Complexity of Implementing Trotter Steps

Guang Hao Low,[1] Yuan Su,[1,*] Yu Tong,[2,3] and Minh C. Tran[4,5]

[1]*Microsoft Quantum, Redmond, Washington 98052, USA*

[2]*Department of Mathematics, University of California, Berkeley, California 94720, USA*

[3]*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[4]*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[5]*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

Quantum dynamics can be simulated on a quantum computer by exponentiating elementary terms from the Hamiltonian in a sequential manner. However, such an implementation of Trotter steps has gate complexity depending on the total Hamiltonian term number, comparing unfavorably to algorithms using more advanced techniques. We develop methods to perform faster Trotter steps with complexity sublinear in the number of terms. We achieve this for a class of Hamiltonians whose interaction strength decays with distance according to power law. Our methods include one based on a recursive block encoding and one based on an average-cost simulation, overcoming the normalization-factor barrier of these advanced quantum simulation techniques. We also realize faster Trotter steps when certain blocks of Hamiltonian coefficients have low rank. Combining with a tighter error analysis, we show that it suffices to use $\left(\eta^{1/3}n^{1/3} + n^{2/3}/\eta^{2/3}\right)n^{1+o(1)}$ gates to simulate uniform electron gas with n spin orbitals and η electrons in second quantization in real space, asymptotically improving over the best previous work. We obtain an analogous result when the external potential of nuclei is introduced under the Born-Oppenheimer approximation. We prove a circuit lower bound when the Hamiltonian coefficients take a continuum range of values, showing that generic n-qubit two-local Hamiltonians with commuting terms require at least $\Omega(n^2)$ gates to evolve with accuracy $\epsilon = \Omega(1/\mathrm{poly}(n))$ for time $t = \Omega(\epsilon)$. Our proof is based on a gate-efficient reduction from the approximate synthesis of diagonal unitaries within the Hamming weight-2 subspace, which may be of independent interest. Our result thus suggests the use of Hamiltonian structural properties as both necessary and sufficient to implement Trotter steps with lower gate complexity.

I. INTRODUCTION

Many-body Hamiltonians can be efficiently simulated on digital quantum computers using either product formulas (such as the Lie-Trotter-Suzuki formulas) or more advanced simulation algorithms. While short steps of product formulas (known as *Trotter steps*) typically require an implementation cost proportional to the number of Hamiltonian terms, a host of techniques have been developed in recent years to implement other quantum simulation algorithms with significantly reduced complexities.

The purpose of this work is to develop methods for performing Trotter steps that go beyond the sequential circuit implementation mentioned above, leading to faster quantum simulation algorithms. We focus on two-local Hamiltonians with power-law decaying interactions—all-to-all interactions whose strengths decay with distance x according to a power law $1/x^\alpha$ $(\alpha > 0)$—for concreteness. Many systems of physical relevance can be modeled by power-law interactions, such as trapped ions, Rydberg atoms, ultracold atoms and molecules, nitrogen-vacancy centers, and superconducting systems.

We first develop a block-encoding-based method to simulate power-law Hamiltonians with efficiently computable coefficients. While a block encoding introduces a slow-down factor proportional to the 1-norm of Hamiltonian coefficients, which is large for power-law interactions, we overcome this barrier by recursively decomposing the system using product formulas to effectively reduce the norm.

*buptsuyuan@gmail.com

We obtain further improvements by simulating commuting Hamiltonian terms with an average combination cost. This gives simulations with complexities almost linear in the spacetime volume for $\alpha \geq 2$ and improved scalings for $\alpha < 2$. We also implement faster Trotter steps when certain off-diagonal blocks of Hamiltonian coefficients exhibit low-rank properties. In particular, we achieve a nearly linear spacetime volume scaling for all $\alpha \geq 1$ using a recursive low-rank decomposition of the Hamiltonian. We extend our recursion techniques to κ-local Hamiltonians and more general fermionic models. In these cases, the exact complexity is determined by the tensor structure of the Hamiltonian coefficients and depends on the problem at hand.

We apply our methods to simulate electronic structure Hamiltonians in second quantization in real space. Combining with a tight Trotter error analysis, we show that $\left(\eta^{1/3}n^{1/3} + n^{2/3}/\eta^{2/3}\right) n^{1+o(1)}$ gates suffice to simulate uniform electron gas with n spin orbitals and η electrons, improving the best results from previous work. An analogous result holds when the external potential of nuclei is included from the Born-Oppenheimer approximation.

Performing faster Trotter steps for Hamiltonians with arbitrary coefficients is a challenging task in general. To confirm this intuition, we prove a gate-complexity lower bound. Specifically, we construct a class of n-qubit two-local Hamiltonians with commuting Pauli-Z terms whose coefficients take a continuum range of values. We show that these Hamiltonians require at least $\Omega(n^2)$ gates to simulate with accuracy $\epsilon = \Omega(1/\operatorname{poly}(n))$ for time $t = \Omega(\epsilon)$; thus the best method one can hope for is to sequentially exponentiate all the $\Theta(n^2)$ terms. Our proof depends on a gate-efficient reduction from the approximate synthesis of diagonal unitaries within the Hamming weight-2 subspace, which we then address by adapting a volume-comparison technique from previous work. Our result thus suggests the use of structural properties of the target Hamiltonian as both necessary and sufficient to achieve lower gate complexity for implementing Trotter steps.

A. Quantum algorithms for quantum simulation

Simulating many-body physical systems is one of the most promising applications of digital quantum computers. Indeed, the idea of quantum computing as originally proposed by Feynman [1], Manin [2], and others is strongly motivated by quantum simulation. Efficient quantum simulations can be used to extract statical and dynamical properties of physical systems, which has potential applications in various areas, such as condensed-matter physics [3], chemistry [4–6], and high-energy physics [7]. Meanwhile, recent developments in quantum simulation algorithms have also provided technical tools that influenced the design of other quantum algorithms [8–12] and proofs of other results in areas beyond quantum computing [13–17].

There are many quantum algorithms one can use to perform quantum simulation. At a high level, these algorithms can be categorized according to their default input models. Common Hamiltonian input models include the following: (i) Linear combinations of Hermitians (LCH), where the target Hamiltonian takes the form

$$H = \sum_{\gamma=1}^{\Gamma} H_\gamma, \tag{1}$$

with H_γ Hermitian and the exponentials e^{-itH_γ} implementable on a quantum computer. Simulation algorithms that work in this input model include one based on product formulas [18,19] (such as the Lie-Trotter formula and its higher-order extensions), as well as a more recent algorithm based on random sampling [20]; and (ii) linear combinations of unitaries (LCU), where the target Hamiltonian takes the form

$$H = \sum_{\gamma=1}^{\Gamma} \beta_\gamma U_\gamma, \tag{2}$$

with U_γ unitary, $\beta_\gamma > 0$ and the controlled operators $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U_\gamma$ implementable on a quantum computer. There are also various algorithms working in the LCU model, including one based on implementing truncated Taylor series [21] and one based on qubitization [22]. For our purpose, we mainly focus on the algorithm based on product formulas as well as the qubitization algorithm, which we review in more detail in Secs. II B and II C.

Naturally, there is no silver-bullet method that solves all simulation problems of interest with the optimal gate complexity. Choices of algorithms should thus be made on a case-by-case basis. There are in fact a few common desirable features shared among many algorithms mentioned above. For instance, it has been well known that many LCU approaches such as the Taylor-series algorithm and qubitization have complexities (nearly) linear in the simulation time and logarithmic in the inverse accuracy. But similar scalings can also be achieved using product formulas by taking linear combinations of formulas with different step sizes and repetition numbers [23]. LCU-type algorithms also have the appealing feature that they can be used to design other functions of Hamiltonians [24], which applies to problems beyond quantum simulation [11,25], such as solving linear systems of equations [8,26], preparing ground states [10,27,28], and performing phase estimation [25,29]. Again, these problems can be well solved using algorithms in the LCH model (and can sometimes be more resource friendly) as demonstrated in recent work, such as Refs. [30,31].

However, there is one question concerning the gate complexities of implementing these approaches, which has not been satisfactorily answered so far. View a sufficiently

long-time Hamiltonian evolution as a concatenation of short-time steps, and consider the complexity of simulating each short evolution. Then a host of techniques are recently developed to implement LCU approaches with cost depending on the target system size as opposed to the number of Hamiltonian terms [32,33]; a similar goal may be realized for the randomized method via an importance sampling [20,34,35]. By contrast, how to provably achieve a similar complexity for product formulas was somewhat under investigated: we review related work on the implementation of Trotter steps in Sec. I B.

For an n-qubit system with two-local interactions, one has that the total number of terms scales like $\Theta(n^2)$ but the system size is only linear in n. This comparison becomes $\Theta(n^\kappa)$ versus n for a κ-local Hamiltonian. In light of this gap, it is natural to ask when and how one can implement product formulas with cost scaling better than the sequential method. We address this question by presenting necessary and sufficient conditions under which faster Trotter steps are possible. As an application, we give a simulation algorithm for electronic structure Hamiltonians in second quantization in real space with complexity $\left(\eta^{1/3}n^{1/3} + n^{2/3}/\eta^{2/3}\right)n^{1+o(1)}$—asymptotically the fastest real-space simulation to date.

B. Previous related work

We now discuss prior work on implementing Trotter steps that are relevant to our paper.

First, there are previous studies on the so-called fast-forwardability of Hamiltonian evolution [36–39]. Some of those techniques, such as introducing efficiently computable phases and diagonalizing quadratic Hamiltonians, can also be used to perform Trotter steps. However, their fundamental goal is quite different from ours. In quantum fast-forwarding, one is asked to simulate the target Hamiltonian for a sufficiently long time, and the goal is to reduce the scaling of time in the gate complexity (potentially at the cost of increasing the system-size scaling). In contrast, each Trotter step approximates only the ideal evolution for a short time, and so the scaling with time is no longer a key contribution to the complexity. Instead, our main goal here is to reduce the dependence on the size of the simulated system.

For electronic structure models represented under arbitrary basis, the number of terms in the Hamiltonian typically scales like $\Theta(n^4)$ for n spin orbitals, so a sequential implementation of Trotter steps would have cost scaling $\Theta(n^4)$. To address this, recent work developed quantum circuits based on low-rank factorizations of such systems [40,41] (see Refs. [42,43] for more recent developments of such a method). Specifically, they apply product formulas to decompose the Hamiltonian into multiple components, each of which has coefficients with certain low-rank properties and can be further implemented by diagonalization.

The gate complexity of the resulting circuits would then depend on the value of rank as opposed to the number of Hamiltonian terms, which significantly reduces the cost per time step. However, those works did not rigorously analyze the total complexity of the proposed methods, and it is unclear how much overall advantage their approach can offer. In fact, their factorization does not seem to preserve the commutation relations between Hamiltonian terms and could potentially introduce a Trotter error larger than the sequential approach (so more Trotter steps would be required to reach the same simulation accuracy).

Another related approach to reducing the complexity of quantum simulation is to truncate Hamiltonian terms of small sizes. Such truncations are useful for not only performing Trotter steps [44–46], but also implementing more advanced quantum simulation algorithms [47–49]. Generally speaking, the error introduced in the truncation will grow linearly with time, so the simulation is accurate only when the evolution is sufficiently short. In particular, for rapidly decaying power-law interactions with exponent $\alpha > 2$, a truncation is possible only when $t = \mathcal{O}\left(n^{(\alpha-2)/(\alpha-1)}\right)$ [44]. For simulations of chemistry and material models, truncation thresholds can often be determined empirically under certain assumptions of the model Hamiltonians.

Here, our work considers simulating two-local Hamiltonians with interaction strength decaying according to power law, and we study the cost of implementing one short Trotter step as well as the entire long-time simulation, using the so-called block-encoding technique and recursive and hierarchical low-rank decompositions [50,51]. Our motivation for using the low-rank decomposition partly overlaps with that of a recent work by Nguyen *et al.* [52], but the main problems we study are different. Instead of Hamiltonian simulation, they studied the block encoding of kernel matrices of the form

$$K = \sum_{j,k=1}^{n} \beta_{j,k}|j\rangle\langle k|, \qquad (3)$$

where $\beta_{j,k}$ can be power-law functions such as $1/|j - k|^\alpha$. The matrix K is an n-dimensional operator and has spectral norm $\|K\| = \|\beta\| = \Theta(1)$ for $\alpha > 1$ [53]. Instead, our problem centers around the simulation of

$$H = \sum_{j,k=1}^{n} \beta_{j,k}X_j Y_k, \qquad (4)$$

which is a 2^n-dimensional operator and has spectral norm generally scaling with the *vector* 1-*norm* of coefficients: $\|H\| = \Theta(\|\beta\|_1) = \Theta(n)$ for $\alpha > 1$. Thus, a naive block encoding of our H will have an intrinsically worse normalization factor than that of their K; see Sec. II C for further explanations of how such normalization factors affect the

complexity of quantum simulation. Nevertheless, we overcome this technical obstacle by recursively decomposing the Hamiltonian using product formulas, which significantly reduces the 1-norm while maintaining the overall scaling of gate count.

Finally, we note that there is a large body of previous work analyzing and optimizing the concrete resources for implementing Trotter steps, for both near-term and fault-tolerant quantum computers (see Refs. [3,54,55] as well as other work citing and cited by these papers). We do not attempt to optimize the constant factors of the complexity of our methods, but we consider such optimizations to be an interesting subject for future investigation.

C. Faster Trotter steps by recursion

Consider a Hamiltonian H with Γ terms $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$. If each Hamiltonian term can be exponentiated on a quantum computer with cost $\mathcal{O}(1)$, then one can simulate the evolution of H for a short time using product formulas, and the complexity would scale like $\mathcal{O}(\Gamma)$. We identify scenarios in which improved implementations of Trotter steps are possible with gate complexities sublinear in Γ.

We focus on a class of two-local Hamiltonians in one spatial dimension with all-to-all interactions and magnitude of the coefficients decaying with distance x according to power law $1/x^\alpha$ ($\alpha > 0$). We describe how our results can be extended to higher spatial dimensions in Appendix C, and to more general local and fermionic models (though the amount of improvement largely depends on the tensor structure of the Hamiltonian coefficients, which has not been fully understood). We restrict to power-law models because product formulas are known to provide the fastest method for simulating this class of Hamiltonian, so we can directly compare our result with the state of the art. Examples of power-law interactions include the Coulomb interaction between charged particles and the dipole-dipole interaction between molecules, both of which are ubiquitous in quantum chemistry—a primary target application of quantum computation. In physics, impressive controls in recent experiments with trapped ions [56,57], Rydberg atoms [58], and ultracold atoms and polar molecules [59,60] have enabled the possibility to study new phases of matter with power-law interactions [61–66] and contributed to a growing interest in simulating such systems. In fact, we describe a direct application of our method in Sec. VI for faster simulations of electronic structure Hamiltonians in real space.

Assume that the coefficients of the target Hamiltonian are efficiently computable. As explained above, there have been a host of techniques developed recently based on the notion of block encoding, which enables simulation with complexities depending only on the system size. One may ask if these techniques also lead to faster Trotter steps with a similar cost scaling. Unfortunately, the answer is negative in general. This is because a block encoding typically introduces a normalization factor proportional to the 1-norm of the Hamiltonian coefficients, and we thus need to repeat a corresponding number of times to perform Trotter steps. For instance, one can block-encode power-law Hamiltonians with gate complexity $\Theta(n)$, but this introduces a normalization factor proportional to the 1-norm, which is generally $\Theta(n)$ for power-law interactions with $\alpha > 1$. Meanwhile, a Trotter step for the power-law models has an almost constant evolution time. So one roughly needs a total number of

$$\underbrace{n}_{\text{block encoding}} \cdot \underbrace{n}_{\text{effective time}} = \underbrace{n^2}_{\text{Trotter step complexity}} \tag{5}$$

gates to implement a single Trotter step, which has no benefit over the sequential implementation. See Sec. III A for a more detailed explanation of this issue [67].

We develop a method based on block encoding that overcomes the above technical issue. The key observation is that product formulas can be used to reduce the 1-norm of Hamiltonian coefficients "almost for free": we apply product formulas to recursively decompose the Hamiltonian into multiple groups, but such a coarse-grained decomposition introduces a Trotter error no larger than the sequential approach. We choose the decomposition to significantly reduce the 1-norm of each group while maintaining the overall scaling of the gate complexity, giving an efficient block-encoding circuit. The resulting simulation has gate complexity $(nt)^{1+o(1)}$ when $\alpha \geq 2$, and $n^{3-\alpha+o(1)}t^{1+o(1)}$ when $\alpha < 2$. We formally state this theorem as Theorem 1 in Sec. III and preview it below.

Theorem 1: (Faster Trotter steps using block encoding). *Consider two-local Hamiltonians*

$$H = \sum_{\sigma,\sigma' \in \{i,x,y,z\}} \sum_{1 \leq j < k \leq n} \beta_{j,k}^{(\sigma,\sigma')} P_j^{(\sigma)} P_k^{(\sigma')},$$

where $\left| \beta_{j,k}^{(\sigma,\sigma')} \right| \leq 1/|j-k|^\alpha$ *for some constant* $\alpha > 0$ *and* $P^{(\sigma)}$ *($\sigma = i,x,y,z$) are the identity and Pauli matrices. Let $t > 0$ be the simulation time and $\epsilon > 0$ be the target accuracy. Assume that the coefficient oracle*

$$O_{\beta,\sigma,\sigma'}|j,k,0\rangle = |j,k,\beta_{j,k}^{(\sigma,\sigma')}\rangle \tag{6}$$

can be implemented with gate complexity $\mathcal{O}$ *(polylog* (nt/ϵ)). *Then H can be simulated using the algorithm of Sec. III C with $\mathcal{O}$ (log(nt/ϵ)) ancilla qubits and gate complexity*

$$\begin{array}{ll} nt \left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq 2, \\ n^{3-\alpha}t \left(\frac{nt}{\epsilon}\right)^{o(1)}, & 0 < \alpha < 2. \end{array} \tag{7}$$

The Hamiltonian decomposition we study has an additional feature that all the terms within each group commute with each other. Such terms can be simulated from block encodings of their subcomponents with an average combination cost, similar to the interaction-picture simulation [26,68] but without an exponentially growing factor. We leverage this observation to further improve the complexity to $n^{2-\alpha/2+o(1)}t^{1+o(1)}$ for $1 \leq \alpha < 2$ and $n^{5/2-\alpha+o(1)}t^{1+o(1)}$ for $\alpha < 1$. We summarize our result in Theorem 2 previewed in the following and describe the details in Sec. IV.

Theorem 2: (Faster Trotter steps using average-cost simulation). *Consider two-local Hamiltonians*

$$H = \sum_{\sigma,\sigma'\in\{i,x,y,z\}} \sum_{1\leq j<k\leq n} \beta_{j,k}^{(\sigma,\sigma')} P_j^{(\sigma)} P_k^{(\sigma')},$$

where $\left|\beta_{j,k}^{(\sigma,\sigma')}\right| \leq 1/|j-k|^\alpha$ *for some constant* $\alpha > 0$ *and* $P^{(\sigma)}$ *($\sigma = i,x,y,z$) are the identity and Pauli matrices. Let $t > 0$ be the simulation time and $\epsilon > 0$ be the target accuracy. Assume that the coefficient oracle*

$$O_{\beta,\sigma,\sigma'}|j,k,0\rangle = |j,k,\beta_{j,k}^{(\sigma,\sigma')}\rangle \qquad (8)$$

can be implemented with gate complexity $\mathcal{O}$ (polylog (nt/ϵ)). *Then H can be simulated using the algorithm of Sec. IV C with* $\mathcal{O}(\log(nt/\epsilon))$ *ancilla qubits and gate complexity*

$$\begin{aligned} n^{2-\alpha/2}t\left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & \quad 1 \leq \alpha < 2, \\ n^{5/2-\alpha}t\left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & \quad 0 < \alpha < 1. \end{aligned} \qquad (9)$$

We also realize faster Trotter steps when certain blocks of Hamiltonian coefficients exhibit low-rank properties. Such assumptions were studied in the context of block-encoding kernel matrices [52]. Under the same hierarchical low-rank assumptions [50,51], we directly implement the diagonalization procedure without using block encoding

to achieve gate complexity $(nt)^{1+o(1)}$ nearly linear in the spacetime volume for all $\alpha \geq 1$. This is summarized in Theorem 3 and restated below. See Sec. V for details. We summarize our improvements (in a simplified form) in Table I for simulating power-law Hamiltonians in general d spatial dimensions.

Theorem 3: (Faster Trotter steps using low-rank decomposition). *Consider two-local Hamiltonians*

$$H = \sum_{\sigma,\sigma'\in\{i,x,y,z\}} \sum_{1\leq j<k\leq n} \beta_{j,k}^{(\sigma,\sigma')} P_j^{(\sigma)} P_k^{(\sigma')},$$

where $\left|\beta_{j,k}^{(\sigma,\sigma')}\right| \leq 1/|j-k|^\alpha$ *for some constant* $\alpha > 0$ *and* $P^{(\sigma)}$ *($\sigma = i,x,y,z$) are the identity and Pauli matrices. Let $t > 0$ be the simulation time and $\epsilon > 0$ be the target accuracy. Then H can be simulated using the algorithm of Sec. V B with* $\mathcal{O}(\log(nt/\epsilon))$ *ancilla qubits and gate complexity*

$$\begin{aligned} \rho n t \left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & \quad \alpha \geq 1, \\ \rho n^{2-\alpha} t \left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & \quad 0 < \alpha < 1. \end{aligned} \qquad (10)$$

Here, $1 \leq \rho \leq n$ *defined in Eq. (127) is the maximum truncation rank of certain off-diagonal blocks of coefficient matrices ($\rho = \mathcal{O}(\log(nt/\epsilon))$ if the coefficient distribution exactly matches a power law in one spatial dimension).*

Although we achieve various speedups for simulating power-law Hamiltonians based on different techniques, we use recursion in the development of all our methods, and the core idea behind our improvements can all be understood through the so-called "master theorem" [69–72]. Specifically, to solve a problem of size n using recursion, we divide the problem into m subproblems, each of which can be seen as an instance of the original problem of size

TABLE I. Comparison of our results and the best previous results for simulating power-law interactions $1/x^\alpha$ in d spatial dimensions. Upper bounds are used for some gate complexity expressions for presentational purpose. The truncation result of Ref. [44] holds only for a sufficiently short time and is thus not compared here. Both the block encoding and the average-cost simulation method assume that the coefficients of the Hamiltonian are efficiently computable. The low-rank method has a dependence on the maximum truncation rank $1 \leq \rho \leq n$, which is polylogarithmic in the input parameters for all power-law models in one spatial dimension and many power-law models, such as Coulomb interactions in higher spatial dimensions.

