$ be the coordinates of sites in Λ. For $\alpha > D + 1$ and $R > \sqrt{D}$, there exists a constant λ that may depend on D but not on R, α such that

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{1}{\|\vec{r}\|^{\alpha}} \leq \frac{\lambda}{(R-\sqrt{D})^{\alpha-D}}. \tag{G1}$$

In particular, it implies that the sum $\sum_{\vec{r}\in\Lambda}$ converges for all $\alpha > D$.

Proof.—The proof of this bound is straightforward. For simplicity, we first assume none of the coordinates of $\vec{r}$ is zero. Since $1/x^{\alpha}$ is a decreasing function of x for all $\alpha > 0$, we can always bound the sum over such $\vec{r}$ by an integral

$$\begin{aligned}\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}}{}' \frac{1}{\|\vec{r}\|^{\alpha}} &\leq \int_{\|\vec{r}\|\geq R-\sqrt{D}} \frac{d^D\vec{r}}{\|\vec{r}\|^{\alpha}} \\ &= \frac{2\pi^{D/2}}{\Gamma(\frac{D}{2})}\int_{R-\sqrt{D}}^{\infty}\frac{dr}{r^{\alpha-D+1}} \leq \frac{g(D)}{(R-\sqrt{D})^{\alpha-D}},\end{aligned} \tag{G2}$$

where $\sum'$ denotes the sum over $\vec{r}$ with no zero coordinate and $g(D) \equiv 2\pi^{D/2}/\Gamma(D/2)$.

Next, consider $\vec{r}$ with exactly one zero coordinate. These sites lie on D hyperplanes, each of dimension $(D-1)$. Therefore, the contribution from them can be evaluated using the above integral with $D \to D-1$:

$$\frac{Dg(D-1)}{(R-\sqrt{D-1})^{\alpha-D+1}} < \frac{Dg(D-1)}{(R-\sqrt{D})^{\alpha-D}}. \tag{G3}$$

By repeating this argument for the sums over $\vec{r}$ with different number of zero coordinates, we arrive at

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{1}{\|\vec{r}\|^{\alpha}} \leq \frac{\lambda}{(R-\sqrt{D})^{\alpha-D}}, \tag{G4}$$

where $\lambda = \sum_{d=0}^{D}\binom{D}{d}g(D-d)$ is a constant independent of R. ■

Lemma 6.—Let Λ be a D-dimensional lattice and $\vec{r}$ be the coordinates of sites in Λ. For all $R > 0$, there exists a constant λ that may depend on β, D but not on R such that

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} \leq \frac{\lambda R^{\beta+D-1}}{e^{R}}, \tag{G5}$$

where β is a positive constant. In particular, it also implies that the sum $\sum_{\vec{r}\in\Lambda}$ converges.

Proof.—The proof of this lemma follows the same idea as of Lemma 5. However, note that the function $x^{\beta}e^{-x}$ is a decreasing function of x only when $x \geq x_0$ for some x_0 that depends only on β. Therefore, if $R \geq x_0$, we follow the exact same lines as in the proof of Lemma 5. For example, if none of the coordinates of $\vec{r}$ is zero, we can bound

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}}{}' \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} \leq \int_{\|\vec{r}\|\geq R-\sqrt{D}} \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} d^D\vec{r}$$

$$= \frac{2\pi^{D/2}}{\Gamma(\frac{D}{2})}\int_{R-\sqrt{D}}^{\infty}\frac{r^{\beta+D-1}dr}{e^{r}} \tag{G6}$$

$$\leq \lambda_1 \frac{(R-\sqrt{D})^{\beta+D-1}}{e^{R-\sqrt{D}}} \leq \lambda_2 \frac{R^{\beta+D-1}}{e^{R}}, \tag{G7}$$

for some constants λ_1, λ_2 that depend only on β, D.