Method	Hamiltonian		
	$\alpha \geq 2d$	$d \leq \alpha < 2d$	$0 < \alpha < d$
Sequential [44]	$n^{2+o(1)}t^{1+o(1)}$	$n^{2+o(1)}t^{1+o(1)}$	$n^{3-\alpha/d+o(1)}t^{1+o(1)}$
Block encoding (Sec. III)	$(nt)^{1+o(1)}$	$n^{3-\alpha/d+o(1)}t^{1+o(1)}$	$n^{3-\alpha/d+o(1)}t^{1+o(1)}$
Average cost (Sec. IV)	—	$n^{2-\alpha/2d+o(1)}t^{1+o(1)}$	$n^{5/2-\alpha/d+o(1)}t^{1+o(1)}$
Low rank (Sec. V)	$\rho(nt)^{1+o(1)}$	$\rho(nt)^{1+o(1)}$	$\rho n^{2-\alpha/d+o(1)}t^{1+o(1)}$

n/m, so

$$\text{cost}_{\text{rec}}(n) = m\,\text{cost}_{\text{rec}}\left(\frac{n}{m}\right) + \text{cost}(n), \qquad (11)$$

where $\text{cost}(n)$ quantifies the additional cost to combine solutions of the subproblems in the current layer of recursion and $\text{cost}_{\text{rec}}(n)$ denotes the total cost of the recursion. Then the master theorem asserts that, under certain assumptions of the cost function, the scaling of $\text{cost}_{\text{rec}}(n)$ is the same as that of $\text{cost}(n)$ up to a logarithmic factor, i.e.,

$$\text{cost}_{\text{rec}}(n) = \mathcal{O}\left(\text{cost}(n)\log(n) + n\right).$$

See Lemma 1 of Sec. II A for a more formal description of this result. However, performing the combination step can often be much simpler than directly solving the full problem, and one then expects to get a better $\text{cost}_{\text{rec}}(n)$ when $\text{cost}(n)$ is improved. We show that improved recursions are indeed possible for power-law systems by reducing the normalization factor of the Hamiltonian and by exploiting low-rank properties of certain blocks of the Hamiltonian coefficients, leading to faster quantum simulation by recursion.

The electronic structure Hamiltonian is one of the most widely studied candidate models in quantum simulation [4–6]. An efficient simulation of such Hamiltonians could provide insights to various problems in chemistry and material science. Here, we focus on a simulation in real space, an idea investigated by Kassal *et al.* [73] and subsequently pursued by later work, such as Refs. [74–78]. Although the full Hamiltonian does not satisfy power laws, the Coulomb potential part can be represented in second quantization with the magnitude of coefficients decaying as $1/x$, to which our method applies. In particular, we can choose $\rho = \mathcal{O}\left(\log(nt/\epsilon)\right)$ in the low-rank decomposition [79] to efficiently implement Trotter steps. Combining with an improved Trotter error analysis, we show that $\left(\eta^{1/3}n^{1/3} + n^{2/3}/\eta^{2/3}\right)n^{1+o(1)}$ gates suffice to simulate the uniform electron gas with n spin orbitals and η electrons; we obtain an analogous result when the external potential of nuclei is introduced under the Born-Oppenheimer approximation. This improves the best previous results for simulating electronic structures. We describe this application in Sec. VI, with the improved Trotter error analysis detailed in Appendix B.

D. Circuit lower bound

It is worth noting that all our above methods hold under certain additional assumptions on the target Hamiltonian: we assume that either the Hamiltonian coefficients are efficiently computable or certain blocks of them have low rank. If no such structural properties are available, we are then faced with Hamiltonians with arbitrary coefficients, and intuitively there would be no implementation of Trotter steps better than the sequential method.

We prove a circuit lower bound to justify this intuition. Specifically, we consider a class of two-local Hamiltonians of the form

$$H = \sum_{1 \le j < k \le n} \beta_{j,k} Z_j Z_k, \qquad (12)$$

where Z_j denotes the Pauli-Z operator on the jth qubit and coefficients are arbitrarily chosen from a continuum range of values $|\beta_{j,k}| \le t$. Subclasses of such Hamiltonians are of interest in areas beyond quantum simulation [80]. Even for such commuting H, we show the gate-complexity lower bound $\Omega\left(n^2/\log(b|\mathcal{K}|)\right)$ to evolve with accuracy $\epsilon = \Omega(1/\text{poly}(n))$ for $t = \Omega(\epsilon)$, using quantum circuits of $b \ge n$ qubits with a gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$. For circuits with a continuous gate set, we can first compile them using a finite universal gate set (say applying the Clifford+T synthesis [81] or the Solovay-Kitaev theorem [82]) and invoke the above bound. We discuss the circuit lower bounds in detail in Sec. VII with the result summarized in Corollary 1 and previewed here.

Corollary 1: (Simulating two-local commuting Hamiltonians). *Consider two-local Hamiltonians*

$$H = \sum_{1 \le j < k \le n} \beta_{j,k} Z_j Z_k,$$

where coefficients take values up to $|\beta_{j,k}| \le t$. Given accuracy $0 < \epsilon < 1/3$, number of qubits $b \ge n$, and two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$, if $t = \Omega(\epsilon)$,

$\min\{g \mid \forall 2\text{-local Hamiltonian } H, \exists \text{ circuit } V \text{ on}$

$b \text{ qubits with } g \text{ gates from } \mathcal{K}, \left\| e^{-iH} - V \right\| \le \epsilon\}$

$$= \Omega\left(\frac{n^2}{\log\left(b|\mathcal{K}|\right)} - n\,\text{polylog}\left(\frac{n}{\epsilon}\right)\right). \qquad (13)$$

Under the same assumption but choosing $\mathcal{K}$ to be the set of arbitrary two-qubit gates,

$\min\{g \mid \forall 2\text{-local Hamiltonian } H, \exists \text{ circuit } V \text{ on}$

$b \text{ qubits with } g \text{ gates from } \mathcal{K}, \left\| e^{-iH} - V \right\| \le \epsilon\}$

$$= \Omega\left(\frac{n^2}{\log b} - n\,\text{polylog}\left(\frac{n}{\epsilon}\right)\right). \qquad (14)$$

Underpinning our circuit lower bound proof is an efficient reduction from the approximate synthesis of diagonal unitaries within the Hamming weight-2 subspace, up to a gate overhead of $\mathcal{O}\left(n\,\text{polylog}(n/\epsilon)\right)$. We describe this reduction in Sec. VII B (with the circuit illustrated

in Fig. 4). The Hamming weight-2 subspace has dimensions $\binom{n}{2} = \Theta(n^2)$, so our problem is reduced to studying diagonal operators

$$\mathcal{D}_{\theta_{\max}} = \left\{ \sum_x e^{i\theta_x} |x\rangle\langle x|, \ |\theta_x| \le \theta_{\max} \right\} \qquad (15)$$

for $\Theta(n^2)$ values of x. We then show in Sec. VII A that such diagonal unitaries require roughly n^2 gates to approximately implement, generalizing a previous lower bound for exact synthesis due to Bullock and Markov [83,84]. See also recent work [85] for a related bound expressed in terms of measure. Our argument is based on an adaption of a technique of Knill [86]. Knill proved asymptotic circuit lower bounds for synthesizing the full unitary group using a volume-comparison technique, but here we consider only diagonal unitaries whose volume can be easily evaluated in closed form (see Sec. II D). The volume of diagonal unitaries predominantly depends on the dimensionality $\Theta(n^2)$, which leads to our desired lower bound scaling.

Additionally, we prove a second lower bound in Appendix A showing that any approximate realization of the coefficient oracle O_β requires $\Omega(n^2)$ gates in the worst case. Combining both our upper and lower bounds, we conclude that the use of structural properties of the Hamiltonian is both necessary and sufficient to achieve lower gate complexity for implementing Trotter steps.

We briefly summarize in Sec. II the preliminaries of our paper and present in Sec. VIII a collection of questions related to our result for future work.

II. PRELIMINARIES

A. Notation and terminology

We now introduce notation and terminology to be used in the remainder of our paper.

We use lowercase Latin letters as well as the Greek alphabet (in both upper and lower cases) to denote scalars, and we save uppercase Latin letters for matrices and operators. For instance, we use n for the system size, t for the evolution time (assuming $t > 0$ without loss of generality), ϵ for the simulation accuracy, and we write H to denote the target Hamiltonian, X, Y, Z to denote Pauli operators, and I to denote the identity matrix. When discussing fermionic Hamiltonians, we use $A^\dagger$, A, and N to represent fermionic creation, annihilation, and occupation-number operators, respectively [87]. We write the commutator of matrices B and C as $[B, C] := BC - CB$ when the multiplications are well defined. When multiplying noncommutative quantities, we use abbreviations like $\prod_{j=1}^n$ to denote the ordering where the smallest index appears on the right, e.g., $\prod_{j=1}^n U_j = U_n \cdots U_1$. We let a summation be zero and a product be one if their lower limits exceed the upper limits. We interchangeably use the decimal representation $|x\rangle$ ($x = 0, \ldots, 2^n - 1$) and the binary representation $|x_{n-1}, \ldots, x_0\rangle$ ($x_j = 0, 1$) of computational basis states if no ambiguity arises.

We also construct vectors of scalars and operators and use subscripts to index them, e.g., $\beta_{j,k}$ and Z_j. We then define the transpose operation $\beta_{j,k}^\top := \beta_{k,j}$. Our focus is on simulating two-local systems throughout the paper, so the Hamiltonian coefficients will typically have no more than two subscript indices. We assume that coefficients in an n-qubit Hamiltonian can be represented in binary using $\mathcal{O}(\log(nt/\epsilon))$ bits, for otherwise we may truncate the binary representation and simulate the truncated Hamiltonian with error at most $\mathcal{O}(\epsilon)$.

Our analysis requires various norms defined for vectors and matrices. For $[\beta_{j,k}]$, we define the vector 1-*norm* $\|\beta\|_1 := \sum_{j,k} |\beta_{j,k}|$, the *max-norm* $\|\beta\|_{\max} := \max_{j,k} |\beta_{j,k}|$, the *Euclidean norm* $\|\beta\| := \sqrt{\sum_{j,k} |\beta_{j,k}|^2}$ and the *induced* 1-*norm* $\|\|\beta\|\|_1 := \max_j \sum_k |\beta_{j,k}|$. We also need a restricted version of the induced 1-norm defined as

$$\|\|\beta\|\|_{1,[\eta]} := \max_j \max_{k_1 < \cdots < k_\eta} \left(|\beta_{j,k_1}| + \cdots + |\beta_{j,k_\eta}| \right). \quad (16)$$

By definition, this norm $\|\|\beta\|\|_{1,[\eta]}$ sums only the largest η elements in a row (maximized over all rows), and is thus always upper bounded by the induced 1-norm. We see later in the quantum chemistry application that the gap between these two norms can be significant. We use $\|B\|$ to denote the *operator norm* of B; this is also known as the *spectral norm* and its value is given by the largest singular value of B.

We use calligraphic uppercase letters to denote (un)structured sets. For instance, we use $\mathcal{H}$ to represent an arbitrary finite-dimensional Hilbert space with (normalized) quantum states $|\psi\rangle \in \mathcal{H}$ (we write $\mathbb{C}^m$ or $\mathbb{R}^m$ if the dimensionality m is explicitly provided). Given an underlying n-qubit system, we denote $\mathcal{W}_\eta$ to be the subspace spanned by computational basis states with Hamming weight η ($\dim(\mathcal{W}_\eta) = \binom{n}{\eta}$); $\mathcal{W}_\eta$ also denotes the subspace spanned by states with η particles for second-quantized fermionic systems. We may then define the operator norm restricted to the Hamming weight-η subspace

$$\|B\|_{\mathcal{W}_\eta} := \max_{|\phi_\eta\rangle, |\psi_\eta\rangle \in \mathcal{W}_\eta} \left| \langle \phi_\eta | B | \psi_\eta \rangle \right| \quad (17)$$

for an arbitrary n-qubit operator B. Given $[\beta_{j,k}]$ and a collection $\mathcal{B} \subseteq \mathbb{Z} \times \mathbb{Z}$ of pairs of indices (u, v), we define the *restricted max-norm* and 1-*norm* as

$$\|\beta\|_{\max, \mathcal{B}} := \max_{(u,v) \in \mathcal{B}} |\beta_{u,v}|, \qquad \|\beta\|_{1,\mathcal{B}} := \sum_{(u,v) \in \mathcal{B}} |\beta_{u,v}|. \quad (18)$$

And as introduced earlier, the set $\mathcal{D}_{\theta_{\max}}$ is the set of diagonal unitaries with phase angles between $-\theta_{\max}$ and $\theta_{\max}$.

We say an operator $G : \mathcal{G} \to \mathcal{H}$ is an *isometry* if $G^\dagger G = I$. By definition, G is necessarily injective and $G^\dagger$ is necessarily surjective, whereas $GG^\dagger$ is an orthogonal projection on $\mathcal{H}$ with image $\mathrm{im}(GG^\dagger) = \mathrm{im}(G)$ and kernel $\ker(GG^\dagger) = \ker(G^\dagger)$. We thus obtain the Hilbert-space isomorphism $\mathcal{G} \cong \mathrm{im}(GG^\dagger) \subseteq \mathcal{H}$ specified by the operators G and $G^\dagger$. Choosing any orthonormal basis that respects this isomorphism, we have the matrix representation

$$G = \begin{bmatrix} I \\ 0 \end{bmatrix}, \qquad G^\dagger = \begin{bmatrix} I & 0 \end{bmatrix}. \tag{19}$$

Examples of isometries include the following: (i) unitary operators U; (ii) quantum states $|\psi\rangle$; (iii) tensor product $G_1 \otimes G_2$ if G_1 and G_2 are isometries; and (iv) composition $G_2 G_1$ if G_1 and G_2 are isometries and the composition is well defined. Isometries will be used later in Sec. II C to describe block encodings and the qubitization algorithm.

Finally, we use $\mathcal{O}(\cdot)$ and $\Omega(\cdot)$ to mean asymptotically bounded above and below respectively, write $\Theta(\cdot)$ if both relations hold, and use the tilde symbol to suppress polylogarithmic factors. This is similar to the notation of a previous work on Trotter error analysis [44], except we do not need their $O(\cdot)$ for order conditions. To analyze the scaling of functions satisfying recurrence relations, we use the following version of the master theorem adapted from Refs. [69–72].

Lemma 1: (Master theorem). *Let* $\mathrm{cost}, \mathrm{cost}_{rec} : \mathbb{Z}_{\geq 1} \to \mathbb{R}_{\geq 0}$ *be non-negative functions defined for positive integers, such that there exist constant* $\gamma \geq 0$, $m_1, m_2 \in \mathbb{Z}_{\geq 1}$, *and* $n_0 \in \mathbb{Z}_{\geq 1}$ *for which*

$$\begin{aligned} \mathrm{cost}_{rec}(n) &\leq \gamma, & 1 \leq n \leq n_0, \\ \mathrm{cost}_{rec}(n) &\leq m_1 \, \mathrm{cost}_{rec}\left(\left\lfloor \tfrac{n}{m_1+m_2} \right\rfloor\right) + m_2 \, \mathrm{cost}_{rec}\left(\left\lceil \tfrac{n}{m_1+m_2} \right\rceil\right) + \mathrm{cost}(n), & n > n_0. \end{aligned} \tag{20}$$

If $\mathrm{cost}(n) = \mathcal{O}\left(n^c \log^k(n)\right)$ *for some* $c \geq 0$ *and* $k \in \mathbb{Z}_{\geq 0}$, *then*

$$\mathrm{cost}_{rec}(n) = \begin{cases} \mathcal{O}(n), & 0 \leq c < 1, \\ \mathcal{O}\left(n \log^{k+1}(n)\right), & c = 1, \\ \mathcal{O}\left(n^c \log^k(n)\right), & c > 1. \end{cases} \tag{21}$$

Thus, we have $\mathrm{cost}_{rec}(n) = \mathcal{O}\left(\mathrm{cost}(n) \log(n) + n\right)$ *in all the cases.*

B. Product formulas

Consider a Hamiltonian given in the LCH form $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$. We can well approximate the evolution under H for a short time using *product formulas* with error high order in time. A longer evolution can then be simulated by repeating the short-time steps. Product formulas provide a simple yet surprisingly efficient approach to quantum simulation. Indeed, recent work has shown that product formulas: (i) can simulate geometrically local lattice systems [88] with nearly optimal gate complexity [89]; (ii) have the lowest asymptotic cost for electronic structure Hamiltonians in second quantization in the plane-wave basis [87]; and (iii) are advantageous for simulating general κ-local Hamiltonians [44] (although our work achieves further improvements regarding this last point). In addition, product formulas have also been widely used in classical simulations of quantum systems and in areas beyond quantum computing [90].

By definition, the cost of the product-formula approach is determined by both the repetition number of short-time steps and the complexity of implementing each step. To elaborate, first consider a simple example where we use the Lie-Trotter formula $e^{-itB}e^{-itA}$ to simulate a two-term Hamiltonian $H = A + B$. Then it holds that

$$e^{-itB}e^{-itA} - e^{-it(A+B)}$$

$$= \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \, e^{-i(t-\tau_1)(A+B)} e^{-i\tau_1 B}$$

$$\times \, e^{i\tau_2 B}\big[iB, iA\big] e^{-i\tau_2 B} e^{-i\tau_1 A}, \tag{22}$$

which implies the *Trotter error* bound

$$\left\| e^{-itB}e^{-itA} - e^{-it(A+B)} \right\| \leq \frac{t^2}{2} \left\| [A, B] \right\|. \tag{23}$$

This approximation error is small for a sufficiently short-time *Trotter step*. For a longer simulation, we apply r repetitions of the same step with time t/r, obtaining

$$\left\| \left(e^{-i(t/r)B} e^{-i(t/r)A} \right)^r - e^{-it(A+B)} \right\|$$

$$\leq r \left\| e^{-i(t/r)B} e^{-i(t/r)A} - e^{-i(t/r)(A+B)} \right\| \leq \frac{t^2}{2r} \left\| [A, B] \right\|. \tag{24}$$

To achieve an error at most ϵ, it thus suffices to take

$$r = \left\lceil \frac{\|[A,B]\| \, t^2}{2\epsilon} \right\rceil. \tag{25}$$

We now have a total of $2r$ elementary exponentials, where r is the number of Trotter steps determined by the error bound Eq. (23).

In general, we can simulate the Hamiltonian $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ using a pth-order product formula $S_p(t)$, where

p is a positive integer that can be arbitrarily large. Similar to above, we need to bound the cost of implementing Trotter steps, as well as the repetition number r, which is in turn determined by the analysis of Trotter error. We introduce the following Trotter error bound with a commutator scaling established by Ref. [44]. Although Trotter error analysis can be further tightened using additional assumptions of the quantum simulation problem [91–99], those improvements are not relevant to our results and are not further discussed here.

Lemma 2: (Trotter error with commutator scaling). *Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a Hamiltonian in the LCH form and $S_p(t)$ be a pth-order product formula with respect to this decomposition. We have*

$$\left\| S_p(t) - e^{-itH} \right\| = \mathcal{O}\left(\sum_{\gamma_1,\gamma_2,\dots,\gamma_{p+1}=1}^{\Gamma} \left\| \left[H_{\gamma_{p+1}}, \dots \left[H_{\gamma_2}, H_{\gamma_1} \right] \right] \right\| t^{p+1} \right). \tag{26}$$

Furthermore, if H is a fermionic Hamiltonian in second quantization and H_γ are number preserving,

$$\left\| S_p(t) - e^{-itH} \right\|_{\mathcal{W}_\eta} = \mathcal{O}\left(\sum_{\gamma_1,\gamma_2,\dots,\gamma_{p+1}=1}^{\Gamma} \left\| \left[H_{\gamma_{p+1}}, \dots \left[H_{\gamma_2}, H_{\gamma_1} \right] \right] \right\|_{\mathcal{W}_\eta} t^{p+1} \right). \tag{27}$$

We denote the sum of nested commutators as

$$\beta_{\text{comm}} := \sum_{\gamma_1,\gamma_2,\dots,\gamma_{p+1}=1}^{\Gamma} \left\| \left[H_{\gamma_{p+1}}, \dots \left[H_{\gamma_2}, H_{\gamma_1} \right] \right] \right\|. \tag{28}$$

Then to simulate for time t with accuracy ϵ, it suffices to take $r = \mathcal{O}\left(\beta_{\text{comm}}^{1/p} t^{1+1/p} / \epsilon^{1/p} \right)$, which simplifies to

$$r = \frac{\beta_{\text{comm}}^{o(1)} t^{1+o(1)}}{\epsilon^{o(1)}} \tag{29}$$

by choosing p sufficiently large. For the fermionic case, we can simply replace the spectral norm $\|\cdot\|$ by its restriction to the Hamming weight-η subspace $\|\cdot\|_{\mathcal{W}_\eta}$.

A sequential implementation of Trotter steps has an asymptotic cost of $\mathcal{O}(\Gamma)$, giving total gate complexity $\mathcal{O}(r\Gamma)$. The purpose of this work is to identify conditions under which the cost of Trotter steps does not explicitly scale with Γ. We present three methods, one based on the block-encoding technique (Sec. III), one based on an average-cost simulation (Sec. IV) and one based on a recursive low-rank decomposition (Sec. V). We review the basic notion of block encoding as well as the related qubitization algorithm in the next subsection.

C. Block encoding and qubitization

Block encodings, together with the related qubitization algorithm, provide a versatile framework for simulating Hamiltonians in the LCU form and beyond. Such techniques enable quantum simulation with a complexity linear in the evolution time and logarithmic in the inverse accuracy [100], and can be extended to solve problems other than simulation [11,25] (although these goals are sometimes also achievable via product formulas as explained in Sec. I A). In addition, these techniques have been utilized by recent work to design classical algorithms for simulating quantum systems [101].

To explain the idea of block encoding and qubitization, we use the notion of isometries. Specifically, consider two isometries $G_0 : \mathcal{G}_0 \to \mathcal{H}$, $G_1 : \mathcal{G}_1 \to \mathcal{H}$ and a unitary $U : \mathcal{H} \to \mathcal{H}$. We say an operator $B : \mathcal{G}_0 \to \mathcal{G}_1$ is *block encoded* by G_0, G_1, and U if

$$B = G_1^\dagger U G_0. \tag{30}$$

Of course, this unitary dilation is mathematically feasible if and only if $\|B\| \leq 1$ [102, 2.7.P2]. However, there are many scenarios where additional normalization factors will be introduced when such block encodings are realized by quantum circuits. If we choose two bases with respect to the orthogonal decompositions $\mathcal{H} = \text{im}(G_0 G_0^\dagger) \oplus \text{im}(I -$

$G_0 G_0^\dagger) = \mathrm{im}(G_1 G_1^\dagger) \oplus \mathrm{im}(I - G_1 G_1^\dagger)$ with $\oplus$ being the orthogonal direct sum, then U has a matrix representation that looks like

$$U = \begin{bmatrix} B & \cdot \\ \cdot & \cdot \end{bmatrix}, \tag{31}$$

where B is exhibited as the top-left matrix block; hence the name *block encoding*.

For quantum simulation, we fix $\mathcal{G}_0 = \mathcal{G}_1 = \mathcal{G}$ to be the target system space and we let $G_1^\dagger U G_0$ be Hermitian to represent physical Hamiltonians. Then, we have the following qubitization algorithm to perform quantum simulation of block-encoded operators [22].

Lemma 3: (Hamiltonian simulation by qubitization). *Let* $G_0, G_1 : \mathcal{G} \to \mathcal{H}$ *be isometries and* $U : \mathcal{H} \to \mathcal{H}$ *be a unitary such that* $G_1^\dagger U G_0$ *is Hermitian. Given a target evolution time t and accuracy ϵ, there exists a unitary* $V_\varphi : \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathcal{H} \to \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathcal{H}$ *parameterized by angles* $\varphi_1, \ldots, \varphi_r$ *such that*

$$\left\| \left(\langle + | \otimes \frac{\langle 0 | \otimes G_0^\dagger + \langle 1 | \otimes G_1^\dagger}{\sqrt{2}} \right) V_\varphi \left(|+\rangle \otimes \frac{|0\rangle \otimes G_0 + |1\rangle \otimes G_1}{\sqrt{2}} \right) - e^{-it G_1^\dagger U G_0} \right\| \leq \epsilon. \tag{32}$$

The number of steps r is an even integer with the asymptotic scaling

$$r = \mathcal{O}\left(t + \log\left(\frac{1}{\epsilon}\right) \right), \tag{33}$$

and V_φ *is explicitly defined as*

$$V_\varphi := \prod_{k=0}^{r/2-1} \left(V_{\varphi_{2k+2}}^\dagger V_{\varphi_{2k+1}} \right),$$

$$V_{\varphi_j} := \left(e^{-i\varphi_j Z/2} \otimes I \right) \left(|+\rangle\langle+| \otimes I + |-\rangle\langle-| \otimes V \right) \left(e^{i\varphi_j Z/2} \otimes I \right),$$

$$V := (X \otimes I) \left(|0\rangle\langle 0| \otimes U + |1\rangle\langle 1| \otimes U^\dagger \right) \left(I - 2 \frac{|0\rangle \otimes G_0 + |1\rangle \otimes G_1}{\sqrt{2}} \frac{\langle 0| \otimes G_0^\dagger + \langle 1| \otimes G_1^\dagger}{\sqrt{2}} \right). \tag{34}$$

Note that by our definition of $\mathcal{O}$, the complexity expression Eq. (33) is valid for t sufficiently large and ϵ sufficiently small. This expression may be modified if other parameter regimes are of interest.

To simulate an n-qubit Hamiltonian $H = \sum_{\gamma=1}^{\Gamma} \beta_\gamma U_\gamma$ in the LCU form, we define

$$G : \mathbb{C}^{2^n} \to \mathbb{C}^\Gamma \otimes \mathbb{C}^{2^n}, \qquad G = \frac{1}{\sqrt{\|\beta\|_1}} \sum_{\gamma=1}^{\Gamma} \sqrt{\beta_\gamma} |\gamma\rangle \otimes I,$$

$$\tag{35}$$

$$U : \mathbb{C}^\Gamma \otimes \mathbb{C}^{2^n} \to \mathbb{C}^\Gamma \otimes \mathbb{C}^{2^n}, \qquad U = \sum_{\gamma=1}^{\Gamma} |\gamma\rangle\langle\gamma| \otimes U_\gamma,$$

which implies

$$G^\dagger U G : \mathbb{C}^{2^n} \to \mathbb{C}^{2^n}, \qquad G^\dagger U G = \frac{H}{\|\beta\|_1}. \tag{36}$$

We thus see that the target Hamiltonian is block encoded by G (a state *preparation subroutine*) and U (an operator *selection subroutine*), but with a normalization factor that depends on the vector 1-norm $\|\beta\|_1$, which slows down quantum simulation. To approximate e^{-itH}, we then need to invoke Lemma 3 with an effective evolution time of

$\|\beta\|_1 t$. This implies that the number of steps will now scale as

$$r = \mathcal{O}\left(\|\beta\|_1 t + \log\left(\frac{1}{\epsilon}\right)\right). \qquad (37)$$

The preparation of a Γ-dimensional state naively requires $\mathcal{O}(\Gamma)$ gates to implement [103] and the selection of Γ terms has cost $\mathcal{O}(\Gamma)$ [3, Appendix G.4], assuming each controlled-U_γ has constant cost. This gives a total complexity of $\mathcal{O}(r\Gamma)$ for qubitization.

Having introduced the qubitization algorithm, we now make a few remarks about its circuit implementation.

First, qubitization approximately simulates the time evolution with a certain probability, and both the approximation error and the success probability can be analyzed using the condition Eq. (32). Indeed, an argument based on the triangle inequality shows that the qubitization algorithm succeeds with probability at least $(1-\epsilon)^2 \geq 1-2\epsilon$, and the postmeasurement state has an error at most $\sqrt{1+\epsilon} - \sqrt{1-\epsilon} \leq \sqrt{2\epsilon}$ [3, Lemma G.4 and Appendix H.1]. It therefore suffices to adjust the value of ϵ to meet the desired error and probability requirements.

Second, in our above LCU example, we define G and U that faithfully block encode $H/\|\beta\|_1$. More realistically, we may consider state preparation and operator selection circuits that have lower cost but introduce error to the block encoding. Specifically, suppose $\widetilde{G}^\dagger \widetilde{U} \widetilde{G} = \widetilde{H}/\|\widetilde{\beta}\|_1$, then we are effectively block encoding an erroneous Hamiltonian with error growing at most linearly in time: $\left\|e^{-it\|\widetilde{\beta}\|_1 \widetilde{G}^\dagger \widetilde{U} \widetilde{G}} - e^{-itH}\right\| \leq t\|H - \widetilde{H}\|$. And we can suppress this error accordingly by increasing the accuracy of block encoding.

We focus on the complexity of the state preparation and operator selection subroutines for the remainder of our paper. There is also an additional cost in qubitization to implement the operations in between these subroutines, but that cost is typically logarithmic in Γ and makes only a mild contribution to the total gate complexity. In particular, one can check that this holds for our block encodings of power-law Hamiltonians to be described in Sec. III and Sec. IV. So there is no loss of rigor to ignore this subdominant contribution in our analysis.

Finally, we point out that a host of techniques have been recently developed to improve the circuit implementation of block encoding and qubitization over the naive approach. For instance, if the target Hamiltonian consists of tensor product of Pauli operators, then the selection subroutine $U = U^\dagger$ is both unitary and Hermitian, so we can reduce 1 ancilla from the qubitization circuit in Lemma 3. Even when this Hermitian condition is not satisfied, one may still improve the implementation of the bidirectional control $|0\rangle\langle 0| \otimes U + |1\rangle\langle 1| \otimes U^\dagger$ using explicit structures of U. Also, the qubitization algorithm we introduce in Lemma 3 is mainly for dynamical simulation. When

qubitization is used as a subroutine in quantum phase estimation, we need only to implement the operator V in Eq. (34) as opposed to the full circuit [104–106]. There are also some flexibilities about block encoding the LCU Hamiltonian beyond the standard approach of Eq. (35), such as allowing garbage registers [107], performing unary encodings [21], or using even more advanced circuit compilation tricks [108]. Additionally, given a block encoding $G_0, G_1 : \mathcal{G} \to \mathcal{H}$ and $U : \mathcal{H} \to \mathcal{H}$, there always exists unitary $W : \mathcal{H} \to \mathcal{H}$ such that $WG_1 = G_0$. Any such W satisfies $G_1^\dagger U G_0 = G_0^\dagger (WU) G_0$ and hence gives an alternative block encoding with only one isometry, simplifying the circuit implementation.

But perhaps the most significant improvement comes from the following observation: terms in the LCU Hamiltonian are sometimes indexed by vectors with a product structure, i.e.,

$$U_\gamma = U_{\gamma_1} \cdots U_{\gamma_s}, \qquad \gamma_j = 1, \dots, \Gamma^{1/s}. \qquad (38)$$

In this case, the selection subroutine takes a product form

$$\sum_{\gamma=1}^{\Gamma} |\gamma\rangle\langle\gamma| \otimes U_\gamma$$

$$= \sum_{\gamma_1, \dots, \gamma_s = 1}^{\Gamma^{1/s}} |\gamma_1 \cdots \gamma_s\rangle\langle\gamma_1 \cdots \gamma_s| \otimes U_{\gamma_1} \cdots U_{\gamma_s} \qquad (39)$$

and can be implemented with complexity $\mathcal{O}(s\Gamma^{1/s})$ as opposed to $\mathcal{O}(\Gamma)$. Meanwhile, we can sometimes exploit the structure of Hamiltonian coefficients to also perform the preparation subroutine faster than the naive approach. This then yields a step of qubitization with complexity sublinearly in the number of terms Γ.

However, the above discussion ignored the scaling of the normalization factor $\|\beta\|_1$, which determines the number of repetition steps and is still the bottleneck of qubitization. In fact, the reduction of the 1-norm has been the central technical problem studied by many recent work of quantum simulation [32,33,109,110]. We show in Sec. III that this 1-norm can be significantly reduced using product formulas "almost for free," which leads to new methods for simulating power-law Hamiltonians that improve the state of the art.

D. Volume of diagonal unitaries

In this section, we review techniques for computing the volume of a parameterized object in a high-dimensional Euclidean space. In particular, we compute the volume of diagonal unitaries with phase angles taking a continuum range of values between $-\theta_{\max}$ and $\theta_{\max}$, which will be used later in Sec. VII to prove our circuit lower bound. These calculations can be seen as generalizations of line and surface integrals defined for low-dimensional spaces.

Let $\mathcal{M} \subset \mathbb{R}^m$ be a subset and $f : \mathcal{M} \to \mathbb{R}^q$ ($m \leq q$) be a mapping. Here, $f = [f_1 \ \cdots \ f_q]$ consists of q scalar functions and we can think of it as a parameterized description of a high-dimensional geometrical object, where $\mathcal{M}$ is the set of all possible parameters. Then, given a subset of parameters $\mathcal{A} \subset \mathcal{M}$, we may compute the volume of $f(\mathcal{A})$ as

$$\text{vol}(f(\mathcal{A})) = \int_{\mathcal{A}} d\theta \sqrt{\det(G_f)}, \tag{40}$$

where G_f is the Gramian whose element in the xth row and yth column is given by the inner product

$$\begin{bmatrix} \frac{\partial f_1}{\partial \theta_x} & \cdots & \frac{\partial f_q}{\partial \theta_x} \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial \theta_y} \\ \vdots \\ \frac{\partial f_q}{\partial \theta_y} \end{bmatrix}. \tag{41}$$

$$G_f = \begin{bmatrix} -\sin\theta_0 & \cos\theta_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -\sin\theta_1 & \cos\theta_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\sin\theta_{m-1} & \cos\theta_{m-1} \end{bmatrix} \begin{bmatrix} -\sin\theta_0 & 0 & \cdots & 0 \\ \cos\theta_0 & 0 & \cdots & 0 \\ 0 & -\sin\theta_1 & \cdots & 0 \\ 0 & \cos\theta_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -\sin\theta_{m-1} \\ 0 & 0 & \cdots & \cos\theta_{m-1} \end{bmatrix} = I,$$
$$\tag{44}$$

and as a result we obtain the volume formula

$$\text{vol}\left(\mathcal{D}_{\theta_{\max}}\right) = (2\theta_{\max})^m. \tag{45}$$

Alternatively, we may also consider the mapping

$$\varphi : \ \theta = [\theta_0 \ \cdots \ \theta_{m-1}] \mapsto \sum_{x=0}^{m-1} e^{i\theta_x}|x\rangle\langle x| \tag{46}$$

that parameterizes all diagonal unitaries with phase angles $-\pi \leq \theta_x < \pi$. This is a bijection, and we may define the pushforward measure on the set of diagonal unitaries using the Lebesgue measure λ as

$$\varphi_*(\lambda)(\mathcal{B}) = \lambda(\varphi^{-1}(\mathcal{B})). \tag{47}$$

The measure $\varphi_*(\lambda)$ can be verified to be invariant under multiplication and, up to a constant factor, is the unique Haar measure for the group of diagonal unitaries. This then gives the volume $\varphi_*(\lambda)\left(\mathcal{D}_{\theta_{\max}}\right) = (2\theta_{\max})^m$ the same as our previous formula.

For the above formula to hold mathematically, we require that $\mathcal{M}$ is open, $\mathcal{A}$ is Lebesgue measurable, and f is continuously differentiable and injective on $\mathcal{A}$. We refer the reader to standard analysis textbooks for a detailed explanation of this formula [111,112].

Now, consider the set of diagonal unitaries

$$\mathcal{D}_{\theta_{\max}} := \left\{ \sum_{x=0}^{m-1} e^{i\theta_x}|x\rangle\langle x|, \ |\theta_x| \leq \theta_{\max} < \pi \right\}. \tag{42}$$

We may view them as objects in the Euclidean space $\mathbb{R}^{2m}$ by choosing the parameterization

$$f(\theta_0, \ldots, \theta_{m-1})$$
$$= \begin{bmatrix} \cos\theta_0 & \sin\theta_0 & \cdots & \cos\theta_{m-1} & \sin\theta_{m-1} \end{bmatrix}. \tag{43}$$

Then, the Gramian matrix is simply the identity matrix

III. FASTER TROTTER STEPS USING BLOCK ENCODING

We now introduce our first method to perform faster Trotter steps based on block encoding. The structure of this section is as follows. We describe how to reduce the 1-norm of Hamiltonian terms using product formulas in Sec. III A, removing the bottleneck of the qubitization algorithm. We then give a circuit implementation of the preparation and selection subroutines in Sec. III B, assuming that the Hamiltonian coefficients are efficiently computable. This gives simulations with complexity nearly linear in the spacetime volume for power-law exponent $\alpha \geq 2$. Readers who wish to see the overall algorithm may skip ahead to Sec. III C.

A. Reducing the 1-norm of Hamiltonian coefficients

Consider an n-qubit Hamiltonian $H = \sum_{1 \leq j < k \leq n} H_{j,k}$, where $H_{j,k}$ are two-local terms acting nontrivially only on sites j and k with norm $\|H_{j,k}\| \leq 1/|j-k|^\alpha$. Our aim is to implement Trotter steps using block encodings and qubitization. To this end, we rewrite the Hamiltonian in the Pauli basis:

$$H = \sum_{1\leq j < k \leq n} \beta_{j,k}^{(xx)} X_j X_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(xy)} X_j Y_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(xz)} X_j Z_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(yx)} Y_j X_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(yy)} Y_j Y_k$$

$$+ \sum_{1\leq j < k \leq n} \beta_{j,k}^{(yz)} Y_j Z_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(zx)} Z_j X_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(zy)} Z_j Y_k + \sum_{1\leq j < k \leq n} \beta_{j,k}^{(zz)} Z_j Z_k + \sum_{1\leq j \leq n} \beta_j^{(x)} X_j$$

$$+ \sum_{1\leq j \leq n} \beta_j^{(y)} Y_j + \sum_{1\leq j \leq n} \beta_j^{(z)} Z_j + \beta^{(i)} I, \tag{48}$$

where the coefficients have the scaling [44, Theorem H.1]

$$\beta_{j,k}^{(\sigma,\sigma')} = \mathcal{O}\left(\frac{1}{|j-k|^\alpha}\right), \quad \beta_j^{(\sigma)} = \begin{cases} \mathcal{O}(1), & \alpha > 1, \\ \mathcal{O}(\log n), & \alpha = 1, \\ \mathcal{O}(n^{1-\alpha}), & 0 < \alpha < 1, \end{cases} \qquad \beta^{(i)} = \begin{cases} \mathcal{O}(n), & \alpha > 1, \\ \mathcal{O}(n\log n), & \alpha = 1, \\ \mathcal{O}(n^{2-\alpha}), & 0 < \alpha < 1, \end{cases} \tag{49}$$

for $\sigma \neq \sigma' \in \{x,y,z\}$. Since $\beta^{(i)} I$ commutes with all the other terms and its evolution only introduces a global phase, we may separate out this term without introducing error. We then use product formulas to make a coarse-grained decomposition of the remaining evolution into exponentials of 12 groups of terms. The commutator norm Eq. (28) corresponding to this decomposition has the scaling [44, Theorem H.2]

$$\beta_{\mathrm{comm}}^{o(1)} = \begin{cases} n^{o(1)}, & \alpha \geq 1, \\ n^{1-\alpha+o(1)}, & 0 < \alpha < 1, \end{cases} \tag{50}$$

which implies

$$r = \begin{cases} \dfrac{n^{o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}, & \alpha \geq 1, \\[2ex] \dfrac{n^{1-\alpha+o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < 1. \end{cases} \tag{51}$$

In particular, the Trotter error bound corresponding to this coarse-grained decomposition is never larger than that of the fine-grained decomposition [in which all $\Theta(n^2)$ terms are split].

Note that all on-site terms from the same group pair-wise commute and can be exponentiated with complexity $\mathcal{O}(n)$. So without loss of generality, we consider only Hamiltonians of the form [113]

$$H = \sum_{1\leq j < k \leq n} \beta_{j,k} X_j Y_k. \tag{52}$$

The coefficients of this Hamiltonian have vector 1-norm scaling like [44, Lemma H.1]

$$\|\beta\|_1 = \sum_{1\leq j < k \leq n} |\beta_{j,k}| = \begin{cases} \mathcal{O}(n), & \alpha > 1, \\ \mathcal{O}(n\log n), & \alpha = 1, \\ \mathcal{O}(n^{2-\alpha}), & 0 < \alpha < 1. \end{cases} \tag{53}$$

For $\alpha > 1$, even if we use an improved qubitization circuit with product structure like in Eq. (39), we still need a cost roughly

$$\underbrace{n}_{\text{block encoding}} \cdot \underbrace{n\left(\frac{t}{r}\right)^{o(1)}}_{\substack{\text{effective time} \\ \text{Trotter step complexity}}} = \underbrace{n^2 \left(\frac{t}{r}\right)^{o(1)}}_{} \tag{54}$$

to simulate for an almost constant time, to implement a single Trotter step. Thus, block encoding does not seem to offer a significant benefit over the sequential approach. In what follows, we show that product formulas can be used to reduce the 1-norm of Hamiltonian coefficients, which resolves this technical obstacle and, when combined with block encoding, provide a fast implementation of Trotter steps.

To simplify the discussion, we assume that the system size n is a power of 2. We also use

$$H_{[j,k]} := \sum_{j \leq u < v \leq k} \beta_{u,v} X_u Y_v \; (1 \leq j < k \leq n),$$

$$H_{[j,k]:[l,m]} := \sum_{\substack{j \leq u \leq k \\ l \leq v \leq m}} \beta_{u,v} X_u Y_v \; (1 \leq j \leq k < l \leq m \leq n) \tag{55}$$

to represent terms within a specific interval and across two disjoint intervals of sites. Then, we define a decomposition via the recurrence relation

$$H_{[j,k]} = H_{[j,\lfloor \frac{j+k}{2}\rfloor]:[\lfloor \frac{j+k}{2}\rfloor+1,k]} + H_{[j,\lfloor \frac{j+k}{2}\rfloor]} + H_{[\lfloor \frac{j+k}{2}\rfloor+1,k]}, \tag{56}$$

which unwraps to

$$
\begin{aligned}
H &= H_{[1,n]} \\
&= H_{[1,n/2]:[n/2+1,n]} + H_{[1,n/2]} + H_{[n/2+1,n]} \\
&= H_{[1,n/2]:[n/2+1,n]} + H_{[1,n/4]:[n/4+1,n/2]} \\
&\quad + H_{[n/2+1,3n/4]:[3n/4+1,n]} + H_{[1,n/4]} \\
&\quad + H_{[n/4+1,n/2]} + H_{[n/2+1,3n/4]} + H_{[3n/4+1,n]} \\
&= \cdots \\
&= \sum_{\ell=1}^{\log n-1} \sum_{b=0}^{2^{\ell-1}-1} \\
&\quad H_{[2b(n/2^\ell)+1,(2b+1)n/2^\ell]:[(2b+1)n/2^\ell+1,2(b+1)n/2^\ell]}. \quad (57)
\end{aligned}
$$

Here, the variable ℓ labels the layer of the decomposition and b indexes the pairs of neighboring intervals within layer ℓ. See Fig. 1 for an illustration of this recursion for $\ell = 1, 2, 3$. Correspondingly, we apply product formulas to decompose the entire evolution into exponentials of each individual $H_{[2b(n/2^\ell)+1,(2b+1)n/2^\ell]:[(2b+1)n/2^\ell+1,2(b+1)n/2^\ell]}$. Similar to the above analysis, this coarse-grained decomposition gives a Trotter error bound no larger than that of the fine-grained decomposition and implies the scaling Eq. (51) as well.

To proceed, we take a closer look at this recursive decomposition and collect some of its features below.

(i) There are $\log n - 1 = \mathcal{O}(\log n)$ layers in the decomposition, all indexed by ℓ.

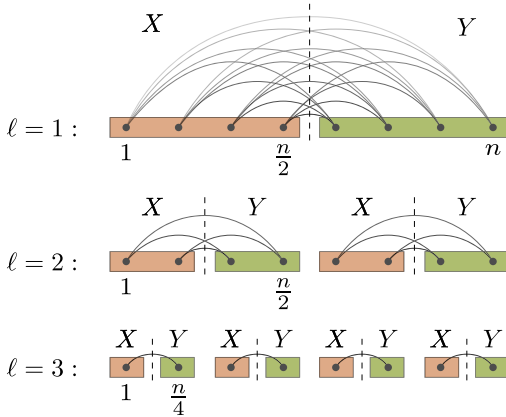(ii) For a fixed layer ℓ, there are 2^ℓ consecutive intervals each of length $n/2^\ell$.