On the other hand, if $R < x_0$, we consider

$$\lambda = \max\left\{\frac{e^R}{R^{\beta+D-1}}, \sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{\|\vec{r}\|^\beta}{e^{\|\vec{r}\|}}\right\}. \tag{G8}$$

The lemma should follow if we can argue that λ can be chosen independently of R. Indeed, since $1 \leq R < x_0$ and from the previous calculation, we know that the sum over $\vec{r}$ converges to a constant that depends only on β, D. This concludes the proof of Lemma 6. ■

Lemma 7.—Let $\vec{a}$, $\vec{b}$, $\vec{c}$ be three distinct sites in a D-dimensional lattice Λ. For all $\alpha > D$,

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{1}{\|\vec{b}-\vec{c}\|^\alpha} \leq \frac{\lambda 2^\alpha}{\|\vec{a}-\vec{c}\|^\alpha}, \tag{G9}$$

where λ is a constant independent of $\vec{a}, \vec{c}, \alpha$.

Proof.—A proof of the lemma is presented in Ref. [4]. ■

Lemma 8.—Let $\vec{a}$, $\vec{b}$, $\vec{c}$ be three distinct sites in a D-dimensional lattice Λ. For all $\alpha > D$, $\gamma \in (0,1)$, and positive integers $\beta \in \mathbb{Z}^+$, there exists a constant $\gamma' \in (0,1)$ such that

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}-\vec{c}\|^\beta}{e^{\gamma\|\vec{b}-\vec{c}\|}} \leq \frac{\lambda(\frac{4}{1-\gamma})^\alpha}{\|\vec{a}-\vec{c}\|^\alpha} + \frac{\lambda'\|\vec{a}-\vec{c}\|^{\beta+D-1}}{e^{\gamma'\|\vec{a}-\vec{c}\|}}, \tag{G10}$$

where λ, λ' are constants that may depend on β, D, but not on $\vec{a}, \vec{c}, \alpha$.

Proof.—Without loss of generality, assume $\vec{c} = 0$. Let $\ell = \|\vec{c}-\vec{a}\| = \|\vec{a}\|$ be the distance between $\vec{c}$ and $\vec{a}$. We need to prove

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}\|^\beta}{e^{\gamma\|\vec{b}\|}} \leq \frac{\lambda(\frac{4}{1-\gamma})^\alpha}{\ell^\alpha} + \frac{\lambda'\ell^{\beta+D-1}}{e^{\gamma'\ell}}. \tag{G11}$$

Let $\mathcal{B}_{\mu\ell}$ be a D ball of radius $\mu\ell$ centered around $\vec{c}$ for some arbitrary constant $\mu \in (0,1)$. We shall divide the sum over $\vec{b}$ into two regimes, corresponding to $\vec{b}$ inside and outside $\mathcal{B}_{\mu\ell}$.

In the first regime where $\vec{b}$ is inside $\mathcal{B}_{\mu\ell}$, we can show using the triangle inequality that $\|\vec{a}-\vec{b}\| \geq (1-\mu)\ell$. Therefore, the sum over these $\vec{b}$ can be bounded by

$$\frac{1}{[(1-\mu)\ell]^\alpha} \sum_{\vec{b}\in\mathcal{B}_{\mu\ell}} \frac{\|\vec{b}\|^\beta}{e^{\|\vec{b}\|}} \leq \frac{\lambda(\frac{2}{1-\mu})^\alpha}{\ell^\alpha}, \tag{G12}$$

where we have used the fact that $\sum_{\vec{b}\in\mathcal{B}_{\mu\ell}}[(\|\vec{b}\|^\beta)/e^{\|\vec{b}\|}]$ converges and is bounded by a constant λ which may depend only on D, β.

In the second regime, we bound $\|\vec{a}-\vec{b}\| \geq 1$ to obtain

$$\sum_{\vec{b}\notin\mathcal{B}_{\mu\ell}} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}\|^\beta}{e^{\gamma\|\vec{b}\|}} \leq \sum_{\vec{b}\notin\mathcal{B}_{\mu\ell}} \frac{\|\vec{b}\|^\beta}{e^{\gamma\|\vec{b}\|}} \leq \lambda' \frac{\ell^{\beta+D-1}}{2^\alpha e^{\gamma\mu\ell}}, \tag{G13}$$

where the last sum is bounded using Lemma 6 and noting that $\mu < 1$.