FIG. 1. Illustration of the recursive decomposition of $\sum_{j<k} \beta_{j,k} X_j Y_k$ with three layers. The black dots represent qubits and the curves correspond to the interactions between them. In each layer, we cut the system into equal halves (dashed lines) and consider only the interactions across the cuts. The colors indicate whether the interaction consists of Pauli X (shaded orange) or Pauli Y (green).

(iii) Within each layer ℓ, even and odd intervals are further grouped into $2^\ell - 1$ pairs (indexed by b). Terms across intervals from the same pair are denoted by $H_{[2b(n/2^\ell)+1,(2b+1)n/2^\ell]:[(2b+1)n/2^\ell+1,2(b+1)n/2^\ell]}$. The total number of pairs of intervals is $\mathcal{O}(n)$ by the master theorem Lemma 1.

(iv) When $\ell = 1, \ldots, \log n - 1$ and $b = 0, \ldots, 2^{\ell-1} - 1$, $H_{[2b(n/2^\ell)+1,(2b+1)n/2^\ell]:[(2b+1)n/2^\ell+1,2(b+1)n/2^\ell]}$ provide a partition of all the terms in the original Hamiltonian H.

We may use these features to further simplify our discussion. For instance, since we have only two-local terms of the same Pauli-type acting across disjoint intervals, we can simultaneously change the basis and consider only Pauli-Z interactions without loss of generality:

$$
\begin{aligned}
&H_{[2b(n/2^\ell)+1,(2b+1)n/2^\ell]:[(2b+1)n/2^\ell+1,2(b+1)n/2^\ell]} \\
&= \sum_{\substack{2b(n/2^\ell)+1\leq u\leq(2b+1)n/2^\ell \\ (2b+1)n/2^\ell+1\leq v\leq2(b+1)n/2^\ell}} \beta_{u,v} Z_u Z_v. \quad (58)
\end{aligned}
$$

Also, due to the nature of the recursive decomposition, it suffices to focus on implementing a specific $H_{[2b(n/2^\ell)+1,(2b+1)n/2^\ell]:[(2b+1)n/2^\ell+1,2(b+1)n/2^\ell]}$: we have similar complexities cost $(\cdot)$ for all pairs of intervals from different layers, so the total complexity

$$
\text{cost}_{\text{rec}}(n) = \sum_{\ell=1}^{\log n-1} 2^{\ell-1} \text{cost}\left(\frac{n}{2^{\ell-1}}\right) \quad (59)
$$

can be immediately bounded by the master theorem Lemma 1. For simplicity, we choose $\ell = 1$, $b = 0$ and we study the complexity cost(n) of exponentiating $H_{[1,n/2]:[n/2+1,n]}$.

We note that prior work has utilized the above features of the recursive decomposition to develop measurement schedules for local quantum observables [114]. For the purpose of quantum simulation, we however need an additional key observation.

(i) Coefficients of $H_{[1,n/2]:[n/2+1,n]}$ have the vector 1-norm scaling:

$$
\|\beta\|_1 = \sum_{\substack{1\leq u\leq n/2 \\ n/2+1\leq v\leq n}} |\beta_{u,v}| = \begin{cases} \mathcal{O}(1), & \alpha > 2, \\ \mathcal{O}(\log n), & \alpha = 2, \\ \mathcal{O}(n^{2-\alpha}), & 0 < \alpha < 2. \end{cases} \quad (60)
$$

Comparing Eq. (60) with Eq. (53), the coefficients of $H_{[1,n/2]:[n/2+1,n]}$ have a significantly smaller 1-norm than that of the full Hamiltonian H. In particular, the 1-norm in Eq. (60) is asymptotically constant for $\alpha > 2$,

a factor of n smaller than that of the original Hamiltonian. Therefore, exponentiating $H_{[1,n/2]:[n/2+1,n]}$ would also take a factor of n fewer quantum gates than exponentiating H. Meanwhile, the gate complexity of block encoding remains asymptotically the same: we show in Sec. III B that block encoding can be implemented for the pair of intervals $[1, n/2]$ and $[n/2 + 1, n]$ with $\text{cost}(n) = \mathcal{O}(n \, \text{polylog}(nt/\epsilon))$, which gives the total cost by applying the master theorem to Eq. (59). This leads to the desired faster Trotter steps for simulating power-law Hamiltonians.

B. Preparation and selection subroutines

In the previous subsection, we reduce the 1-norm of power-law Hamiltonians by applying a recursive decomposition using product formulas. For completeness, we now give an explicit circuit for block encoding and apply qubitization to simulate the decomposed Hamiltonian.

To be precise, our goal is to simulate the following Hamiltonian acting across the intervals $[1, n/2]$ and $[n/2 + 1, n]$:

$$H_{[1,n/2]:[n/2+1,n]} = \sum_{\substack{1 \leq u \leq n/2 \\ n/2+1 \leq v \leq n}} \beta_{u,v} Z_u Z_v, \qquad (61)$$

where the coefficients $|\beta_{u,v}| \leq 1/|u - v|^\alpha$. Additionally, we assume that the coefficients are *logarithmically*

computable, meaning the oracle

$$O_\beta |u, v, 0\rangle = |u, v, \beta_{u,v}\rangle \qquad (62)$$

can be implemented with cost $\mathcal{O}(\text{polylog}(nt/\epsilon))$ [115]. Note that by making this assumption, we implicitly assume certain structural properties of the Hamiltonian coefficients. For example, when $\beta_{u,v} = 1/|u - v|^\alpha$ decays exactly as a power law, the oracle O_β can be implemented with cost $\mathcal{O}(\text{polylog}(nt/\epsilon))$, fulfilling this requirement. This can be achieved by performing elementary arithmetics such as subtraction, multiplication, and division on the binary representation of u and v [which has length $\mathcal{O}(\log(nt/\epsilon))$]. However, when Hamiltonian coefficients are arbitrarily given, we prove a lower bound in Appendix A showing that at least approximately n^2 gates are required to implement O_β in the circuit model.

In what follows, we analyze the number of queries to O_β as well as additional interquery gates. We aim to block encode $H_{[1,n/2]:[n/2+1,n]}$ with its 1-norm scaling close to Eq. (60). As suggested in Sec. II C, we may simply choose the selection subroutine to have the product structure

$$\sum_{\substack{1 \leq u \leq n/2 \\ n/2+1 \leq v \leq n}} |uv\rangle\langle uv| \otimes Z_u Z_v. \qquad (63)$$

However, some extra efforts are needed to define the preparation subroutine. In particular, the following black-box state-preparation subroutine

$$\frac{1}{n/2} \sum_{\substack{1 \leq u \leq \frac{n}{2} \\ \frac{n}{2}+1 \leq v \leq n}} |u\rangle|v\rangle \overset{O_\beta}{\mapsto} \frac{1}{n/2} \sum_{\substack{1 \leq u \leq n/2 \\ n/2+1 \leq v \leq n}} |u\rangle|v\rangle|\beta_{u,v}\rangle$$

$$\mapsto \frac{1}{n/2} \sum_{\substack{1 \leq u \leq n/2 \\ n/2+1 \leq v \leq n}} |u\rangle|v\rangle|\beta_{u,v}\rangle \left(\sqrt{|\beta_{u,v}|}|0\rangle + \sqrt{1 - |\beta_{u,v}|}|1\rangle \right)$$

$$\overset{O_\beta^\dagger}{\mapsto} \frac{1}{n/2} \sum_{\substack{1 \leq u \leq n/2 \\ n/2+1 \leq v \leq n}} \left(\sqrt{|\beta_{u,v}|}|u\rangle|v\rangle|0\rangle + \sqrt{1 - |\beta_{u,v}|}|\psi_\perp\rangle|1\rangle \right) \qquad (64)$$

does not work since it enlarges the 1-norm of block encoding by a factor of $n/2$, preventing us from achieving the scaling in Eq. (60).

The issue with the naive preparation subroutine is that we prepare uniform superposition states at the beginning which ignore the power-law decaying pattern of $\beta_{u,v}$. To address this issue, we group the coefficients into multiple

boxes. We take the number of boxes to be logarithmic in the system size n so we can prepare superposition states over these boxes efficiently. Meanwhile, we ensure that the coefficients within each box are approximately uniform so we can again prepare the corresponding superposition state efficiently. This idea was previously used for the efficient block encoding of the plane-wave-basis

electronic structure Hamiltonian in first quantization [116]. Here, we describe a variant of this technique to efficiently block-encode power-law Hamiltonians.

To simplify the notation, we shift the lattice and consider the rectangle region $(-n/2, -1] \times [1, n/2)$. We then divide it into boxes

$$\mathcal{B}_{\mu,\nu} := \left\{ (u,v)| -2^{\mu+1} < u \le -2^\mu, 2^\nu \le v < 2^{\nu+1} \right\} \tag{65}$$

for $\mu, \nu = 0, 1, \ldots, \log n - 2$. With the restricted max norm

$$\|\beta\|_{\max, \mathcal{B}_{\mu,\nu}} = \max_{\substack{-2^{\mu+1} < u \le -2^\mu \\ 2^\nu \le v < 2^{\nu+1}}} |\beta_{u,v}|, \tag{66}$$

we describe a state-preparation subroutine whose complexity depends on the ratio between $\|\beta\|_1$ and

$$\|\beta\|_{1,\text{box}} := \sum_{\mu,\nu=0}^{\log n - 2} 2^{\mu+\nu} \|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}. \tag{67}$$

Thus, our approach is efficient if the coefficients are almost uniformly distributed within each box. The boundary terms can be handled separately without changing the asymptotic complexity scaling.

We start by preparing the state

$$\frac{1}{\sqrt{\|\beta\|_{1,\text{box}}}} \sum_{\mu,\nu=0}^{\log n - 2} \sqrt{2^{\mu+\nu} \|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}} |e_\mu\rangle |e_\nu\rangle, \tag{68}$$

where we use unary encoding to represent μ and ν. This state can be prepared using $\mathcal{O}(\log^2 n)$ gates. Conditioned on the unary value of μ and ν, we prepare the uniform superposition

$$\frac{1}{\sqrt{\|\beta\|_{1,\text{box}}}} \sum_{\mu,\nu=0}^{\log n - 2} \sqrt{2^{\mu+\nu} \|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}} |e_\mu\rangle |e_\nu\rangle \frac{1}{\sqrt{2^\mu}}$$

$$\otimes \sum_{-2^{\mu+1} < u \le -2^\mu} |u\rangle \frac{1}{\sqrt{2^\nu}} \sum_{2^\nu \le v < 2^{\nu+1}} |v\rangle, \tag{69}$$

where we represent u and v in binary. We can achieve this by first transforming $|e_\mu\rangle$ and $|e_\nu\rangle$ into run-length unary representation, and performing a sequence of $\mathcal{O}(\log n)$ controlled Hadamard gates to generate the superposition.

We now introduce an auxiliary uniform superposition state

$$\frac{1}{\sqrt{\|\beta\|_{1,\text{box}}}} \sum_{\mu,\nu=0}^{\log n - 2} \sqrt{2^{\mu+\nu} \|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}} |e_\mu\rangle |e_\nu\rangle \frac{1}{\sqrt{2^{\mu+\nu}}}$$

$$\otimes \sum_{\substack{-2^{\mu+1} < u \le -2^\mu \\ 2^\nu \le v < 2^{\nu+1}}} |u\rangle |v\rangle \frac{1}{\sqrt{\Xi}} \sum_{\xi=0}^{\Xi-1} |\xi\rangle \tag{70}$$

and test the inequality

$$\frac{\xi}{\Xi} < \frac{|\beta_{u,v}|}{\|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}}. \tag{71}$$

This can be performed with one query to the coefficient oracle O_β and mild uses of other gates as follows. We first apply O_β to load the coefficients $\beta_{u,v}$. We also load the values of $\|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}$ via the transformation

$$|e_\mu, e_\nu, 0\rangle \mapsto |e_\mu, e_\nu, \|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}\rangle. \tag{72}$$

Note that there are only $\mathcal{O}(\log^2 n)$ different values of $\|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}$, each of which can be represented using at most $\mathcal{O}(\log(nt/\epsilon))$ bits, so we can efficiently perform this transformation. The inequality test can then be performed with a complexity scaling with the maximum size of the inputs. Assuming Ξ is sufficiently large, the state after the inequality test (omitting the failure part and garbage registers) becomes

$$\frac{1}{\sqrt{\|\beta\|_{1,\text{box}}}} \sum_{\mu,\nu=0}^{\log n - 2} \sqrt{2^{\mu+\nu} \|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}} |e_\mu\rangle |e_\nu\rangle \frac{1}{\sqrt{2^{\mu+\nu}}}$$

$$\otimes \sum_{\substack{-2^{\mu+1} < u \le -2^\mu \\ 2^\nu \le v < 2^{\nu+1}}} \sqrt{\frac{|\beta_{u,v}|}{\|\beta\|_{\max, \mathcal{B}_{\mu,\nu}}}} |u\rangle |v\rangle$$

$$= \frac{1}{\sqrt{\|\beta\|_{1,\text{box}}}} \sum_{\mu,\nu=0}^{\log n - 2} \sum_{\substack{-2^{\mu+1} < u \le -2^\mu \\ 2^\nu \le v < 2^{\nu+1}}} \sqrt{|\beta_{u,v}|} |e_\mu\rangle |e_\nu\rangle |u\rangle |v\rangle. \tag{73}$$

In practice, we choose a finite value of Ξ, which results in an erroneous block encoding that can be analyzed as in our second remark in Sec. II C. To achieve an overall error of at most ϵ, it suffices to set $\Xi = \text{poly}(nt/\epsilon)$. Under this nested-boxes representation, the selection subroutine

becomes

$$\sum_{\mu,\nu=0}^{\log n-2} \sum_{\substack{-2^{\mu+1}<u\leq-2^{\mu} \\ 2^{\nu}\leq v<2^{\nu+1}}} |e_\mu e_\nu\rangle\langle e_\mu e_\nu| \otimes |uv\rangle\langle uv| \otimes Z_{\mu,u}Z_{\nu,v}.$$

$$(74)$$

Again, this operation has a product structure and can be implemented with gate complexity $\mathcal{O}(n\log n)$ per Sec. II C.

It is clear that we have prepared a state proportional to the desired state for block encoding when all $\beta_{u,v} > 0$. If $\beta_{u,v} < 0$, we simply introduce a minus sign in the implementation of the selection subroutine. The success probability is given by

$$\frac{\sum_{\mu,\nu=0}^{\log n-2} \sum_{\substack{-2^{\mu+1}<u\leq-2^{\mu} \\ 2^{\nu}\leq v<2^{\nu+1}}} |\beta_{u,v}|}{\|\beta\|_{1,\text{box}}} = \frac{\|\beta\|_1}{\|\beta\|_{1,\text{box}}}. \qquad (75)$$

To boost this success probability to close to 1, we can perform $\mathcal{O}(\sqrt{\|\beta\|_{1,\text{box}}/\|\beta\|_1})$ steps of amplitude amplification. In other words, the complexity of state preparation will indeed depend on the ratio between $\|\beta\|_{1,\text{box}}$ and $\|\beta\|_1$ as previously claimed.

If the distribution of Hamiltonian coefficients exactly matches a power law, we have

$$\|\beta\|_1 = \sum_{\substack{-n/2<u\leq-1 \\ 1\leq v<n/2}} |\beta_{u,v}| = \begin{cases} \Theta(1), & \alpha > 2, \\ \Theta(\log n), & \alpha = 2, \\ \Theta(n^{2-\alpha}), & 0 < \alpha < 2. \end{cases}$$

$$(76)$$

Furthermore, $\|\beta\|_1$ differs from $\|\beta\|_{1,\text{box}}$ by at most a constant factor. This is because

$$\sum_{\mu,\nu=0}^{\log n-2} 2^{\mu+\nu} \min_{\substack{-2^{\mu+1}<u\leq-2^{\mu} \\ 2^{\nu}\leq v<2^{\nu+1}}} |\beta_{u,v}|$$

$$\leq \sum_{\mu,\nu=0}^{\log n-2} \sum_{\substack{-2^{\mu+1}<u\leq-2^{\mu} \\ 2^{\nu}\leq v<2^{\nu+1}}} |\beta_{u,v}| = \|\beta\|_1$$

$$\leq \|\beta\|_{1,\text{box}} = \sum_{\mu,\nu=0}^{\log n-2} 2^{\mu+\nu} \max_{\substack{-2^{\mu+1}<u\leq-2^{\mu} \\ 2^{\nu}\leq v<2^{\nu+1}}} |\beta_{u,v}|, \qquad (77)$$

where the first and last quantities differ termwise by at most a factor of $2^\alpha = \mathcal{O}(1)$. Therefore, the number of amplitude amplification steps is constant. The cost of the preparation subroutine is then dominated by the query to the oracle O_β, which costs $\mathcal{O}(\text{polylog}(nt/\epsilon))$ to implement by our assumption.

We now use qubitization (Lemma 3) to simulate $H_{[1,n/2]:[n/2+1,n]}$ for time t/r with accuracy $\mathcal{O}(\epsilon/(rn))$, where

$$r = \begin{cases} \dfrac{n^{o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & \alpha \geq 1, \\[3mm] \dfrac{n^{1-\alpha+o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < 1, \end{cases} \qquad (78)$$

with the selection and preparation subroutine defined above. We bound the gate complexity $\text{cost}(n)$ as follows:

$$\text{cost}(n) = \underbrace{\mathcal{O}\left(\|\beta\|_1\frac{t}{r} + \log\left(\frac{rn}{\epsilon}\right)\right)}_{\text{number of qubitization steps}} \cdot \left(\underbrace{\mathcal{O}(n\log n)}_{\text{selection}} + \underbrace{\mathcal{O}\left(\sqrt{\frac{\|\beta\|_{1,\text{box}}}{\|\beta\|_1}}n\,\text{polylog}(nt/\epsilon)\right)}_{\text{preparation}}\right)$$

$$= \begin{cases} \mathcal{O}\left(\left(\frac{t}{r}+1\right)n\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & \alpha \geq 2, \\[3mm] \mathcal{O}\left(\left(n^{2-\alpha}\frac{t}{r}+1\right)n\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & 0 < \alpha < 2. \end{cases} \qquad (79)$$

By using the same evolution and target accuracy with different values of n and invoking the master theorem (Lemma 1), we obtain the gate complexity of implementing one Trotter step

$$\text{cost}_{\text{rec}}(n) = \sum_{\ell=1}^{\log n-1} 2^{\ell-1}\text{cost}\left(\frac{n}{2^{\ell-1}}\right) = \begin{cases} \mathcal{O}\left(\left(\frac{t}{r}+1\right)n\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & \alpha \geq 2, \\[3mm] \mathcal{O}\left(\left(n^{2-\alpha}\frac{t}{r}+1\right)n\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & 0 < \alpha < 2. \end{cases} \qquad (80)$$

This then gives the complexity of the entire quantum simulation:

$$
\begin{aligned}
&nt\left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq 2,\\
&\mathcal{O}\left(n^{3-\alpha}t\,\mathrm{polylog}\left(\tfrac{nt}{\epsilon}\right)\right) + nt\left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & 1 \leq \alpha < 2,\\
&\mathcal{O}\left(n^{3-\alpha}t\,\mathrm{polylog}\left(\tfrac{nt}{\epsilon}\right)\right) + n^{2-\alpha}t\left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & 0 < \alpha < 1.
\end{aligned}
\tag{81}
$$

In general, the complexity of block encoding will depend on the closeness of distribution of Hamiltonian coefficients to a power-law distribution. To quantify this, we let λ_{block} be the maximum ratio between $\|\beta\|_{1,\mathrm{box}}$ and $\|\beta\|_1$, maximized over all pairs of intervals in the recursive decomposition Eq. (57). More explicitly,

$$
\lambda_{\mathrm{block}} := \max_{\sigma\neq\sigma'\in\{x,y,z\}} \max_{\ell=1,\dots,\log n-1} \max_{b=0,\dots,2^{\ell-1}-1} \frac{\left\|\beta^{(\sigma,\sigma')}\right\|_{1,\mathrm{box},\mathcal{I}_{\ell,b}}}{\left\|\beta^{(\sigma,\sigma')}\right\|_{1,\mathcal{I}_{\ell,b}}},
\tag{82}
$$

where $\|\cdot\|_{1,\mathcal{I}_{\ell,b}}$ and $\|\cdot\|_{1,\mathrm{box},\mathcal{I}_{\ell,b}}$ are the above norms defined with respect to the region

$$
\mathcal{I}_{\ell,b} := \left\{(u,v)\mid 2b\frac{n}{2^\ell}+1 \leq u \leq (2b+1)\frac{n}{2^\ell}, (2b+1)\frac{n}{2^\ell}+1 \leq v \leq 2(b+1)\frac{n}{2^\ell}\right\}.
\tag{83}
$$

Then, the complexity of quantum simulation should be revised to

$$
\begin{aligned}
&\left(\sqrt{\lambda_{\mathrm{block}}}+n\right)t\left(\tfrac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq 2,\\
&\left(\sqrt{\lambda_{\mathrm{block}}}+n\right)\left(\mathcal{O}\left(n^{2-\alpha}t\,\mathrm{polylog}\left(\tfrac{nt}{\epsilon}\right)\right)+t\left(\tfrac{nt}{\epsilon}\right)^{o(1)}\right), & 1 \leq \alpha < 2,\\
&\left(\sqrt{\lambda_{\mathrm{block}}}+n\right)\left(\mathcal{O}\left(n^{2-\alpha}t\,\mathrm{polylog}\left(\tfrac{nt}{\epsilon}\right)\right)+n^{1-\alpha}t\left(\tfrac{nt}{\epsilon}\right)^{o(1)}\right), & 0 < \alpha < 1.
\end{aligned}
\tag{84}
$$

It is clear that when the distribution of coefficients is close to a power law, Eq. (77) implies that $\lambda_{\mathrm{block}} = \mathcal{O}(1)$ so we have recovered the cost scaling claimed earlier. When coefficients deviate significantly from a power-law distribution, our gate complexity will enlarge by a factor of $\sqrt{\lambda_{\mathrm{block}}}$ due to the use of amplitude amplification. However, we always have

$$
\frac{\|\beta\|_{1,\mathrm{box}}}{\|\beta\|_1} = \frac{\sum_{\mu,\nu=0}^{\log n-2} 2^{\mu+\nu}\max_{\substack{-2^{\mu+1}<u\leq-2^\mu\\2^\nu\leq v<2^{\nu+1}}}|\beta_{u,v}|}{\sum_{\mu,\nu=0}^{\log n-2}\sum_{\substack{-2^{\mu+1}<u\leq-2^\mu\\2^\nu\leq v<2^{\nu+1}}}|\beta_{u,v}|} \leq \frac{n^2\sum_{\mu,\nu=0}^{\log n-2}\max_{\substack{-2^{\mu+1}<u\leq-2^\mu\\2^\nu\leq v<2^{\nu+1}}}|\beta_{u,v}|}{\sum_{\mu,\nu=0}^{\log n-2}\max_{\substack{-2^{\mu+1}<u\leq-2^\mu\\2^\nu\leq v<2^{\nu+1}}}|\beta_{u,v}|} = n^2,
$$

which implies $\sqrt{\lambda_{\mathrm{block}}} \leq n$. So the preparation subroutine still costs less than the selection subroutine, even when we perform the amplitude amplification. Thus our asymptotic gate complexity remains the same.