Combining Eqs. (G12) and (G13), we arrive at a bound

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}\|^\beta}{e^{\gamma\|\vec{b}\|}} \leq \frac{\lambda(\frac{2}{1-\mu})^\alpha}{\ell^\alpha} + \frac{\lambda'\ell^{\beta+D-1}}{e^{\gamma\mu\ell}}. \tag{G14}$$

Let $\gamma' = \gamma\mu$ and take $\mu \leq [1/(2-\gamma)]$; it is straightforward to show that $[2/(1-\mu)] \leq [4/(1-\gamma')]$, and therefore,

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}\|^\beta}{e^{\gamma\|\vec{b}\|}} \leq \frac{\lambda(\frac{4}{1-\gamma'})^\alpha}{\ell^\alpha} + \frac{\lambda'\ell^{\beta+D-1}}{e^{\gamma'\ell}}. \tag{G15}$$

Note that if we choose $\mu = [1/(2-\gamma)]$, then $\gamma' = [\gamma/(2-\gamma)]$ takes on a value between 0 and 1, which can be arbitrarily close to 1. ■

2. Parametrizing a convex set

In this section, we show how we evaluate the sum over $\vec{a}$ in Eq. (A3). First, we parametrize a convex set by the distance to its boundary. The following lemma simplifies a sum over every site in a convex set to a sum over the above distance, multiplied by the boundary area of the set.

Lemma 9.—Let $A \subset \mathbb{R}^D$ be a compact and convex set in $\mathbb{R}^D$ with nonempty interior. Let $C \subset \mathbb{R}^D$ be another subset disjoint from A, and let $\ell = \mathrm{dist}(A, C)$ be the smallest distance between elements of the two sets. Furthermore, we denote by $\ell_{\vec{a}} = \mathrm{dist}(\vec{a}, C)$ the minimal distance from a given lattice site $\vec{a}$ in A to C. For a decreasing function $f: \mathbb{R} \to \mathbb{R}$, we shall have

$$\sum_{\vec{a}\in A\cap\Lambda} f(\ell_{\vec{a}}) \leq 2\eta\Phi(A) \sum_{\mu=0}^{\infty} f(\ell+\mu), \tag{G16}$$

where η is a constant that may depend only on D and $\Phi(A)$ is the boundary area of A.

Proof.—Let us divide the set $A \in \mathbb{R}^D$ into disjoint subsets $S_\mu = \{\vec{a}\in A : \mu \leq \mathrm{dist}(\vec{a}, \partial A) \leq \mu+1\}$ for $\mu = 0, 1, \ldots$. Note that the assumption that the interior of A is nonempty implies that $\mathrm{dist}(\vec{a}, \partial A)$ is not uniformly zero. Roughly speaking, S_μ contains the sites in A whose distances to the boundary ∂A are between μ and $\mu+1$. Therefore, $\ell_{\vec{a}} \geq \ell+\mu$ for all $\vec{a} \in S_\mu$. We then obtain

$$\sum_{\vec{a}\in A\cap\Lambda} f(\ell_{\vec{a}}) = \sum_{\mu=0}^{\infty}\sum_{\vec{a}\in S_\mu\cap\Lambda} f(\ell_{\vec{a}}) \tag{G17}$$

$$\leq \sum_{\mu=0}^{\infty} f(\ell+\mu)|S_\mu\cap\Lambda|, \tag{G18}$$

where $|S_\mu \cap \Lambda|$ is the number of lattice sites that lie within S_μ.

Let $A_\mu = \{\vec{a}\in A : \mathrm{dist}(\vec{a},\partial A)\geq\mu\}$ be a subset of A containing sites at least a distance μ from the boundary of A. Clearly, $S_\mu = (A_\mu\backslash A_{\mu+1})\cup\partial A_{\mu+1}$ and $\partial S_\mu = \partial A_\mu \cup \partial A_{\mu+1}$. Roughly speaking, S_μ is a shell with the outer surface A_μ, the inner surface $A_{\mu+1}$, and a unity thickness. The number of lattice sites in S_μ will be bounded by $\eta\Phi(S_\mu) = \eta[\Phi(A_\mu)+\Phi(A_{\mu+1})]$ (see Sec. G 2 a for the definition of the constant η). Since A is compact and convex, $\Phi(A_{\mu+1}) < \Phi(A_\mu) < \Phi(A)$ (see Sec. G 2 b). Therefore, we arrive at the lemma. ■

a. Number of lattice sites in a compact region

In this section, we shall provide an upper bound on the number of lattice sites inside a compact set $A\subset\mathbb{R}^D$. We use this bound in Eq. (G18) to estimate the number of lattice sites in the set $|S_\mu\cap\Lambda|$ by its boundary area. Let $A_> = \{\vec{a}\in A\cap\Lambda : \mathrm{dist}(a,\partial A) > \frac{1}{3}\}$ be the set of lattice sites that are at least a distance $\frac{1}{3}$ away from the boundary ∂A, and let $A_\leq = A\backslash A_>$ be the other lattice sites of A.