C. Summary of the algorithm

We now summarize the block-encoding method for implementing faster Trotter steps:

1. Construct a preamplified preparation subroutine for Eq. (73).
2. Perform $\mathcal{O}(\sqrt{\lambda_{\mathrm{block}}})$ steps of amplitude amplification to construct the actual preparation subroutine with λ_{block} defined in Eq. (82).
3. Define the selection subroutine according to Eq. (74).
4. Use qubitization (Lemma 3) to simulate for time t/r with accuracy $\mathcal{O}(\epsilon/(rn))$, with r scaling like Eq. (51).
5. Perform qubitization for all combinations of Pauli operators $\sigma,\sigma' = x,y,z$, layers indexed by $\ell = 1,\dots,\log n-1$ and pairs of intervals indexed by $b = 0,\dots,2^{\ell-1}-1$, to implement a single Trotter step.
6. Repeat r Trotter steps to simulate the entire evolution.

Theorem 1: (Faster Trotter steps using block encoding). *Consider two-local Hamiltonians*

$$H = \sum_{\sigma,\sigma' \in \{i,x,y,z\}} \sum_{1 \leq j < k \leq n} \beta_{j,k}^{(\sigma,\sigma')} P_j^{(\sigma)} P_k^{(\sigma')},$$

where $\left| \beta_{j,k}^{(\sigma,\sigma')} \right| \leq 1/|j-k|^\alpha$ *for some constant* $\alpha > 0$ *and* $P^{(\sigma)}$ *($\sigma = i, x, y, z$) are the identity and Pauli matrices. Let $t > 0$ be the simulation time and $\epsilon > 0$ be the target accuracy. Assume that the coefficient oracle*

$$O_{\beta,\sigma,\sigma'} |j, k, 0\rangle = |j, k, \beta_{j,k}^{(\sigma,\sigma')}\rangle \tag{85}$$

can be implemented with gate complexity $\mathcal{O}(\mathrm{polylog}(nt/\epsilon))$. *Then H can be simulated using the algorithm of Sec.* III C *with* $\mathcal{O}(\log(nt/\epsilon))$ *ancilla qubits and gate complexity*

$$\begin{cases} nt \left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq 2, \\ \mathcal{O}\left(n^{3-\alpha} t \, \mathrm{polylog}\left(\frac{nt}{\epsilon}\right)\right) + nt \left(\frac{nt}{\epsilon}\right)^{o(1)}, & 1 \leq \alpha < 2, \\ \mathcal{O}\left(n^{3-\alpha} t \, \mathrm{polylog}\left(\frac{nt}{\epsilon}\right)\right) + n^{2-\alpha} t \left(\frac{nt}{\epsilon}\right)^{o(1)}, & 0 < \alpha < 1. \end{cases} \tag{86}$$

IV. FASTER TROTTER STEPS USING AVERAGE-COST SIMULATION

In the previous section, we have described a block-encoding-based method to simulate power-law Hamiltonians with efficiently computable coefficients. The complexity of our method is almost linear in the spacetime volume $(nt)^{1+o(1)}$ for $\alpha \geq 2$ and is close to $n^{3-\alpha+o(1)} t^{1+o(1)}$ for $\alpha < 2$, both of which improve the best results from previous work.

In this section, we obtain a further improvement by performing an average-cost quantum simulation of commuting terms. We explain the basic idea of this technique in Sec. IV A, with further details on the circuit implementation presented in Sec. IV B. Readers may skip ahead to Sec. IV C for a summary of the entire algorithm.

A. Simulating commuting terms with average combination cost

We now show that gate complexities can be further reduced, using the simple fact that commuting Hamiltonian terms can be simulated with an *average-cost linear combination*. To elaborate, consider a Hamiltonian $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ with Hermitian H_γ and suppose we have block encodings $G_{\gamma,1}^\dagger U_\gamma G_{\gamma,0} = H_\gamma / \beta_\gamma$, where $\beta_\gamma > 0$, and $G_{\gamma,0}, G_{\gamma,1} : \mathcal{G} \to \mathcal{H}$ and $U_\gamma : \mathcal{H} \to \mathcal{H}$ can be implemented with cost $c_\gamma > 0$. Then, H can be block encoded by defining

$$\begin{aligned} G_0 &: \mathcal{G} \to \mathbb{C}^\Gamma \otimes \mathcal{H}, & G_0 &= \frac{1}{\sqrt{\|\beta\|_1}} \sum_{\gamma=1}^{\Gamma} \sqrt{\beta_\gamma} |\gamma\rangle \otimes G_{\gamma,0} \\[2mm] G_1 &: \mathcal{G} \to \mathbb{C}^\Gamma \otimes \mathcal{H}, & G_1 &= \frac{1}{\sqrt{\|\beta\|_1}} \sum_{\gamma=1}^{\Gamma} \sqrt{\beta_\gamma} |\gamma\rangle \otimes G_{\gamma,1} \\[2mm] U &: \mathbb{C}^\Gamma \otimes \mathcal{H} \to \mathbb{C}^\Gamma \otimes \mathcal{H}, & U &= \sum_{\gamma=1}^{\Gamma} |\gamma\rangle\langle\gamma| \otimes U_\gamma, \\[2mm] G_1^\dagger U G_0 &: \mathcal{G} \to \mathcal{G}, & G_1^\dagger U G_0 &= \frac{H}{\|\beta\|_1}. \end{aligned} \tag{87}$$

This block encoding has a normalization factor of $\|\beta\|_1$ and can be implemented with cost $\sum_\gamma c_\gamma$ (plus some additional cost for preparing the ancilla state $1/\sqrt{\|\beta\|_1} \sum_{\gamma=1}^{\Gamma} \sqrt{\beta_\gamma} |\gamma\rangle$). Invoking Lemma 3, we can simulate H for time t with

accuracy ϵ with a cost scaling like

$$\mathcal{O}\left(\left(\|\beta\|_1\, t + \log\left(\frac{1}{\epsilon}\right)\right) \cdot \left(\sum_{\gamma=1}^{\Gamma} c_\gamma\right)\right), \tag{88}$$

which reduces to approximately $t\left(\sum_\gamma \beta_\gamma\right)\left(\sum_\gamma c_\gamma\right)$ ignoring the error scaling. This is a worst-case combination because we are paying the same total cost $\sum_\gamma c_\gamma$ for each of the approximately $t\sum_\gamma \beta_\gamma$ qubitization steps.

This worst-case cost scaling may sometimes be avoided by recursively performing simulation in the interaction picture [26,68]. Roughly speaking, the recursive interaction-picture approach has a cost scaling like approximately $t\left(\sum_\gamma \beta_\gamma c_\gamma\right) \log^{2\Gamma-1}\left(t\|\beta\|_{\max}\right)$. Thus we achieve the desired average-case combination cost, but also pick up a factor that scales exponentially with the number of terms Γ, which prevents the approach from being useful in many cases.

Instead, we make the following simple yet important observation about simulating commuting terms using block encodings and qubitization.

Lemma 4: (Simulating commuting terms with average combination cost). *Let $G_{\gamma,0}, G_{\gamma,1}: \mathcal{G} \to \mathcal{H}$ be isometries and $U_\gamma: \mathcal{H} \to \mathcal{H}$ be unitaries such that $G_{\gamma,1}^\dagger U_\gamma G_{\gamma,0} = H_\gamma/\beta_\gamma$ are Hermitian with $\beta_\gamma > 0$. Assume that H_γ pairwise commute. Given a target evolution time t and accuracy ϵ, there exist unitaries $V_{\varphi_\gamma}: \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathcal{H} \to \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathcal{H}$ parameterized by angles $\varphi_{\gamma,1}, \ldots, \varphi_{\gamma,r_\gamma}$ such that*

$$\left\|\prod_{\gamma=1}^{\Gamma}\left(\left(\langle+|\otimes \frac{\langle 0|\otimes G_{\gamma,0}^\dagger + \langle 1|\otimes G_{\gamma,1}^\dagger}{\sqrt{2}}\right) V_{\varphi_\gamma}\left(|+\rangle \otimes \frac{|0\rangle \otimes G_{\gamma,0} + |1\rangle \otimes G_{\gamma,1}}{\sqrt{2}}\right)\right) - e^{-itH}\right\| \leq \epsilon. \tag{89}$$

The number of steps r_γ are even integers with the asymptotic scaling

$$r_\gamma = \mathcal{O}\left(\beta_\gamma t + \log\left(\frac{\Gamma}{\epsilon}\right)\right), \tag{90}$$

and V_{φ_γ} are obtained by applying Lemma 3 to simulate $G_{\gamma,1}^\dagger U_\gamma G_{\gamma,0}$ for time $\beta_\gamma t$ with accuracy ϵ/Γ.

In essence, we are just using the first-order Lie-Trotter formula $e^{-itH} = e^{-itH_\Gamma}\cdots e^{-itH_1}$ with each exponential further simulated by the qubitization algorithm. Because Hamiltonian terms pairwise commute, there is no Trotter error introduced in this decomposition. As a result, we can simulate the target Hamiltonian with an average-case combination cost approximately $t\left(\sum_\gamma \beta_\gamma c_\gamma\right)$, without the unwanted exponential factor from the interaction-picture approach [117].

For power-law Hamiltonians with $\alpha < 2$, recall from Sec. III A that we can without loss of generality consider

$$H_{(-n/2,-1]:[1,n/2)} = \sum_{\substack{-n/2 < u \leq -1 \\ 1 \leq v < n/2}} \beta_{u,v} Z_u Z_v, \tag{91}$$

where we drop the subdominant terms and shift the intervals for notational convenience. Such a Hamiltonian term is directly block encoded and simulated by qubitization in Sec. III B. We now show how that result can be further improved using the average-cost simulation technique. To this end, we divide each interval into m subintervals at $\pm l_1, \ldots, \pm l_m$ and define

$$\mathcal{C}_{j,k} := \left\{(u,v)|-l_{j+1} < u \leq -l_j, l_k \leq v < l_{k+1}\right\}. \tag{92}$$

See Fig. 2 for an illustration of the corresponding decomposition of Hamiltonian. For terms corresponding to $\mathcal{C}_{j,k}$, we have that the 1-norm is asymptotically bounded by

$$\|\beta\|_{1,\mathcal{C}_{j,k}} = \mathcal{O}\left(\frac{(l_{j+1}-l_j)(l_{k+1}-l_k)}{(l_j+l_k)^\alpha}\right). \tag{93}$$
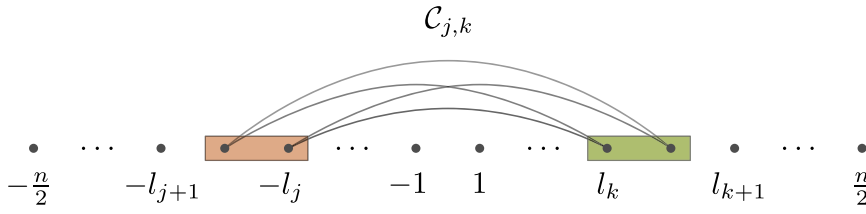
$\mathcal{C}_{j,k}$

FIG. 2. Illustration of the interaction pairs in $\mathcal{C}_{j,k}$ defined in Eq. (92). For convenience, we relabel the sites to $-n/2, \ldots, -1$ and $1, \ldots, n/2$.

We describe in Sec. IV B how to block encode the Hamiltonian terms within $\mathcal{C}_{j,k}$ with 1-norm scaling exactly as above. Specifically, we show that up to polylogarithmic factors the preparation and selection subroutines have a total cost of

$$\widetilde{\mathcal{O}}\left(1 + \left(l_{j+1} - l_j\right) + \left(l_{k+1} - l_k\right)\right). \tag{94}$$

In our above analysis, we omit a factor of λ_{avg} due to the use of amplitude amplification. Just like Sec. III B, this factor is close to 1 when the distribution of Hamiltonian coefficients are close to a power-law distribution. We assume this is the case to simplify the following discussion, and present the full complexity expression in Sec. IV B.

We use the uniform division $l_j = \Theta\left(j\left(n/m\right)\right)$ here for simplicity, although other divisions may lead to circuits with lower cost. Choosing the evolution time t/r and ignoring the polylog(nt/ϵ) scaling, we estimate the cost of simulating $H_{(-n/2,-1]:[1,n/2)}$ (for $\alpha < 2$) as

$$\widetilde{\mathcal{O}}\left(\sum_{j,k=1}^{m}\left(\frac{\left(\frac{n}{m}\right)^2 \frac{t}{r}}{\left(j\frac{n}{m} + k\frac{n}{m}\right)^\alpha} + 1\right)\left(1 + \frac{n}{m}\right)\right) = \widetilde{\mathcal{O}}\left(\left(\sum_{j,k=1}^{m}\frac{\left(\frac{n}{m}\right)^{2-\alpha}\frac{t}{r}}{(j+k)^\alpha} + m^2\right)\left(1 + \frac{n}{m}\right)\right)$$

$$= \widetilde{\mathcal{O}}\left(\left(\left(\sum_{j=1}^{m}\frac{1}{j^{\alpha/2}}\right)^2\left(\frac{n}{m}\right)^{2-\alpha}\frac{t}{r} + m^2\right)\left(1 + \frac{n}{m}\right)\right)$$

$$= \widetilde{\mathcal{O}}\left(\left(\left(m^{1-\frac{\alpha}{2}}\right)^2\left(\frac{n}{m}\right)^{2-\alpha}\frac{t}{r} + m^2\right)\left(1 + \frac{n}{m}\right)\right)$$

$$= \widetilde{\mathcal{O}}\left(\left(n^{2-\alpha}\frac{t}{r} + m^2\right)\left(1 + \frac{n}{m}\right)\right). \tag{95}$$

We balance the two scalings in the first parentheses to optimize the gate complexity, which implies

$$m = \begin{cases} \Theta\left(n^{1-\alpha/2}\right), & 1 \le \alpha < 2, \\ \Theta\left(n^{\frac{1}{2}}\right), & 0 < \alpha < 1. \end{cases} \tag{96}$$

For all values of $\alpha < 2$, one can verify that $m = \Omega(1)$ and $m = \mathcal{O}(n)$ [in fact $m = \mathcal{O}(\sqrt{n})$], so this choice of m is indeed valid. We give a slightly better (yet more complicated) choice of m in Sec. IV B.

The remaining analysis proceeds similarly as in Sec. III B. We use qubitization (Lemma 3) to simulate all m^2 pairs of subintervals in $H_{[1,n/2]:[n/2+1,n]}$ for time t/r with accuracy $\mathcal{O}\left(\epsilon/(rm^2n)\right)$, where

$$r = \begin{cases} \frac{n^{o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & \alpha \ge 1, \\ \frac{n^{1-\alpha+o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < 1, \end{cases} \tag{97}$$

with an average-case combination cost as discussed above. We have the cost function

$$\begin{aligned} \mathcal{O}\left(n^{2-\alpha/2}\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & \quad 1 \le \alpha < 2, \\ \mathcal{O}\left(n^{3/2}\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & \quad 0 < \alpha < 1. \end{aligned} \tag{98}$$

By using the same evolution time and target accuracy with different values of n and invoking the master theorem (Lemma 1), we obtain the gate complexity of implementing one Trotter step

$$\text{cost}_{\text{rec}}(n) = \sum_{l=1}^{\log n - 1} 2^{l-1} \text{cost}\left(\frac{n}{2^{l-1}}\right) = \begin{cases} \mathcal{O}\left(n^{2-\alpha/2}\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & 1 \le \alpha < 2, \\ \mathcal{O}\left(n^{3/2}\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right), & 0 < \alpha < 1. \end{cases} \quad (99)$$

This then gives the complexity of the entire quantum simulation:

$$\begin{cases} \dfrac{n^{2-\alpha/2+o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}, & 1 \le \alpha < 2, \\ \dfrac{n^{5/2-\alpha+o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < 1. \end{cases} \quad (100)$$

B. Preparation and selection subroutines

We now describe a circuit that achieves the average-case combination cost for block-encoding power-law Hamiltonians claimed in the previous subsection.

Specifically, let $l_j, l_{j+1}, l_k, l_{k+1}$ be arbitrary integers such that $1 \le l_j < l_{j+1} \le n/2$ and $1 \le l_k < l_{k+1} \le n/2$. Our goal is to block encode the Hamiltonian

$$H_{(-l_{j+1}, -l_j]:[l_k, l_{k+1})} = \sum_{\substack{-l_{j+1} < u \le -l_j \\ l_k \le v < l_{k+1}}} \beta_{u,v} Z_u Z_v, \quad (101)$$

where $|\beta_{u,v}| \le 1/|u - v|^\alpha$ and we shift the intervals for notational convenience. Additionally, we assume that the coefficients are *logarithmically computable* [118], meaning the oracle

$$O_\beta |u, v, 0\rangle = |u, v, \beta_{u,v}\rangle \quad (102)$$

can be implemented with cost $\mathcal{O}\left(\text{polylog}(nt/\epsilon)\right)$. By making this oracle assumption, we implicitly assume certain underlying structure of the Hamiltonian coefficients: in Appendix A we show that one needs approximately n^2 gates to implement O_β in the circuit model when structural properties of coefficients are unavailable. The selection subroutine can simply be chosen as

$$\sum_{-l_{j+1} < u \le -l_j} \sum_{l_k \le v < l_{k+1}} |u, v\rangle\langle u, v| \otimes Z_u Z_v. \quad (103)$$

Due to the product structure of this operation, we can implement it with gate complexity $\mathcal{O}(n/m)$.

In what follows, we analyze the preparation subroutine. Here, we consider only the uniform division $l_j = \Theta(j(n/m))$, as this is enough to justify the gate complexity claimed in Sec. IV A. We use the naive black-box state-preparation technique, because coefficients within the divided subintervals are already close to uniform. We start by preparing the uniform superposition

$$\frac{1}{\sqrt{l_{j+1} - l_j}} \sum_{-l_{j+1} < u \le -l_j} |u\rangle \frac{1}{\sqrt{l_{k+1} - l_k}} \sum_{l_k \le v < l_{k+1}} |v\rangle. \quad (104)$$

Then we invoke the black-box state-preparation subroutine

$$\frac{1}{\sqrt{(l_{j+1} - l_j)(l_{k+1} - l_k)}} \sum_{\substack{-l_{j+1} < u \le -l_j \\ l_k \le v < l_{k+1}}} |u, v\rangle$$

$$\overset{O_\beta}{\mapsto} \frac{1}{\sqrt{(l_{j+1} - l_j)(l_{k+1} - l_k)}} \sum_{\substack{-l_{j+1} < u \le -l_j \\ l_k \le v < l_{k+1}}} |u, v\rangle |\beta_{u,v}\rangle$$

$$\mapsto \frac{1}{\sqrt{(l_{j+1} - l_j)(l_{k+1} - l_k)}} \sum_{\substack{-l_{j+1} < u \le -l_j \\ l_k \le v < l_{k+1}}} |u, v\rangle |\beta_{u,v}\rangle \left(\sqrt{\frac{|\beta_{u,v}|}{\|\beta\|_{\max}}} |0\rangle + \sqrt{\frac{\|\beta\|_{\max} - |\beta_{u,v}|}{\|\beta\|_{\max}}} |1\rangle\right), \quad (105)$$

where the last step can be realized using inequality test in a way similar to Sec. III B. There is no need to uncompute the ancilla register $|\beta_{u,v}\rangle$, as the uncomputation is automatically performed in the qubitization algorithm. It is clear that we prepare a state proportional to the desired state for block encoding when all $\beta_{u,v} > 0$. If some $\beta_{u,v} < 0$, we simply introduce a minus sign in the implementation of the selection subroutine. The success probability is given by

$$\frac{\sum_{\substack{-l_{j+1}<u\leq -l_j \\ l_k\leq v<l_{k+1}}} |\beta_{u,v}|}{\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)\|\beta\|_{\max}} = \frac{\|\beta\|_1}{\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)\|\beta\|_{\max}}. \tag{106}$$

To boost this probability to close to 1, we can perform $\mathcal{O}\left(\sqrt{\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)\|\beta\|_{\max}/\|\beta\|_1}\right)$ steps of amplitude amplification.

If the distribution of Hamiltonian coefficients exactly matches a power law, we have

$$\|\beta\|_1 = \sum_{\substack{-l_{j+1}<u\leq -l_j \\ l_k\leq v<l_{k+1}}} |\beta_{u,v}| \geq \frac{\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)}{\left(l_{j+1}+l_{k+1}\right)^\alpha}, \qquad \|\beta\|_{\max} \leq \frac{1}{\left(l_j+l_k\right)^\alpha}, \tag{107}$$

so the number of amplification steps scales like

$$\mathcal{O}\left(\sqrt{\frac{\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)\|\beta\|_{\max}}{\|\beta\|_1}}\right) = \mathcal{O}\left(\sqrt{\frac{\left(l_{j+1}+l_{k+1}\right)^\alpha}{\left(l_j+l_k\right)^\alpha}}\right) = \mathcal{O}(1), \tag{108}$$

which justifies the previous claim in Eq. (94). In general, the complexity of block encoding will depend on the closeness of the distribution of Hamiltonian coefficients to a power-law distribution. To quantify this, we let λ_{avg} be the maximum ratio between $\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)\|\beta\|_{\max}$ and $\|\beta\|_1$, maximized over all pairs of intervals in the decomposition Eq. (57) with a further uniform division. More explicitly,

$$\lambda_{\mathrm{avg}} := \max_{\sigma\neq\sigma'\in\{x,y,z\}} \max_{\ell=1,\ldots,\log n-1} \max_{b=0,\ldots,2^{\ell-1}-1} \max_{j,k=1,\ldots,m} \frac{\left(l_{\ell,b,j+1}-l_{\ell,b,j}\right)\left(l_{\ell,b,k+1}-l_{\ell,b,k}\right)\left\|\beta^{(\sigma,\sigma')}\right\|_{\max,\mathcal{I}_{\ell,b,j,k}}}{\left\|\beta^{(\sigma,\sigma')}\right\|_{1,\mathcal{I}_{\ell,b,j,k}}}, \tag{109}$$

where $l_{\ell,b,j}$ are the uniform division points, and $\|\cdot\|_{\max,\mathcal{I}_{\ell,b,j,k}}$ and $\|\cdot\|_{1,\mathcal{I}_{\ell,b,j,k}}$ are the max- and 1-norm restricted to the region

$$\mathcal{I}_{\ell,b,j,k} := \left\{(u,v)|\ (2b+1)\frac{n}{2^\ell}-l_{\ell,b,j+1}<u\leq(2b+1)\frac{n}{2^\ell}-l_{\ell,b,j},(2b+1)\frac{n}{2^\ell}+l_{\ell,b,k}\leq v\leq(2b+1)\frac{n}{2^\ell}+l_{\ell,b,k+1}\right\}. \tag{110}$$

With this definition, the cost of simulating $H_{(-n/2,-1]:[1,n/2)}$ for time t/r should be revised to

$$\mathrm{cost}(n) = \widetilde{\mathcal{O}}\left(\left(n^{2-\alpha}\frac{t}{r}+m^2\right)\left(\sqrt{\lambda_{\mathrm{avg}}}+\frac{n}{m}\right)\right). \tag{111}$$

However, we always have

$$\frac{\left(l_{j+1}-l_j\right)\left(l_{k+1}-l_k\right)\|\beta\|_{\max}}{\|\beta\|_1} \leq \frac{\left(\frac{n}{m}\right)^2\|\beta\|_{\max}}{\|\beta\|_{\max}} = \left(\frac{n}{m}\right)^2,$$

which implies $\sqrt{\lambda_{\mathrm{avg}}}\leq n/m$. So the preparation subroutine still costs less than the selection subroutine, even when we perform the amplitude amplification. We balance the first term by choosing

$$m = \max\left\{\Theta\left(n^{1-\alpha/2}\left(\frac{t}{r}\right)^{1/2}\right),1\right\}, \tag{112}$$

which implies through the master theorem (Lemma 1) that the total simulation has complexity

$$\widetilde{\mathcal{O}}\left(\left(n^{2-\alpha}t+r\right)\min\left\{n^{\alpha/2}\left(\frac{r}{t}\right)^{1/2},n\right\}\right). \tag{113}$$

Inserting the scaling of r from Eq. (51), we obtain the cost scaling

$$\min\left\{n^{2-\frac{\alpha}{2}}t\left(\frac{nt}{\epsilon}\right)^{o(1)},\ \mathcal{O}\left(n^{3-\alpha}t\,\mathrm{polylog}\left(\frac{nt}{\epsilon}\right)\right)+nt\left(\frac{nt}{\epsilon}\right)^{o(1)}\right\},\qquad 1\leq\alpha<2,$$
$$\min\left\{n^{5/2-\alpha}t\left(\frac{nt}{\epsilon}\right)^{o(1)},\ \mathcal{O}\left(n^{3-\alpha}t\,\mathrm{polylog}\left(\frac{nt}{\epsilon}\right)\right)+n^{2-\alpha}t\left(\frac{nt}{\epsilon}\right)^{o(1)}\right\},\quad 0<\alpha<1. \tag{114}$$

C. Summary of the algorithm

We now summarize the average-cost simulation method for implementing faster Trotter steps:

1. Construct a preamplified preparation subroutine for Eq. (105).
2. Perform $\mathcal{O}(\sqrt{\lambda_{\mathrm{avg}}})$ steps of amplitude amplification to construct the actual preparation subroutine with λ_{avg} defined in Eq. (109).
3. Define the selection subroutine according to Eq. (103).
4. Use qubitization (Lemma 3) to simulate for time t/r with accuracy $\mathcal{O}\left(\epsilon/(rn^2)\right)$, with r scaling like Eq. (51).
5. Perform qubitization for all combinations of Pauli operators $\sigma,\sigma'=x,y,z$, layers indexed by $\ell=1,\ldots,\log n-1$, intervals indexed by $b=0,\ldots,2^{\ell-1}-1$ and pairs of subintervals indexed by $j,k=1,\ldots,m$, to implement a single Trotter step.
6. Repeat r Trotter steps to simulate the entire evolution.

Theorem 2: (Faster Trotter steps using average-cost simulation). *Consider two-local Hamiltonians*

$$H=\sum_{\sigma,\sigma'\in\{i,x,y,z\}}\sum_{1\leq j<k\leq n}\beta_{j,k}^{(\sigma,\sigma')}P_j^{(\sigma)}P_k^{(\sigma')},$$

where $\left|\beta_{j,k}^{(\sigma,\sigma')}\right|\leq 1/|j-k|^{\alpha}$ *for some constant* $\alpha>0$ *and* $P^{(\sigma)}$ *($\sigma=i,x,y,z$) are the identity and Pauli matrices. Let* $t>0$ *be the simulation time and* $\epsilon>0$ *be the target accuracy. Assume that the coefficient oracle*

$$O_{\beta,\sigma,\sigma'}|j,k,0\rangle=|j,k,\beta_{j,k}^{(\sigma,\sigma')}\rangle \tag{115}$$

can be implemented with gate complexity $\mathcal{O}\left(\mathrm{polylog}(nt/\epsilon)\right)$. *Then* H *can be simulated using the algorithm of Sec. IV C with* $\mathcal{O}\left(\log(nt/\epsilon)\right)$ *ancilla qubits and gate complexity*

$$\min\left\{n^{2-\frac{\alpha}{2}}t\left(\frac{nt}{\epsilon}\right)^{o(1)},\ \mathcal{O}\left(n^{3-\alpha}t\,\mathrm{polylog}\left(\frac{nt}{\epsilon}\right)\right)+nt\left(\frac{nt}{\epsilon}\right)^{o(1)}\right\},\qquad 1\leq\alpha<2,$$
$$\min\left\{n^{5/2-\alpha}t\left(\frac{nt}{\epsilon}\right)^{o(1)},\ \mathcal{O}\left(n^{3-\alpha}t\,\mathrm{polylog}\left(\frac{nt}{\epsilon}\right)\right)+n^{2-\alpha}t\left(\frac{nt}{\epsilon}\right)^{o(1)}\right\},\quad 0<\alpha<1. \tag{116}$$

V. FASTER TROTTER STEPS USING LOW-RANK DECOMPOSITION

In the previous sections, we show how the 1-norm of Hamiltonian coefficients can be reduced via a recursive decomposition using product formulas, which results in faster circuit implementation of Trotter steps. We further improve our result by simulating commuting terms with an average-case combination cost. Assuming Hamiltonian coefficients are efficiently computable, these techniques together enable simulations of power-law systems with complexity nearly linear in the spacetime volume for $\alpha\geq 2$, whereas the cost becomes $n^{2-\alpha/2+o(1)}t^{1+o(1)}$ for $1\leq\alpha<2$ and $n^{5/2-\alpha+o(1)}t^{1+o(1)}$ for $0<\alpha<1$.

In this section, we describe a method for implementing a Trotter step through a recursive decomposition of the Hamiltonian using its hierarchical low-rank structure [50,51]. This low-rank structure was previously used in Ref. [52] to block-encode kernel matrices. Here, we directly use the recursive decomposition in Sec. V A to construct circuits without block encoding. The overall algorithm and its complexity are then summarized in Sec. V B.

A. Recursive low-rank decomposition

The initial steps of our method are the same as that of Sec. III A. In particular, we expand the power-law Hamiltonian in the Pauli basis as in Eq. (48), and use product formulas to perform a coarse-grained decomposition. Without loss of generality, we may focus on two-local terms

$$H = \sum_{1 \le j < k \le n} \beta_{j,k} X_j Y_k, \tag{117}$$

as the on-site terms can be implemented with subdominant cost and the remaining two-local terms can be handled similarly by a change of basis. As before, we assume that the system size n is a power of 2 and use

$$H_{[j,k]} := \sum_{j \le u < v \le k} \beta_{u,v} X_u Y_v \ (1 \le j < k \le n),$$

$$H_{[j,k]:[l,m]} := \sum_{\substack{j \le u \le k \\ l \le v \le m}} \beta_{u,v} X_u Y_v \ (1 \le j \le k < l \le m \le n) \tag{118}$$

to represent terms within a specific interval and across two disjoint intervals of sites.

However, we now use a decomposition different from that of Sec. III A. Specifically, we use the recurrence relation

$$H_{[j,k]} = H_{[j,j+\delta-1]:[j+2\delta,j+3\delta-1]} + H_{[j,j+\delta-1]:[j+3\delta,k]} + H_{[j+\delta,j+2\delta-1]:[j+3\delta,k]} + H_{[j,j+2\delta-1]} + H_{[j+\delta,j+3\delta-1]} + H_{[j+2\delta,k]}$$
$$- H_{[j+\delta,j+2\delta-1]} - H_{[j+2\delta,j+3\delta-1]} \tag{119}$$

with $\delta = \lfloor (k - j + 1)/4 \rfloor$, which unwraps to layer $h = \Theta(\log n)$ as

$$H = H_{[1,n]}$$
$$= H_{[1,n/4]:[n/2+1,3n/4]} + H_{[1,n/4]:[3n/4+1,n]} + H_{[n/4+1,n/2]:[3n/4+1,n]} + H_{[1,n/2]} + H_{[n/4+1,3n/4]} + H_{[n/2+1,n]}$$
$$- H_{[n/4+1,n/2]} - H_{[n/2+1,3n/4]}$$
$$= \cdots$$
$$= \sum_{\ell=2}^{h} \sum_{b=0}^{2^{\ell-1}-2} \Big(H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+2)n/2^\ell,(2b+3)n/2^\ell]} + H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+3)n/2^\ell,(2b+4)n/2^\ell]}$$
$$+ H_{[1+(2b+1)n/2^\ell,(2b+2)n/2^\ell]:[1+(2b+3)n/2^\ell,(2b+4)n/2^\ell]} \Big) + \sum_{b=0}^{2^h-1} H_{[1+b(n/2^\ell),(b+1)n/2^\ell]}$$
$$+ \sum_{b=0}^{2^h-2} H_{[1+b(n/2^\ell),(b+1)n/2^\ell]:[1+(b+1)n/2^\ell,(b+2)n/2^\ell]}. \tag{120}$$

See Fig. 3 for an illustration of this decomposition at the first nontrivial layer $\ell = 2$.

We observe the following features of the decomposition that are helpful to describe our circuit implementation:

1. There are $h - 1 = \Theta(\log n)$ layers in the decomposition, all indexed by ℓ.
2. For a fixed layer ℓ, there are 2^ℓ consecutive intervals each of length $n/2^\ell$.
3. Within each layer ℓ, intervals are further divided into $2^{\ell-1}$ blocks (indexed by b). Within each pair of consecutive blocks, we keep only the three terms that act on intervals with distance at least $n/2^\ell$. They are

$$H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+2)n/2^\ell,(2b+3)n/2^\ell]},$$
$$H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+3)n/2^\ell,(2b+4)n/2^\ell]},$$
$$H_{[1+(2b+1)n/2^\ell,(2b+2)n/2^\ell]:[1+(2b+3)n/2^\ell,(2b+4)n/2^\ell]}.$$

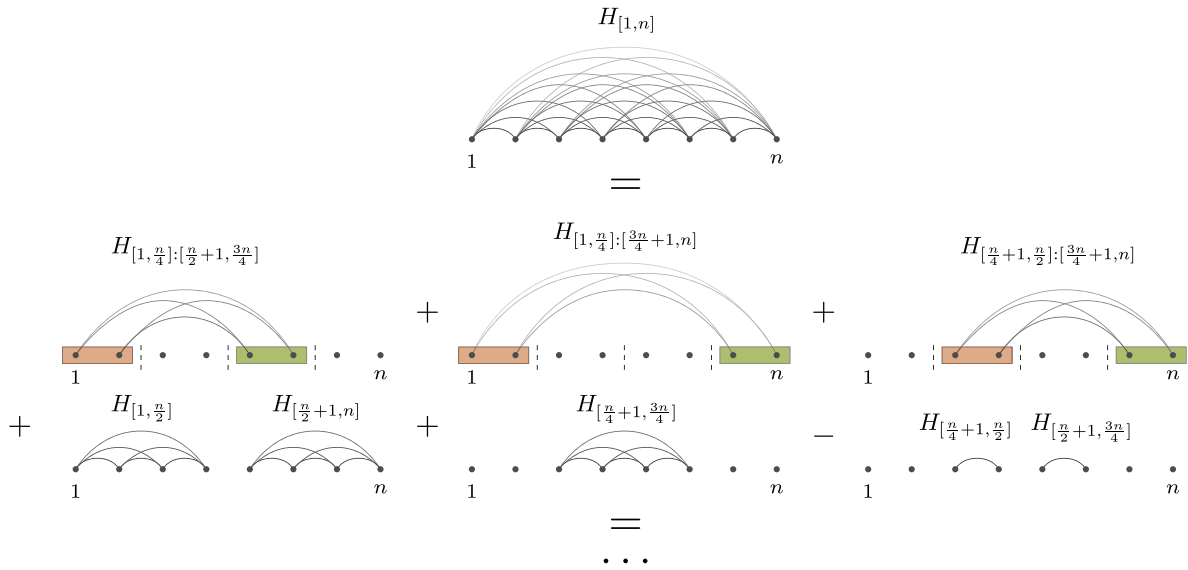The total number of pairs of intervals is $\mathcal{O}(n)$ by the master theorem Lemma 1.

FIG. 3. Illustration of one iteration [Eq. (119)] in the low-rank decomposition. We repeatedly apply Eq. (119) to the remaining terms in the second row to arrive at the final decomposition Eq. (120).

4. When $\ell = 2, \ldots, h$ and $b = 0, \ldots, 2^{\ell-1} - 2$, the decomposition provides a partition of all the terms in the Hamiltonian with distance at least $n/2^h$.

As in Sec. III A, we may use these features to simplify our analysis. For instance, since we have only two-local terms with the same type of Pauli operators acting across disjoint intervals, we can simultaneously change the basis and consider only Pauli-Z interactions, e.g.,

$$H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+2)n/2^\ell,(2b+3)n/2^\ell]} = \sum_{\substack{1+2b(n/2^\ell) \leq u \leq (2b+1)n/2^\ell \\ 1+(2b+2)n/2^\ell \leq v \leq (2b+3)n/2^\ell}} \beta_{u,v} Z_u Z_v. \tag{121}$$

Also, due to the nature of the recursive decomposition, it suffices to focus on the implementation of a specific term such as $H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+2)n/2^\ell,(2b+3)n/2^\ell]}$. We show momentarily that all the decomposed terms can be implemented with similar complexities cost$(\cdot)$, while the remaining terms such as $H_{[1+b(n/2^\ell),(b+1)n/2^\ell]}$ and $H_{[1+b(n/2^\ell),(b+1)n/2^\ell]:[1+(b+1)n/2^\ell,(b+2)n/2^\ell]}$ act on constant-size intervals and can be handled by a sequential implementation using product formulas. This means the total complexity can be bounded as

$$\text{cost}_{\text{rec}}(n) = \sum_{\ell=2}^{h} 3 \left(2^{\ell-1} - 1 \right) \text{cost} \left(\frac{n}{2^{\ell-1}} \right) + \mathcal{O}(n), \tag{122}$$

which again reduces to the study of cost$(\cdot)$ because of the master theorem (Lemma 1). For simplicity, we choose $\ell = 2$, $b = 0$ and we study the complexity cost$(n/2)$ of simulating $H_{[1,n/4]:[n/2+1,3n/4]}$.

Our main motivation to consider this decomposition is made clear through the following rank assumption of $H_{[1,n/4]:[n/2+1,3n/4]}$:

(i) Coefficients of $H_{[1,n/4]:[n/2+1,3n/4]}$, when organized into an $n/4$-by-$n/4$ matrix, have rank at most ρ, i.e., rank$\left([\beta_{u,v}] \right) \leq \rho$.

Because of this low-rank property, the coefficient matrix admits the thin singular value decomposition [102, Theorem 7.3.2]

$$\beta_{u,v} = \sum_{s=1}^{\rho} \mu_{u,s} \sigma_s \nu_{v,s} \tag{123}$$

for $u = 1\ldots, n/4$ and $v = n/2 + 1, \ldots, 3n/4$, where μ and ν, when viewed as $n/4$-by-ρ matrices, have real orthonormal columns, and σ_s are singular values bounded by the induced 1-norm of β and $\beta^\top$: $0 \leq \sigma_s \leq \sqrt{\|\|\beta\|\|_1 \|\|\beta^\top\|\|_1}$ [102, 5.6.P21]. Correspondingly, the exponential of $H_{[1,n/4]:[n/2+1,3n/4]}$ can be rewritten as

$$e^{-itH_{[1,n/4]:[n/2+1,3n/4]}} = e^{-it\sum_{s=1}^{\rho}\sigma_s\left(\sum_{u=1}^{n/4}\mu_{u,s}Z_u\right)\left(\sum_{v=n/2+1}^{3n/r}\nu_{v,s}Z_v\right)}. \tag{124}$$

With respect to the computational basis, our target exponential has the action

$$e^{-itH_{[1,n/4]:[n/2+1,3n/4]}}|z_n, \ldots, z_1\rangle = e^{-it\sum_{s=1}^{\rho}\sigma_s\left(\sum_{u=1}^{n/4}\mu_{u,s}(-1)^{z_u}\right)\left(\sum_{v=n/2+1}^{3n/r}\nu_{v,s}(-1)^{z_v}\right)}|z_n, \ldots, z_1\rangle. \tag{125}$$

This suggests a circuit implementation as follows. We first introduce an ancilla register of size $\mathcal{O}(\log(nt/\epsilon))$, and compute the following function in superposition:

$$\begin{bmatrix} z_n & \cdots & z_1 \end{bmatrix} \mapsto \sum_{s=1}^{\rho}\sigma_s\left(\sum_{u=1}^{n/4}\mu_{u,s}(-1)^{z_u}\right)\left(\sum_{v=n/2+1}^{3n/r}\nu_{v,s}(-1)^{z_v}\right). \tag{126}$$

Note that the function values are real numbers of size at most $\mathcal{O}(\text{poly}(n))$ and can thus be approximately stored in the $\mathcal{O}(\log(nt/\epsilon))$-qubit ancilla register. We then use a sequence of controlled rotations to introduce the phase. We may now uncompute the ancilla register by reverting the circuit. This implements the desired exponential $e^{-itH_{[1,n/4]:[n/2+1,3n/4]}}$.

Our above circuit has a gate complexity of cost $(n/2) = \mathcal{O}(n\rho\,\text{polylog}(nt/\epsilon))$ for simulating one pair of intervals in the recursive decomposition. To implement the Trotter step, we redefine ρ to be the maximum truncation rank of coefficient matrices, maximized over all pairs of intervals in the decomposition Eq. (57). Explicitly,

$$\rho := \max_{\sigma \neq \sigma' \in \{x,y,z\}} \max_{\ell=2,\ldots,h} \max_{b=0,\ldots,2^{l-1}-2} \rho_{\ell,b}^{(\sigma,\sigma')}, \tag{127}$$

where $\rho_{\ell,b}$ is the largest truncation rank for the terms

$$H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+2)n/2^\ell,(2b+3)n/2^\ell]}, \quad H_{[1+2b(n/2^\ell),(2b+1)n/2^\ell]:[1+(2b+3)n/2^\ell,(2b+4)n/2^\ell]},$$
$$H_{[1+(2b+1)n/2^\ell,(2b+2)n/2^\ell]:[1+(2b+3)n/2^\ell,(2b+4)n/2^\ell]}.$$

By the master theorem (Lemma 1), this implies the circuit for Trotter steps has the same asymptotic cost

$$\text{cost}_{\text{rec}}(n) = \sum_{\ell=2}^{h} 3\left(2^{\ell-1} - 1\right)\text{cost}\left(\frac{n}{2^{\ell-1}}\right) + \mathcal{O}(n) = \mathcal{O}\left(n\rho\,\text{polylog}\left(\frac{nt}{\epsilon}\right)\right). \tag{128}$$

By simulating for time t/r in each step and repeating r steps where

$$r = \begin{cases} \dfrac{n^{o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & \alpha \geq 1, \\[3mm] \dfrac{n^{1-\alpha+o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < 1, \end{cases} \tag{129}$$

we obtain the total gate complexity of the low-rank simulation method

$$\begin{cases} \dfrac{\rho(nt)^{1+o(1)}}{\epsilon^{o(1)}}, & \alpha \geq 1, \\[3mm] \dfrac{\rho n^{2-\alpha+o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < 1. \end{cases} \tag{130}$$

Before ending this section, we discuss the important question of how the rank ρ is determined in the above decomposition. For various classes of power-law interactions including the Coulomb interaction, there are rigorous analyses based on the multipole expansion showing that $\rho = \mathcal{O}(\log(nt/\epsilon))$ suffices to guarantee that the simulation is ϵ accurate [119], which leads to the gate complexity claimed in Table I. In fact, we soon examine an application of this method to simulating the real-space electronic structure Hamiltonian in Sec. VI. In practice, one may also run numerical simulations to empirically determine the rank value ρ. We refer the reader to Ref. [79] and papers citing this work for detailed studies of the low-rank decomposition in the classical setting.

B. Summary of the algorithm

We now summarize the low-rank method for implementing faster Trotter steps:

1. Use diagonalization to implement the low-rank matrix exponential Eq. (125).
2. Perform matrix exponentials for all combinations of Pauli operators $\sigma, \sigma' = x, y, z$, layers indexed by $\ell = 2, \ldots, h$ and blocks indexed by $b = 0, \ldots, 2^{\ell-1} - 2$, as well as the constant-size blocks, to implement a single Trotter step.
3. Repeat r Trotter steps to simulate the entire evolution.