First, note that for every lattice site $\vec{a}$ in $A_>$, there exists a D ball $\mathcal{B}_{1/4}(\vec{a})$ of radius $\frac{1}{4}$ that contains no other lattice site and $\mathcal{B}_{1/4}(\vec{a})\subset A$. Therefore, the number of lattice sites in $A_>$ is at most $\mathcal{V}(A)/\mathcal{V}[\mathcal{B}_{1/4}(\vec{a})] = \eta_1\mathcal{V}(A)$, where $\mathcal{V}(A)$ is the volume of A in $\mathbb{R}^D$ and $\eta_1 = \mathcal{V}[\mathcal{B}_{1/4}(\vec{a})]^{-1}$.

Next, to count the lattice sites in $A_\leq$, we note that for every $\vec{a}\in A_\leq$, we can select a point $f(\vec{a})\in\partial A$ on the boundary such that $\|f(\vec{a})-\vec{a}\|\leq\frac{1}{3}$. We now argue that $\|f(\vec{a})-f(\vec{b})\|\geq\frac{1}{3}$ for all distinct lattice sites $\vec{a}\neq\vec{b}$ in $A_\leq$. Indeed, since $\vec{a},\vec{b}$ are distinct lattice sites, the least distance between them is 1; i.e., $\|\vec{a}-\vec{b}\|\geq 1$. Using a triangle inequality, we can show that

$$\begin{aligned}\|f(\vec{a})-f(\vec{b})\| &\geq \|\vec{a}-\vec{b}\| - \|f(\vec{a})-\vec{a}\| - \|f(\vec{b})-\vec{b}\| \\ &\geq 1-\frac{1}{3}-\frac{1}{3}=\frac{1}{3}.\end{aligned} \tag{G19}$$

Therefore, a D ball $\mathcal{B}_{1/6}[f(\vec{a})]$ around $f(\vec{a})\in\partial A$ shall contain no $f(\vec{b})$ of any other lattice site $\vec{b}\in A_\leq$. Therefore, the number of lattice sites in $A_\leq$ is at most $\eta_2\Phi(A)$, where $\Phi(A) = |\partial A|$ is the boundary area of A and η_2 is the area of a $(D-1)$-dimensional disk of radius $1/6$.

In summary, the number of lattice sites in A is therefore at most $\eta_1 V(A) + \eta_2\Phi(A)$. In particular, for a shell A whose volume $\mathcal{V}(A)$ can be upper bounded by $\eta_3\Phi(A)$, the number of lattice sites will be at most $\eta\Phi(A)$, where $\eta = \eta_1\eta_3 + e\eta_2$.

b. Convex sets in $\mathbb{R}^D$ are shrinkable

In the proof of Lemma 9 [see the discussion after Eq. (G18)], we used the fact that $\Phi(A_\mu) < \Phi(A)$. In this section, we show that this property of A—which we term *shrinkability*—holds if A belongs to the class of convex and compact sets in $\mathbb{R}^D$. The formal definition is as follows.

Definition 1 (Shrinkable set).—A compact set $A\subset\mathbb{R}^D$ with boundary ∂A is shrinkable if, for all $r>0$, $A_r = \{\vec{a}\in A : \mathrm{dist}(\vec{a},\partial A)\geq r\}$, we have that $\Phi(A_r) = |\partial A_r| \leq |\partial A| = \Phi(A)$.

In other words, a set is shrinkable if the surface area of the boundary of $A_r\subseteq A$ is no larger than that A. In this section, we will prove that convexity is a sufficient condition for shrinkability. Recall that a set is compact if it is both closed and bounded, whereas convexity is usually defined as follows.

Definition 2.—A set A is convex if for any $x,y\in A$ and any θ such that $0\leq\theta\leq 1$, we have $\theta x + (1-\theta)y\in A$.

Examples of convex sets include D balls and hyperrectangles, which are also shrinkable. To prove this holds in general, we will first show that if A is convex, then A_r is also convex (or empty) for all $r>0$. To do this, we formulate an equivalent definition of a convex set as an intersection of half-spaces.

Definition 3.—A half-space $\mathcal{H}$ is given by the points $\{x\in\mathbb{R}^D | a^T x\geq b\}$, where $a\in\mathbb{R}^D\backslash\{0\}$.

From this definition, it follows that half-spaces are convex sets. A folk lemma [56] states that a closed set A is convex if and only if

$$A = \bigcap_{k\in I}\{\mathcal{H}_k | \mathcal{H}_k \text{ half-space}, A\subseteq\mathcal{H}_k\},$$

for some countable index set I. In other words, A is equivalent to the intersection of all half-spaces that contain it. Since convexity is preserved under arbitrary intersection, this implies that A is convex. The converse follows from the separating hyperplane theorem—see Ref. [56] for details.

With this equivalent definition of convexity in hand, we will prove that A_r is also convex.

Lemma 10.—If a compact set $A\subset\mathbb{R}^D$ is convex, then $A_r = \{\vec{a}\in A : \mathrm{dist}(\vec{a},\partial A)\geq r\}$ is convex (or empty) for all $r>0$.

Proof.—Write A as the intersection of half-spaces $\mathcal{H}_k = \{x\in\mathbb{R}^D | a_k^T x\geq b_k\}$, for $k\in I$. Then A_r is the intersection of the half-spaces given by $H_k^r = \{x\in\mathbb{R}^D | a_k^T x\geq b_k + r\}$. By the converse of the above lemma, A_r is convex (or empty). ■

To show that A is shrinkable, we must show that $\Phi(A_r) = |\partial A_r| \leq |\partial A| = \Phi(A)$. Following a standard technique in the literature, we define the nearest-point projection of $\mathbb{R}^D$ onto a convex set and then show that it is a

contraction. The following lemma implies that such a mapping is well defined.

Lemma 11.—Given a nonempty, compact and convex set $A \subseteq \mathbb{R}^D$ and a point $x \in \mathbb{R}^D$, there exists a unique point $p_A(x) \in A$ such that

$$p_A(x) = \arg\min_{y \in A} \|x - y\|.$$

Proof.—Since A is compact, the continuous function $d_x(y) = \|x - y\|$ must achieve its minimum value on A.

Now suppose that the minimum value of d_x occurs at a point $y \in A$. We will show that y is unique. Assume for the sake of contradiction that there exists some point $\tilde{y} \in A$ such that $d_x(y) = d_x(\tilde{y})$, but $y \neq \tilde{y}$. Then the set of points x, y, and $\tilde{y}$ form an isosceles triangle, with $\overline{y\tilde{y}}$ as the base. Dropping an altitude from x intersects this line segment at the midpoint m such that $\|x - m\| < \|x - y\| = \|x - \tilde{y}\|$. But m is a convex combination of y and $\tilde{y}$, i.e., $m = \frac{1}{2}(y + \tilde{y}) \in A$, so we have reached a contradiction. Thus, y must be unique, and, therefore, $p_A(x)$ is well defined. ■

The projection function $p_A(x)$ can be interpreted as generalizing the concept of the orthogonal projection into an affine subspace. It is also well known that the nearest-point projection p_A is a contraction mapping.

Lemma 12.—Given a nearest-point projection $p_A : \mathbb{R}^D \to A$ onto a convex set A, it holds for all $x, y \in \mathbb{R}^D$ that

$$\|p_A(x) - p_A(y)\| \leq \|x - y\|.$$

Proof.—While the lemma can be proved for all $x, y \in \mathbb{R}^D$, for our purposes, we only need to consider $x, y \notin A$. Assume that $p_A(x) \neq p_A(y)$. Then consider the hyperplanes H_x and H_y that pass through $p_A(x)$ and $p_A(y)$, respectively, and are perpendicular to the line segment $\overline{p_A(x)p_A(y)}$. (See the geometric diagram in Fig. 10.)

We prove by contradiction that x (y) and $p_A(y)$ [$p_A(x)$] lie on opposite sides of H_x (H_y). Suppose without loss of generality that x and $p_A(y)$ lie on the same side of H_x.