Theorem 3: (Faster Trotter steps using low-rank decomposition). *Consider two-local Hamiltonians*

$$H = \sum_{\sigma,\sigma' \in \{i,x,y,z\}} \sum_{1 \leq j < k \leq n} \beta_{j,k}^{(\sigma,\sigma')} P_j^{(\sigma)} P_k^{(\sigma')},$$

where $\left| \beta_{j,k}^{(\sigma,\sigma')} \right| \leq 1/|j-k|^\alpha$ for some constant $\alpha > 0$ and $P^{(\sigma)}$ ($\sigma = i,x,y,z$) are the identity and Pauli matrices. Let $t > 0$ be the simulation time and $\epsilon > 0$ be the target accuracy. Then H can be simulated using the algorithm of Sec. V B with $\mathcal{O}(\log(nt/\epsilon))$ ancilla qubits and gate complexity

$$\begin{cases} \rho n t \left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq 1, \\ \rho n^{2-\alpha} t \left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha < 1. \end{cases} \quad (131)$$

Here, $1 \leq \rho \leq n$ defined in Eq. (127) is the maximum truncation rank of certain off-diagonal blocks of coefficient matrices ($\rho = \mathcal{O}(\log(nt/\epsilon))$ if the coefficient distribution exactly matches a power law in one spatial dimension).

VI. APPLICATIONS TO REAL-SPACE QUANTUM SIMULATION

Simulating electronic structure Hamiltonians is one of the most widely studied problems in quantum simulation [4–6]. An efficient solution of the electronic structure problem could lead to better understandings of catalysts and materials, which has applications in numerous subareas of physics and chemistry. Here, we consider mapping the electronic structure Hamiltonian on a grid and performing simulation in the second quantization in real space. Compared to general molecular basis Hamiltonians [32], the grid-based Hamiltonian contains much fewer terms with well-structured coefficients, which is useful for reducing the resource requirement of quantum simulation. We present an algorithm combining our method for Trotter-step implementation with a tighter error analysis, which improves the best simulation results from previous work.

We consider the following class of Hamiltonians:

$$H = T + V := \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} v_{l,m} N_l N_m, \quad (132)$$

where $A_j^\dagger$, A_k are the fermionic creation and annihilation operators, N_l are the occupation-number operators, τ, v are coefficient matrices, and the summations are over n spin orbitals. We can represent the real-space electronic structure Hamiltonians in the above form with specific choices of coefficients τ and v; we come back to this point momentarily. Then, we simulate the Hamiltonian using product formulas by splitting e^{-itH} into products of e^{-itT} and e^{-itV}. The Trotter error corresponding to this splitting was studied in Ref. [87]. There, they found that a pth-order product formula $S_p(t)$ has error scaling like

$$\begin{aligned} &\left\| S_p(t) - e^{-itH} \right\|_{\mathcal{W}_\eta} \\ &= \mathcal{O}\left((\|\tau\| + \|v\|_{\max}\,\eta)^{p-1} \|\tau\| \|v\|_{\max}\,\eta^2 t^{p+1} \right), \end{aligned} \quad (133)$$

where recall $\|\cdot\|_{\max}$ is the max norm denoting the largest matrix element in absolute value, $\|\cdot\|$ is the operator norm and

$$\|X\|_{\mathcal{W}_\eta} := \max_{|\psi_\eta\rangle, |\phi_\eta\rangle \in \mathcal{W}_\eta} \left| \langle \phi_\eta | X | \psi_\eta \rangle \right| \quad (134)$$

is the restriction of the operator norm to the η-electron subspace. Furthermore, if the coefficient matrices τ and v are s sparse (with at most s nonzero elements in each row and column), then it holds that

$$\begin{aligned} &\left\| S_p(t) - e^{-itH} \right\|_{\mathcal{W}_\eta} \\ &= \mathcal{O}\left((\|\tau\|_{\max} + \|v\|_{\max})^{p-1} \|\tau\|_{\max} \|v\|_{\max}\, s^{p+1} \eta t^{p+1} \right). \end{aligned} \quad (135)$$

The new bound we prove is as follows:

$$\begin{aligned} &\left\| S_p(t) - e^{-itH} \right\|_{\mathcal{W}_\eta} \\ &= \mathcal{O}\left(\left(|||\tau|||_1 + |||v|||_{1,[\eta]} \right)^{p-1} |||\tau|||_1 |||v|||_{1,[\eta]}\, \eta t^{p+1} \right), \end{aligned} \quad (136)$$

where we use the induced 1-norm and its restricted version

$$|||\tau|||_1 := \max_j \sum_k |\tau_{j,k}|,$$

$$|||v|||_{1,[\eta]} := \max_j \max_{k_1 < \cdots < k_\eta} \left(|v_{j,k_1}| + \cdots + |v_{j,k_\eta}| \right). \quad (137)$$

This new bound is strictly better than Eq. (135) because for an s-sparse coefficient matrix β, it holds that

$\||\beta\||_{1,[\eta]} \leq \||\beta\||_1 \leq s \|\beta\|_{\max}$ and the equalities are not always attainable. In fact, as we see momentarily, the gap between the induced 1-norm $\||\beta\||_{1,[\eta]}$ and the max norm $s \|\beta\|_{\max}$ is quite significant for the Coulomb interaction, which is relevant for the electronic structure simulation. Compared with Eq. (133), our induced 1-norm scaling $\||\nu\||_{1,[\eta]}$ is again better than $\|\nu\|_{\max} \eta$, but our dependence on τ is slightly worse since $\||\tau\||_1 \geq \|\tau\|$ for a Hermitian τ. Fortunately, this is not an issue with the real-space electronic structure Hamiltonian. To avoid diluting the main message of our work, we state our Trotter error bound below and leave its proof to Appendix B.

Theorem 4: (Trotter error with fermionic induced 1-norm scaling). *Let* $H = T + V := \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} \nu_{l,m} N_l N_m$ *be an interacting-electronic Hamiltonian, and $S_p(t)$ be a pth-order product formula splitting the evolutions under T and V. Then,*

$$
\begin{aligned}
&\left\| S_p(t) - e^{-itH} \right\|_{\mathcal{W}_\eta} \\
&= \mathcal{O}\left(\left(\||\tau\||_1 + \||\nu\||_{1,[\eta]} \right)^{p-1} \||\tau\||_1 \||\nu\||_{1,[\eta]} \eta t^{p+1} \right).
\end{aligned}
\tag{138}
$$

Here, $\||\tau\||_1$ and $\||\nu\||_{1,[\eta]}$ are the (restricted) fermionic induced 1-norm defined by

$$
\||\tau\||_1 = \max_j \sum_k \left| \tau_{j,k} \right|,
$$

$$
\||\nu\||_{1,[\eta]} = \max_j \max_{k_1 < \cdots < k_\eta} \left(\left| \nu_{j,k_1} \right| + \cdots + \left| \nu_{j,k_\eta} \right| \right). \tag{139}
$$

We now apply the above theorem to the simulation of electronic structure Hamiltonians in real space. In this case, we have $H = T + U + V$, where T represents the kinetic term, U represents the external potential term introduced by the Born-Oppenheimer approximation, and V represents the Coulomb potential term. We first consider the uniform electron gas without the external potential U. The Coulomb interaction takes the form [77, Eq. (K4)]

$$
V = \frac{n^{1/3}}{2\omega^{1/3}} \sum_{l,m} \frac{1}{\|l - m\|} N_l N_m, \tag{140}
$$

where n is the number of spin orbitals, ω is the volume of the computational cell, l and m are three-dimensional (3D) vectors with each coordinate taking integer values between $-n^{1/3} - 1/2$ and $n^{1/3} - 1/2$. This term can be represented as $V = \sum_{l,m} \nu_{l,m} N_l N_m$ and it is easy to check that $\|\nu\|_{\max} = \mathcal{O}\left(n^{1/3}/\omega^{1/3}\right)$. On the other hand, we can estimate the (restricted) induced 1-norm as follows. For a fixed value of l, the η spin orbitals nearest to l, if occupied, will have the largest possible coefficients $1/\|l - m\|$. These η spin orbitals form a ball of radius $\mathcal{O}(\eta^{1/3})$, which

is also inside of a cube of linear size $\mathcal{O}(\eta^{1/3})$. This implies that [44, Lemma H.1]

$$
\||\nu\||_{1,[\eta]} = \mathcal{O}\left(\frac{\eta^{2/3} n^{1/3}}{\omega^{1/3}} \right). \tag{141}
$$

We thus see that the scaling of $\||\nu\||_{1,[\eta]}$ is strictly less than $\eta \|\nu\|_{\max}$, giving a factor of $\eta^{1/3}$ improvement over the best previous result. Meanwhile, the scaling for the kinetic term is [120]

$$
\||\tau\||_1 = \mathcal{O}\left(\frac{n^{2/3}}{\omega^{2/3}} \right). \tag{142}
$$

This determines the asymptotic scaling of the number of Trotter steps as

$$
\left(\frac{\eta^{2/3} n^{1/3}}{\omega^{1/3}} + \frac{n^{2/3}}{\omega^{2/3}} \right) \frac{n^{o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}. \tag{143}
$$

To implement each Trotter step, one could sequentially exponentiate all terms in the Hamiltonian with cost $\Theta(n^2)$. However, note that under the Jordan-Wigner encoding, the Coulomb potential term is translated to the spin Hamiltonian

$$
V = \frac{n^{1/3}}{2\omega^{1/3}} \sum_{l,m} \frac{1}{\|l - m\|} \frac{I - Z_l}{2} \frac{I - Z_m}{2}, \tag{144}
$$

with interaction coefficients decaying with distance according to power law. This can be handled by the recursive low-rank method described in Sec. V for the one-dimensional (1D) case with straightforward generalizations in Appendix C 1 to higher spatial dimensions. One can show using the multipole expansion that the decomposed terms can be approximated using matrices with rank polylogarithmic in the input parameters [79,121]. So the complexity of implementing the Coulomb potential term is $\mathcal{O}(n \, \text{polylog}(nt/\epsilon))$.

For the kinetic part, we can either represent it using a finite-difference scheme [76], or using a diagonal form in the kinetic basis and fermionic-Fourier-transforming back [77]. In the first representation, we have geometrically local terms and we can further split them using product formulas without increasing the asymptotic scaling of Trotter error. Alternatively, we can modify boundary terms (using product formulas) so that the coefficients are cyclic, and the kinetic operator can then be implemented using the fermionic fast Fourier transform [76] (we can also use this fermionic-Fourier-transform implementation for the diagonal-form representation in the kinetic basis). As the first derivatives of the electronic wave function are not continuous at electron cusps [122], we note that the error introduced by even a first-order difference scheme

should not differ asymptotically from high-order difference schemes, or even an exact diagonalization of the kinetic part, in the limit of large n. In either case, we can exponentiate the kinetic part with an asymptotic cost of $\mathcal{O}(n \text{ polylog}(nt/\epsilon))$ as well. This gives the total gate complexity

$$\left(\frac{\eta^{2/3} n^{4/3}}{\omega^{1/3}} + \frac{n^{5/3}}{\omega^{2/3}} \right) \frac{n^{o(1)} t^{1+o(1)}}{\epsilon^{o(1)}} \qquad (145)$$

for the uniform electron gas, which simplifies to $\left(\eta^{1/3} n^{1/3} + n^{2/3}/\eta^{2/3} \right) n^{1+o(1)}$ assuming the density of electrons $\eta/\omega = \mathcal{O}(1)$ is fixed.

In second quantization, electronic structure Hamiltonians can be simulated in real space by implementing sequential Trotter steps, which is ancilla-free with the asymptotic gate complexity $\left(n^{8/3}/\eta^{2/3} + n^{7/3}\eta^{2/3} \right) n^{o(1)}$. This complexity can be further improved by diagonalizing the Coulomb potential using the (classical) fast Fourier transform [68,87], although the ancilla space complexity increases to $\widetilde{\mathcal{O}}(n)$. By contrast, our recursive low-rank method combined with the tighter Trotter error bound achieves a strictly lower complexity while using only $\mathcal{O}(\text{polylog}(n))$ ancillas. In first quantization, recent work [77] gives an algorithm based on the interaction-picture simulation [68] with complexity $\widetilde{\mathcal{O}}\left(\eta^{8/3} n^{1/3} \right)$. So our result, while not dominating in all parameter regimes, is asymptotically better as long as $n = \eta^{7/3-o(1)}$. See Table II for a more detailed comparison.

For the general case, we have the external potential term

$$U = \sum_m \left(\sum_l \frac{\zeta_l}{\|\widetilde{r}_l - r_m\|} \right) N_m \qquad (146)$$

introduced under the Born-Oppenheimer approximation, where ζ_l are nuclear charges, $\widetilde{r}_l$ are nuclear coordinates, and r_m are the positions of electrons ($r_m = \omega^{1/3}/n^{1/3} m$). In this case, we have a complexity similar to Eq. (145), but with an additional contribution from the external potential

that depends on

$$\max_m \sum_l \frac{\zeta_l}{\|\widetilde{r}_l - r_m\|}. \qquad (147)$$

We note that this is strictly better than the interaction-picture algorithm from previous work [77], which instead depends on a larger quantity

$$\left(\sum_l \zeta_l \right) \max_{l,m} \frac{1}{\|\widetilde{r}_l - r_m\|} = \mathcal{O}\left(\frac{\eta n^{1/3}}{\omega^{1/3}} \right). \qquad (148)$$

A study of how much this improvement is for practical electronic structure problems will be left as a subject for future work.

VII. FIRST CIRCUIT LOWER BOUND

We have presented three methods for implementing faster Trotter steps and identified applications to electronic structure simulation in second quantization in real space. It is worth mentioning that all our methods make use of structural properties of the Hamiltonian in an essential way. Without such guarantees, the coefficient oracle in the block-encoding and average-cost simulation described in Secs. III and IV would have an implementation cost that depends on the number of terms $\Theta(n^2)$, whereas the blocks in the recursive decomposition of Sec. V would no longer have low rank. Consequently, all our methods will reduce to the sequential circuit implementation of Trotter steps.

In this section, we construct a class of two-local Hamiltonians with only Pauli-Z interactions with coefficients taking a continuum range of values. We show that one needs at least $\Omega(n^2)$ gates to evolve this class of Hamiltonians with accuracy $\epsilon = \Omega(1/\text{poly}(n))$ for time $t = \Omega(\epsilon)$. This suggests that using structural properties of the Hamiltonian is both necessary and sufficient for performing faster Trotter steps. We prove a related circuit lower bound in Appendix A to justify the necessity of structural properties to efficiently implement the coefficient oracle O_β with approximate circuit synthesis.

TABLE II. Comparison of our result and previous results for simulating the uniform electron gas with n spin orbitals and η electrons. We use $\widetilde{\mathcal{O}}(\cdot)$ to suppress polylog(n) factors in the complexity scaling. The complexity of the recursive low-rank method follows from Theorem 3 (with its 3D extension discussed in Appendix C 1), as well as the tighter Trotter error bound described in Theorem 4.

Simulation algorithms	Gate complexity		Space complexity	
	n, η	$\eta = \Theta(n)$	System	Ancilla
First-quantized qubitization [116]	$\widetilde{\mathcal{O}}\left(n^{2/3}\eta^{4/3} + n^{1/3}\eta^{8/3} \right)$	$\widetilde{\mathcal{O}}\left(n^3 \right)$	$\widetilde{\mathcal{O}}(\eta)$	$\widetilde{\mathcal{O}}(\eta)$
First-quantized interaction picture [116]	$\widetilde{\mathcal{O}}\left(n^{1/3}\eta^{8/3} \right)$	$\widetilde{\mathcal{O}}\left(n^3 \right)$	$\widetilde{\mathcal{O}}(\eta)$	$\widetilde{\mathcal{O}}(\eta)$
Second-quantized interaction picture [68]	$\widetilde{\mathcal{O}}\left(n^{8/3}/\eta^{2/3} \right)$	$\widetilde{\mathcal{O}}\left(n^2 \right)$	$\widetilde{\mathcal{O}}(n)$	$\widetilde{\mathcal{O}}(n)$
Second-quantized Trotterization [87] (sequential)	$\left(n^{8/3}/\eta^{2/3} + n^{7/3}\eta^{2/3} \right) n^{o(1)}$	$n^{3+o(1)}$	$\widetilde{\mathcal{O}}(n)$	$\widetilde{\mathcal{O}}(1)$
Second-quantized Trotterization [87] (fast Fourier transform)	$\left(n^{5/3}/\eta^{2/3} + n^{4/3}\eta^{2/3} \right) n^{o(1)}$	$n^{2+o(1)}$	$\widetilde{\mathcal{O}}(n)$	$\widetilde{\mathcal{O}}(n)$
Second-quantized Trotterization (recursive low rank)	$\left(n^{5/3}/\eta^{2/3} + n^{4/3}\eta^{1/3} \right) n^{o(1)}$	$n^{5/3+o(1)}$	$\widetilde{\mathcal{O}}(n)$	$\widetilde{\mathcal{O}}(1)$

The proof of our circuit lower bound is based on a gate-efficient reduction to the problem of approximately synthesizing diagonal unitaries in the Hamming weight-2 subspace, which we describe in detail in Sec. VII B. With this strategy in mind, we first examine the approximate synthesis of diagonal unitaries in Sec. VII A.

A. Approximate synthesis of diagonal unitaries

We now consider the gate complexity lower bound for synthesizing diagonal unitaries. The exact version of this problem was considered by Bullock and Markov [83]. Here, we generalize their lower bound to the approximate synthesis, by adapting a volume comparison technique of Knill [86].

To be specific, our goal is to implement μ-qubit diagonal unitaries with accuracy δ

$$|x_\mu, \ldots, x_2, x_1\rangle \mapsto e^{i\theta_x}|x_\mu, \ldots, x_2, x_1\rangle, \quad (149)$$

where the phase angles take values up to $|\theta_x| \leq \theta_{\max} < \pi$. Recall that the set of all such unitaries is denoted by $\mathcal{D}_{\theta_{\max}}$. Our circuit acts on a total number of $b \geq \mu$ qubits. We first prove a lower bound for a two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$, and then bootstrap the result to obtain a lower bound for arbitrary two-qubit gates.

As aforementioned, this lower bound is proved using a volume comparison argument: we compare the volume of all the diagonal unitaries within δ distance to the set of constructable quantum circuits, against the volume of our target set of diagonal unitaries. A similar argument was used by Knill to prove a circuit lower bound for synthesizing the full unitary group $\mathcal{U}(2^\mu)$ [86], but our proof here is significantly simplified as we restrict to the set of diagonal unitaries whose volume is easier to compute (see Sec. II D).

We start by bounding the number of b-qubit circuits that can be constructed using g gates from a finite two-qubit gate set $\mathcal{K}$. For each of the g gates, we have $\binom{b}{2}$ choices for the locations in which the two-qubit gate acts and, fixing the locations, we have $|\mathcal{K}|$ choices for the specific gate. This gives the upper bound

$$\left(\binom{b}{2}|\mathcal{K}|\right)^g \quad (150)$$

for the number of distinct circuits we can construct.

Now, to synthesize diagonal unitaries with accuracy δ, we need all of them to fall into the δ ball of some quantum circuits (in spectral norm distance). For each circuit V, the volume of diagonal unitaries within its δ ball can be estimated as follows. We take any two diagonal unitaries U and $\widetilde{U}$ from the δ ball centered at V. By the triangle

inequality,

$$\|U - V\|, \|\widetilde{U} - V\| \leq \delta \quad \Rightarrow \quad \|\widetilde{U} - U\| \leq 2\delta, \quad (151)$$

which implies

$$\left|e^{i\widetilde{\theta}_x} - e^{i\theta_x}\right| \leq 2\delta \quad (152)$$

when restricting to the specific diagonal elements indexed by x. Solving for $\widetilde{\theta}_x$ and θ_x when $\delta < 1$, we have

$$\left|\widetilde{\theta}_x - \theta_x\right| \leq \arcsin\left(2\delta\sqrt{1 - \delta^2}\right). \quad (153)$$

This means $\max_x \max_{\widetilde{\theta}_x, \theta_x} \left|\widetilde{\theta}_x - \theta_x\right| \leq \arcsin\left(2\delta\sqrt{1 - \delta^2}\right)$, so the volume of such diagonal unitaries is upper bounded by $\left(\arcsin\left(2\delta\sqrt{1 - \delta^2}\right)\right)^{2^\mu}$. We thus have that the total volume of synthesizable diagonal unitaries is at most

$$\text{vol}\left(\{U | U \text{ is diagonal unitary}, \exists \text{ circuit } V, \|U - V\| \leq \delta\}\right)$$

$$\leq \left(\binom{b}{2}|\mathcal{K}|\right)^g \max_{\text{circuit } V} \text{vol}\left(\{U | U \text{ is diagonal unitary}, \right.$$

$$\left. \|U - V\| \leq \delta\}\right)$$

$$\leq \left(\binom{b}{2}|\mathcal{K}|\right)^g \text{vol}\left(\mathcal{D}_{\frac{\arcsin\left(2\delta\sqrt{1-\delta^2}\right)}{2}}\right) = \left(\binom{b}{2}|\mathcal{K}|\right)^g$$

$$\times \left(\arcsin\left(2\delta\sqrt{1 - \delta^2}\right)\right)^{2^\mu}. \quad (154)$$

On the other hand, we know that the total volume of the target diagonal unitaries is

$$\text{vol}\left(\mathcal{D}_{\theta_{\max}}\right) = (2\theta_{\max})^{2^\mu}. \quad (155)$$

For our synthesis to succeed, we require that

$$\{U | U \text{ is diagonal unitary}, \exists \text{ circuit } V, \|U - V\| \leq \delta\}$$
$$\supseteq \mathcal{D}_{\theta_{\max}}, \quad (156)$$

which implies

$$\left(\binom{b}{2}|\mathcal{K}|\right)^g \left(\arcsin\left(2\delta\sqrt{1 - \delta^2}\right)\right)^{2^\mu} \geq (2\theta_{\max})^{2^\mu},$$

$$g \geq 2^\mu \cdot \frac{\log\left(\frac{2\theta_{\max}}{\arcsin\left(2\delta\sqrt{1-\delta^2}\right)}\right)}{\log\left(\binom{b}{2}|\mathcal{K}|\right)}. \quad (157)$$

Note that this lower bound is only useful when $\theta_{\max} = \Omega(\delta)$, for otherwise the right-hand side of the bound becomes negative. Intuitively, we would simply use the

identity operator in the circuit synthesis if the target diagonal unitary itself is close to identity. On the other hand, the restriction $\theta_{\max} < \pi$ is not fundamental: if $\theta_{\max} \geq \pi$, we can simply reset it to a smaller value.