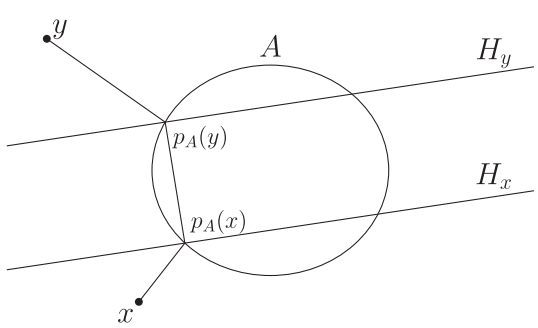

FIG. 10. The nearest-point projection p_A of two points x and y onto a compact set A (oval). Also depicted are the line segment connecting the two image points $p_A(x)$ and $p_A(y)$, as well as the two hyperplanes orthogonal to it.

Then the point where the altitude from x intersects the line segment $\overline{p_A(x)p_A(y)}$ would lie in A, contradicting the fact that $p_A(x)$ is the nearest point in A to x. Thus, x (y) must lie on the opposite side of H_x (H_y) from $p_A(y)$ [$p_A(x)$]. Then, as shown in Fig. 10, the points x and y must fall outside the rectangular strip between the two hyperplanes. From this we conclude that $\|p_A(x) - p_A(y)\| \leq \|x - y\|$. ■

The above result proves that the projection $p_A(x)$ is indeed a contraction. Since contraction mappings do not increase lengths, we can use this fact to demonstrate that the boundary of A_r is less than that of A.

Theorem 3.—If the set $A \subset \mathbb{R}^D$ is compact and convex, then $\Phi(A_r) = |\partial A_r| \leq |\partial A| = \Phi(A)$.

Proof.—Consider the projection $p_{A_r} : A \to A_r$. Note that for $r > 0$, we have that $A_r = \{x \in A | d(x, A^c) \geq r\}$ is entirely contained in the interior of A, which implies that $A_r \cap \partial A = \varnothing$. Thus, our situation satisfies the assumption we made in the proof of Lemma 12.

Under the action of p_{A_r}, any point in $\mathbb{R}^D$ outside of A_r will get mapped to ∂A_r. In particular, since the map is onto, ∂A will get mapped to ∂A_r, i.e., $p(\partial A) = \partial A_r$. Using the fact that p_{A_r} is contractive, we have that

$$\Phi(A_r) = |\partial A_r| = |p(\partial A)| \leq |\partial A| = \Phi(A),$$

from which we conclude that A is a shrinkable set. ■

This provides the final step in our proof of Lemma 9. Note that we do not require an explicit formula for the surface area of the boundary of a D-dimensional convex set. In general, one may use the Cauchy-Crofton formula to calculate this quantity—for more details, see Theorem 5.5.2 of Ref. [57].

[1] E. H. Lieb and D. W. Robinson, *The Finite Group Velocity of Quantum Spin Systems*, Commun. Math. Phys. **28**, 251 (1972).

[2] B. Nachtergaele, Y. Ogata, and R. Sims, *Propagation of Correlations in Quantum Lattice Systems*, J. Stat. Phys. **124**, 1 (2006).

[3] B. Nachtergaele and R. Sims, *Lieb-Robinson Bounds and the Exponential Clustering Theorem*, Commun. Math. Phys. **265**, 119 (2006).

[4] M. B. Hastings and T. Koma, *Spectral Gap and Exponential Decay of Correlations*, Commun. Math. Phys. **265**, 781 (2006).

[5] 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).

[6] 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).

[7] D.-M. Storch, M. Van Den Worm, and M. Kastner, *Interplay of Soundcone and Supersonic Propagation in Lattice*

Models With Power Law Interactions, New J. Phys. **17**, 063021 (2015).

[8] B. Nachtergaele, H. Raz, B. Schlein, and R. Sims, *Lieb-Robinson Bounds for Harmonic and Anharmonic Lattice Systems*, Commun. Math. Phys. **286**, 1073 (2009).

[9] I. Prémont-Schwarz, J. Hnybida, I. Klich, and F. Markopoulou-Kalamara, *Lieb-Robinson Bounds for Commutator-Bounded Operators*, Phys. Rev. A **81**, 040102(R) (2010).

[10] I. Prémont-Schwarz and J. Hnybida, *Lieb-Robinson Bounds on the Speed of Information Propagation*, Phys. Rev. A **81**, 062107 (2010).

[11] B. Nachtergaele and R. Sims, *Much Ado About Something: Why Lieb-Robinson Bounds Are Useful*, arXiv:1102.0835.

[12] 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).

[13] M. Cheneau, P. Barmettler, D. Poletti, M. Endres, P. Schauß, T. Fukuhara, C. Gross, I. Bloch, C. Kollath, and S. Kuhr, *Light-Cone-like Spreading of Correlations in a Quantum Many-Body System*, Nature (London) **481**, 484 (2012).

[14] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, *Towards the Fast Scrambling Conjecture*, J. High Energy Phys. 04 (2013) 022.

[15] M. Kliesch, C. Gogolin, and J. Eisert, *Lieb-Robinson Bounds and the Simulation of Time-Evolution of Local Observables in Lattice Systems*, in *Many-Electron Approaches in Physics, Chemistry and Mathematics, Mathematical Physics Studies*, edited by V. Bach and L. Delle Site (Springer International Publishing, Switzerland, 2014), p. 301.

[16] E. Hamza, R. Sims, and G. Stolz, *Dynamical Localization in Disordered Quantum Spin Systems*, Commun. Math. Phys. **315**, 215 (2012).

[17] P. Barmettler, D. Poletti, M. Cheneau, and C. Kollath, *Propagation Front of Correlations in an Interacting Bose Gas*, Phys. Rev. A **85**, 053625 (2012).

[18] M. B. Hastings, *Light-Cone Matrix Product*, J. Math. Phys. (N.Y.) **50**, 095207 (2009).

[19] U. Schollwöck, *The Density-Matrix Renormalization Group in the Age of Matrix Product States*, Ann. Phys. (Amsterdam) **326**, 96 (2011).

[20] T. Enss and J. Sirker, *Light Cone Renormalization and Quantum Quenches in One-Dimensional Hubbard Models*, New J. Phys. **14**, 023008 (2012).

[21] M. P. Woods, M. Cramer, and M. B. Plenio, *Simulating Bosonic Baths With Error Bars*, Phys. Rev. Lett. **115**, 130401 (2015).

[22] M. P. Woods and M. B. Plenio, *Dynamical Error Bounds for Continuum Discretisation via Gauss Quadrature Rules—A Lieb-Robinson Bound Approach*, J. Math. Phys. (N.Y.) **57**, 022105 (2016).

[23] M. B. Hastings, *An Area Law for One-Dimensional Quantum Systems*, J. Stat. Mech. (2007) P08024.

[24] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area Laws for the Entanglement Entropy*, Rev. Mod. Phys. **82**, 277 (2010).

[25] Z.-X. Gong, M. Foss-Feig, F. G. S. L. Brandão, and A. V. Gorshkov, *Entanglement Area Laws for Long-Range Interacting Systems*, Phys. Rev. Lett. **119**, 050501 (2017).

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

[27] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low, *Quantum Algorithm for Simulating Real Time Evolution of Lattice Hamiltonians*, in *Proceedings of the Fifty-Ninth IEEE Annual Symposium on Foundations of Computer Science, FOCS'18* (IEEE Computer Society, Washington, DC, 2018), pp. 350–360.

[28] J. W. Britton, B. C. Sawyer, A. C. Keith, C. C. J. 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).

[29] K. Kim, S. Korenblit, R. Islam, E. E. Edwards, M.-S. Chang, C. Noh, H. Carmichael, G.-D. Lin, L.-M. Duan, C. C. J. Wang, J. K. Freericks, and C. Monroe, *Quantum Simulation of the Transverse Ising Model with Trapped Ions*, New J. Phys. **13**, 105003 (2011).

[30] M. Saffman, T. G. Walker, and K. Mølmer, *Quantum Information with Rydberg Atoms*, Rev. Mod. Phys. **82**, 2313 (2010).

[31] J. S. Douglas, H. Habibian, C.-L. Hung, A. v. Gorshkov, H. J. Kimble, and D. E. Chang, *Quantum Many-Body Models with Cold Atoms Coupled to Photonic Crystals*, Nat. Photonics **9**, 326 (2015).

[32] B. 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).

[33] J. R. Maze, A. Gali, E. Togan, Y. Chu, A. Trifonov, E. Kaxiras, and M. D. Lukin, *Properties of Nitrogen-Vacancy Centers in Diamond: The Group Theoretic Approach*, New J. Phys. **13**, 025025 (2011).

[34] D. Otten, S. Rubbert, J. Ulrich, and F. Hassler, *Universal Power-Law Decay of Electron-Electron Interactions due to Nonlinear Screening in a Josephson Junction Array*, Phys. Rev. B **94**, 115403 (2016).

[35] 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).

[36] M. C. Tran, J. R. Garrison, Z.-X. Gong, and A. V. Gorshkov, *Lieb-Robinson Bounds on N-Partite Connected Correlation Functions*, Phys. Rev. A **96**, 052334 (2017).

[37] A. I. Larkin and Y. N. Ovchinnikov, *Quasiclassical Method in the Theory of Superconductivity*, Sov. Phys. JETP **28**, 1200 (1969).

[38] A. Kitaev and S. J. Suh, *The Soft Mode in the Sachdev-Ye-Kitaev Model and Its Gravity Dual*, J. High Energy Phys. 05 (2018) 183.

[39] S. Aaronson and A. Arkhipov, *The Computational Complexity of Linear Optics*, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing, STOC'11* (Association for Computing Machinery, New York, 2011), pp. 333–342.

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

[41] S. Michalakis, *Stability of the Area Law for the Entropy of Entanglement*, arXiv:1206.6900.

[42] J. Eisert and D. Gross, *Supersonic Quantum Communication*, Phys. Rev. Lett. **102**, 240501 (2009).

[43] J. Jünemann, A. Cadarso, D. Pérez-García, A. Bermudez, and J. J. García-Ripoll, *Lieb-Robinson Bounds for Spin-Boson Lattice Models and Trapped Ions*, Phys. Rev. Lett. **111**, 230404 (2013).

[44] S. Lloyd, *Universal Quantum Simulators*, Science **273**, 1073 (1996).

[45] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, *Efficient Quantum Algorithms for Simulating Sparse Hamiltonians*, Commun. Math. Phys. **270**, 359 (2007).

[46] G. H. Low and I. L. Chuang, *Optimal Hamiltonian Simulation by Quantum Signal Processing*, Phys. Rev. Lett. **118**, 010501 (2017).

[47] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, *Simulating Hamiltonian Dynamics with a Truncated Taylor Series*, Phys. Rev. Lett. **114**, 090502 (2015).

[48] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe, *Chemical Basis of Trotter-Suzuki Errors in Quantum Chemistry Simulation*, Phys. Rev. A **91**, 022311 (2015).

[49] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, *Toward the First Quantum Simulation with Quantum Speedup*, Proc. Natl. Acad. Sci. U.S.A. **115**, 9456 (2018).

[50] D. J. Luitz and Y. Bar Lev, *Emergent Locality in Systems with Power-Law Interactions*, Phys. Rev. A **99**, 010105(R) (2019).

[51] N. Maskara, A. Deshpande, M. C. Tran, A. Ehrenberg, B. Fefferman, and A. V. Gorshkov, *Complexity Phase Diagram for Interacting and Long-Range Bosonic Hamiltonians*, arXiv:1906.04178.

[52] M. Serbyn, Z. Papić, and D. A. Abanin, *Quantum Quenches in the Many-Body Localized Phase*, Phys. Rev. B **90**, 174302 (2014).

[53] D. V. Else, F. Machado, C. Nayak, and N. Y. Yao, *An Improved Lieb-Robinson Bound for Many-Body Hamiltonians with Power- Law Interactions*, arXiv:1809.06369.

[54] Note that there is a difference between the definition of the exponent α in Ref. [53] and in this paper. For two-body Hamiltonians, the exponent α in Ref. [53] is equal to our α minus the dimension D.

[55] J. Haah, *Product Decomposition of Periodic Functions in Quantum Signal Processing*, arXiv:1806.10236.

[56] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004), p. 36.

[57] D. A. Klain and G.-C. Rota, *Introduction to Geometric Probability* (Cambridge University Press, New York, 1997), pp. 46–47.