We have so far proved a circuit lower bound when the underlying gate set $\mathcal{K}$ is of finite size. If we are allowed to use arbitrary two-qubit gates, then the bound can be modified as follows. We first synthesize the circuit with respect to a fixed finite universal gate set $\mathcal{K}'$, say Clifford+T. This gives a new circuit approximating the original one to accuracy δ with a slightly larger gate complexity $cg \log(g/\delta)$ for some constant $c > 0$. We then invoke the above bound with the number of gates $cg \log(g/\delta)$, accuracy 2δ and the gate set $\mathcal{K}' = $ Clifford+T of constant size. This gives

$$cg \log \left(\frac{g}{\delta} \right) = \Omega \left(2^{\mu} \cdot \frac{\log \left(\frac{\theta_{\max}}{\delta} \right)}{\log b} \right). \quad (158)$$

Solving g using the Lambert-W function [11, Lemma 59], we obtain the same asymptotic scaling

$$g = \Omega \left(2^{\mu} \cdot \frac{\log \left(\frac{\theta_{\max}}{\delta} \right)}{\log b} \right). \quad (159)$$

We summarize this bound in the following theorem.

Theorem 5: (Approximate synthesis of diagonal unitaries). *Consider diagonal unitaries*

$$\mathcal{D}_{\theta_{\max}} = \left\{ \sum_x e^{i\theta_x} |x_{\mu}, \ldots, x_1\rangle\langle x_{\mu}, \ldots, x_1| \mid |\theta_x| \leq \theta_{\max} < \pi \right\}$$

on μ qubits with phase angles at most $\theta_{\max} < \pi$. Given accuracy $0 < \delta < 1$, number of qubits $b \geq \mu$, and two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$,

$$\min \{g \mid \forall U \in \mathcal{D}_{\max}, \exists \text{ circuit } V \text{ on } b \text{ qubits with } g \text{ gates}$$

$$\text{from set } \mathcal{K}, \|U - V\| \leq \delta\}$$

$$\geq 2^{\mu} \cdot \frac{\log \left(\frac{2\theta_{\max}}{\arcsin \left(2\delta\sqrt{1-\delta^2} \right)} \right)}{\log \left(\binom{b}{2} |\mathcal{K}| \right)}. \quad (160)$$

Under the same assumption (with $0 < \delta < 1/2$ and $\theta_{\max} = \Omega(\delta)$) but choosing $\mathcal{K}$ to be the set of arbitrary two-qubit gates,

$$\min \{g \mid \forall U \in \mathcal{D}_{\max}, \exists \text{ circuit } V \text{ on } b \text{ qubits with } g \text{ gates}$$

$$\text{from set } \mathcal{K}, \|U - V\| \leq \delta\}$$

$$= \Omega \left(\frac{2^{\mu}}{\log b} \right). \quad (161)$$

B. Reduction from simulation in the Hamming weight-2 subspace

We now show a circuit lower bound for simulating a class of two-local Hamiltonians with only Pauli-Z interactions. Specifically, we consider $H = \sum_{1 \leq j < k \leq n} \beta_{j,k} Z_j Z_k$, where coefficients take a continuum range of values up to $|\beta_{j,k}| \leq t$. Suppose first that this evolution can be simulated to accuracy δ using circuits acting on $b \geq n$ qubits with gates chosen from a two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$. An analogous lower bound holds when the gate set contains arbitrary two-qubit gates.

The key technical ingredient of our proof is a gate-efficient reduction from the approximate synthesis of diagonal unitaries within the Hamming weight-2 subspace $\mathcal{W}_2$. To elaborate, we rewrite the evolution as

$$e^{-iH} = e^{i\sum_{j<k} \beta_{j,k} I} e^{-i\sum_{j<k} \beta_{j,k} Z_j} e^{-i\sum_{j<k} \beta_{j,k} Z_k}$$

$$\times e^{-i\sum_{j<k} 4\beta_{j,k} \frac{I-Z_j}{2} \frac{I-Z_k}{2}}. \quad (162)$$

It suffices for us to consider only the last exponential since all the remaining ones can be performed with gate complexity $\mathcal{O}(n)$ (with only logarithmic overhead when compiled with respect to a fixed finite universal gate set, say Clifford+T). We have

$$\sum_{j<k} \beta_{j,k} \frac{I-Z_j}{2} \frac{I-Z_k}{2} |e_u \oplus e_v\rangle = \beta_{u,v} |e_u \oplus e_v\rangle, \quad (163)$$

for all $1 \leq u < v \leq n$, where e_u represents the n-bit string with 1 at the uth position and 0 elsewhere and $\oplus$ denotes the modulo-2 addition. Using a conversion mapping that is introduced below, we can then relate the problem of simulating two-local Hamiltonians with coefficients $\beta_{j,k}$, to the synthesis of diagonal unitaries with phase angles $\beta_{j,k}$, within the Hamming weight-2 subspace.

To complete the reduction, we need to convert the usual computational basis states to the basis states of the Hamming weight-2 subspace $\mathcal{W}_2$. This can be done using an inequality test, followed by a binary-to-unary conversion. Specifically, suppose we have two quantum registers each of length $\log n$ holding the binary representation of j and k (assuming n is a power of 2). We then use an inequality test with cost $\mathcal{O}(\log n)$ to implement

$$|j\rangle|k\rangle \mapsto |j\rangle|k\rangle|j\,?k\rangle = \begin{cases} |j\rangle|k\rangle|j = k\rangle, & j = k, \\ |j\rangle|k\rangle|j < k\rangle, & j < k, \\ |j\rangle|k\rangle|j > k\rangle, & j > k. \end{cases} \quad (164)$$

We can introduce arbitrary phases for the first case using $\mathcal{O}(n)$ gates (which can again be compiled using Clifford+T with only logarithmic overhead). For the case where $j < k$, we apply the binary-to-unary conversion with cost $\mathcal{O}(n)$
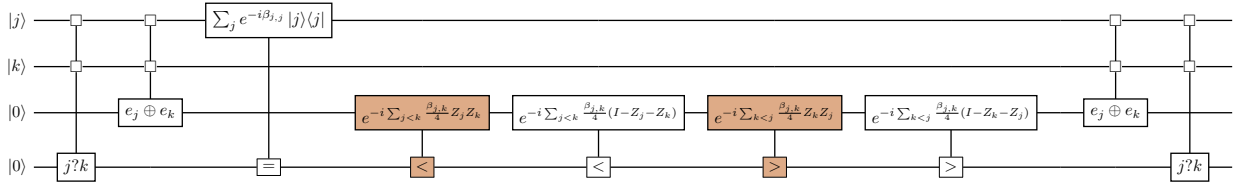
FIG. 4.   Illustration of the reduction. This circuit implements diagonal unitaries on registers $|j\rangle$ and $|k\rangle$ by simulating two-local commuting Hamiltonians on an ancilla register (shaded orange). The reduction is gate efficient, so the asymptotic cost of quantum simulation is the same as that of synthesizing diagonal unitaries.

to get

$$|j\rangle|k\rangle|j < k\rangle \mapsto |j\rangle|k\rangle|e_j \oplus e_k\rangle|j < k\rangle, \qquad (165)$$

after which we introduce the desired phases by simulating the corresponding two-local Hamiltonian with accuracy ϵ. The last case $j > k$ can be handled similarly by swapping the role of j and k. See Fig. 4 for an illustration of this reduction.

To summarize, we can use the above circuit to implement arbitrary diagonal unitaries on the two quantum registers $|j\rangle|k\rangle$ of total length $\log n + \log n = \log(n^2)$ to accuracy 3ϵ with phase angles taking values up to $|\beta_{j,k}| \leq t$. Our circuit uses two instances of controlled simulations of two-local Hamiltonians, together with $\mathcal{O}(n \operatorname{polylog}(n/\epsilon))$ additional Clifford+T gates, acting on a total number of $\mathcal{O}(b)$ qubits (recall $b \geq n$). In other words, if the simulation uses g gates from $\mathcal{K}$, then we can implement arbitrary diagonal unitaries using

$$2g + \mathcal{O}(n \operatorname{polylog}(n/\epsilon)) \qquad (166)$$

gates from the gate set

$$\mathcal{K}' = \text{controlled-}\mathcal{K} \cup (\text{Clifford+}T). \qquad (167)$$

However, our above lower bound indicates that

$$\Omega\left(\frac{n^2}{\log(b|\mathcal{K}'|)}\right) \qquad (168)$$

gates are necessary to implement diagonal unitaries acting on $b = \log(n^2)$ qubits with accuracy 3ϵ for $t = \Omega(\epsilon)$. Thus, a similar asymptotic bound applies to the uncontrolled simulation as well by observing that $|\mathcal{K}'| = \Theta(|\mathcal{K}|)$.

Corollary 1: (Simulating two-local commuting Hamiltonians). *Consider two-local Hamiltonians*

$$H = \sum_{1 \leq j < k \leq n} \beta_{j,k} Z_j Z_k,$$

where coefficients take values up to $|\beta_{j,k}| \leq t$. Given accuracy $0 < \epsilon < 1/3$, the number of qubits $b \geq n$, and

two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$, if $t = \Omega(\epsilon)$,

$$\min\left\{g \mid \forall \text{ 2-local Hamiltonian } H, \exists \text{ circuit } V \text{ on}\right.$$
$$\left. b \text{ qubits with } g \text{ gates from } \mathcal{K}, \left\|e^{-iH} - V\right\| \leq \epsilon\right\}$$
$$= \Omega\left(\frac{n^2}{\log(b|\mathcal{K}|)} - n\operatorname{polylog}\left(\frac{n}{\epsilon}\right)\right). \qquad (169)$$

Under the same assumption but choosing $\mathcal{K}$ to be the set of arbitrary two-qubit gates,

$$\min\left\{g \mid \forall \text{two-local Hamiltonian } H, \exists \text{ circuit } V \text{ on}\right.$$
$$\left. b \text{ qubits with } g \text{ gates from } \mathcal{K}, \left\|e^{-iH} - V\right\| \leq \epsilon\right\}$$
$$= \Omega\left(\frac{n^2}{\log b} - n\operatorname{polylog}\left(\frac{n}{\epsilon}\right)\right). \qquad (170)$$

To apply our bound for implementing Trotter steps, we first divide the evolution into r steps and simulate each step using product formulas for time t/r with accuracy ϵ/r. Depending on the specific definition of the Hamiltonian, we may perform further decompositions like in Secs. III–V and apply the bound only to commuting terms. We note that when the Hamiltonian coefficients have nonuniform magnitudes, the minimum effective evolution time can often be much smaller. For instance, the effective time for power-law interactions scale like $\Theta(t/n^\alpha r)$ as the smallest terms have size $\Theta(1/n^\alpha)$. To apply our bound, we would then require that $t = \Omega(n^\alpha \epsilon)$, which is a pessimistic estimate. Of course, one can replace the volume formula in Eq. (155) with a better estimate that takes account of the nonuniformness of the Hamiltonian coefficients. However, we emphasize that this modification will not lead to a significant improvement to our bound as the dependence on the magnitude of coefficients is only logarithmic. We leave a detailed study of such improvement for future work.

VIII. DISCUSSION

In this work, we have developed circuit implementations of Trotter steps going beyond the sequential approach that cycles through all terms in the Hamiltonian. We have presented three methods, one based on the block-encoding

technique, one based on an average-cost simulation and one based on a recursive low-rank decomposition. This allows us to simulate systems with power-law decaying interactions with significantly lower cost. As an application, we have given an algorithm for simulating electronic structure Hamiltonians in real space, with complexity $\left(\eta^{1/3} n^{1/3} + n^{2/3}/\eta^{2/3}\right) n^{1+o(1)}$ for the uniform electron gas and a similar cost when the external potential is included, improving over the best results from previous work. In achieving these speedups, we have made essential use of structural properties of the system. We further construct a class of two-local commuting Hamiltonians on n qubits with continuum range of coefficients whose simulation requires $\Omega\left(n^2\right)$ gates. Our results thus suggest that the use of structural properties of the Hamiltonian is both necessary and sufficient to implement Trotter steps with lower asymptotic gate complexity.

Although we have employed different techniques to develop various faster simulation algorithms, the core idea behind all our improvements is the use of recursion. With suitable recursive decompositions, we have shown that the target Hamiltonian can be significantly simplified with lower 1-norm and smaller rank value, both of which are useful for quantum simulation. Meanwhile, because of the master theorem (Lemma 1), recursions introduce only logarithmic factors to the overall gate complexity, thus simplifying the target problem "almost for free." We expect that recursions will find further applications in reducing the cost of simulating many-body Hamiltonians.

We have focused on simulating two-local spin models throughout our paper, and it would be interesting to study general κ-local models as well as fermionic systems. For this purpose, previous work [41] utilized different low-rank decompositions for one- and two-body fermionic operators, which led to Trotter steps with lower cost. However, an issue not rigorously addressed in their work is that the decomposition does not respect the commutation relations of Hamiltonian terms and could lead to a Trotter error larger than the naive approach. Thus, some additional efforts are needed to justify the utility of those methods in quantum simulation [123]. Note that this drawback can be slightly alleviated using the phase-gradient construction [54] if the fault-tolerant resource cost is of interest. On the other hand, our techniques do apply to other κ-local models for constant κ and to fermionic systems, as long as the Hamiltonian can be recursively decomposed as in Secs. III–V. Preliminary studies of such generalizations were available in the context of measuring quantum observables [114], and we hope future work could develop more efficient Hamiltonian decompositions along this line.

We have demonstrated that advanced quantum simulation techniques, such as block encoding and qubitization, can be used to improve the performance of product formulas. Such techniques can be directly applied to implement Trotter steps; however, this would introduce a slow-down factor proportional to the 1-norm of Hamiltonian coefficients, offering no benefit over the sequential circuit implementation. We have shown how to overcome this obstacle for power-law interactions by recursively decomposing the Hamiltonian using product formulas to reduce the norm and by implementing an average-case cost simulation. Our result is complementary to recent results on the hybridized simulation [124,125], where quantum algorithms are "hybridized" with the interaction-picture method to achieve better performance. Our finding indicates that product formulas can be just as effective in hybridizing with other algorithms for quantum simulation.

In our low-rank method, we have performed a recursive decomposition of the target Hamiltonian where certain off-diagonal blocks of coefficients have rank value logarithmic in the input parameters. Such a decomposition was used in a different context by Ref. [52] to block-encode kernel matrices. Although their results are potentially useful for improving electronic structure simulation in first quantization, we are not aware of a simple circuit to realize this idea. In our problem, the decomposed Hamiltonian terms can be handled directly, so block encoding is no longer needed (it would be inefficient anyway due to the normalization factor issue mentioned above). We apply this method in Sec. VI to implement the Coulomb potential term of the electronic structure Hamiltonian in second quantization in real space, obtaining improvements over the best previous simulation results. As can be seen from Table I, this method achieves the lowest gate complexity among all our algorithms as long as the low-rank condition is satisfied. This condition can be rigorously justified using the multipole expansion for all power-law models in 1D and many power-law models such as Coulomb interactions in higher spatial dimensions. For other Hamiltonians where the rank condition does not hold, faster Trotter steps are still possible using methods based on block encoding and average-cost simulation.

When no structural property is available in the target Hamiltonian, then intuitively the best method one can hope for is to exponentiate the terms one by one. We justify this intuition in Sec. VII with a circuit lower bound, showing that a class of two-local commuting Hamiltonians on n sites requires at least $\Omega\left(n^2\right)$ gates to simulate. However, we emphasize that our lower bound is useful only when the effective evolution time is sufficiently long: if the effective time is short compared to the accuracy threshold, our bound becomes trivial. However, such failures are often indications that improved implementations of Trotter steps exist, as one can truncate the Hamiltonian and replace small-angle exponentials by identity operators. Previous work has studied various truncation schemes for product formulas as well as more advanced quantum algorithms [44–49]. By exploiting the relative scaling between evolution time and system size, it is plausible to get faster quantum simulation algorithms with better dependence on the spacetime volume.

Although our lower bound is mostly used to argue the difficulty of implementing product formulas, it is applicable to other simulation algorithms as well. In particular, it applies to the sampling-based approach [20], as one can always reproduce classical computations on a quantum computer, and the use of randomness can be represented unitarily via the Stinespring dilation. A similar argument then lower bounds both the classical and quantum resources used by the simulation. But given the difficulty of realizing a scalable quantum computer, it would be beneficial if certain computations can be offloaded classically, which is exactly a main goal of the sampling-based algorithm. Looking forward, it would be interesting to further study the tradeoffs between classical and quantum resources in designing quantum algorithms.

Several other related questions are worth further investigation. To simplify our analysis, we have used boxes with exponentially growing sizes to group each qubit index in the block-encoding approach, but this may be improved by using different box definitions. We have also chosen a uniform division of intervals in the realization of average-cost simulation, but improvements might be possible through a nonuniform division scheme. We have almost exclusively considered the number of arbitrary two-qubit gates as the cost metric for quantum simulation, and future work may consider other metrics such as circuit depth, locality [126] and fault-tolerant gate count (although we expect some of our techniques are useful under those metrics as well). With these questions in mind, we hope our work will motivate continuing progress on finding the most efficient methods to solve instances of quantum simulation problems with practical interest.

APPENDIX A: SECOND CIRCUIT LOWER BOUND

In Sec. VII, we proved a lower bound on the gate complexity of approximately synthesizing diagonal unitaries, and bootstrapped it to get a lower bound on the simulation of two-local commuting Hamiltonians. In this Appendix, we prove a related lower bound on the cost of synthesizing the coefficient oracle O_β used in Secs. III and IV for block encoding. To this end, we first prove a lower bound for the approximate synthesis of discrete diagonal unitaries in Appendix A 1, and then do a reduction from applications of the coefficient oracle in Appendix A 2. Our result extends a previous bound of Ref. [107] by allowing approximate circuit synthesis and use of arbitrary two-qubit gates.

1. Approximate synthesis of discrete diagonal unitaries

We first consider the gate complexity of synthesizing discrete diagonal unitaries of the form

$$|x\rangle \mapsto e^{i\beta_{x,m-1}2^{m-1}+\cdots+\beta_{x,0}/2^m}|x\rangle, \tag{A1}$$

with accuracy δ. Here, x takes 2^μ possible values, and β_x is an m-bit string parameterized by x (so $\beta_{x,j} \in \{0, 1\}$). We use a setup similar to that of Sec. VII. Specifically, our circuit acts on a total number of $b \geq \mu$ qubits. We first prove a lower bound for a two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$, and then bootstrap the result to obtain a lower bound for arbitrary two-qubit gates.

As before, the number of distinct b-qubit circuits that can be constructed using g gates from gate set $\mathcal{K}$ is at most

$$\left(\binom{b}{2}|\mathcal{K}|\right)^g. \tag{A2}$$

We now repeat our volume argument. To synthesize all discrete diagonal unitaries with accuracy δ, we need all of them to fall into the spectral norm δ ball of some quantum circuits. Fixing a specific circuit V, we consider any diagonal unitaries U and $\widetilde{U}$ from the δ ball centered at V. Because of the triangle inequality, we must have

$$\left\|\widetilde{U} - U\right\| \leq 2\delta. \tag{A3}$$

However, the new observation is that our target set is discrete and finite. For $\delta = \mathcal{O}(1)$ and $\delta = \Omega(1/2^m)$, these unitaries have the xth diagonal entries

$$\left|e^{i\widetilde{\beta}_{x,m-1}2^{m-1}+\cdots+\widetilde{\beta}_{x,0}/2^m} - e^{i\beta_{x,m-1}2^{m-1}+\cdots+\beta_{x,0}/2^m}\right| \leq 2\delta. \tag{A4}$$

Solving for the (integers) $\widetilde{\beta}_x$ and β_x, we have

$$\left|\widetilde{\beta}_x - \beta_x\right| \leq 2^m \arcsin\left(2\delta\sqrt{1 - \delta^2}\right), \tag{A5}$$

which implies that the number of such discrete diagonal unitaries is at most

$$\mathcal{O}\left(\left(2^m\delta\right)^{2^\mu}\right). \tag{A6}$$

Using the counting measure $\#\{\cdot\}$, we thus have that the total number of synthesizable discrete diagonal unitaries is at most

$$\#\{U|U \text{ is discrete diagonal unitary}, \exists \text{ circuit } V, \|U - V\| \leq \delta\}$$

$$\leq \left(\binom{b}{2}|\mathcal{K}|\right)^g \max_{\text{circuit } V} \#\{U|U \text{ is discrete diagonal unitary}, \|U - V\| \leq \delta\}$$

$$= \left(\binom{b}{2}|\mathcal{K}|\right)^g \mathcal{O}\left((2^m\delta)^{2^\mu}\right) = \mathcal{O}\left(b^{2g}|\mathcal{K}|^g 2^{m2^\mu} \delta^{2^\mu}\right). \tag{A7}$$

On the other hand, by a counting argument, we know that the total number of discrete diagonal unitaries is exactly

$$\#\{U|U \text{ is discrete diagonal unitary}\} = \left(2^m\right)^{2^\mu} = 2^{m2^\mu}. \tag{A8}$$

For our synthesis to succeed, we require that

$$\{U \text{ is discrete diagonal unitary} \mid \exists \text{ circuit } V, \|U - V\| \leq \delta\}$$

$$\supseteq \{U \text{ is discrete diagonal unitary}\}, \tag{A9}$$

which implies

$$g = \Omega\left(2^\mu \frac{\log\left(\frac{1}{\delta}\right)}{\log\left(b|\mathcal{K}|\right)}\right). \tag{A10}$$

If $\mathcal{K}$ is the arbitrary two-qubit gate set, we can use the Lambert-W function in a similar way as in Sec. VII A to get

$$g = \Omega\left(2^\mu \frac{\log\left(\frac{1}{\delta}\right)}{\log b}\right). \tag{A11}$$

We have thus proved the following.

Theorem 6: (Approximate synthesis of discrete diagonal unitaries). *Consider discrete diagonal unitaries*

$$\left\{\sum_x e^{i\beta_{x,m-1}2^{m-1}+\cdots+\beta_{x,0}/2^m}|x\rangle\langle x|\right\}$$

on μ qubits with phase angles determined by m-bit strings β_x. Given accuracy $\delta = \mathcal{O}(1)$ and $\delta = \Omega(1/2^m)$, the number of qubits $b \geq \mu$, and two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$,

$$\min\{g \mid \forall \text{ discrete diagonal unitary } U, \exists \text{ } b\text{-qubit circuit}$$

$$V \text{ with } g \text{ gates from set } \mathcal{K}, \|U - V\| \leq \delta\}$$

$$= \Omega\left(\frac{2^\mu \log\left(\frac{1}{\delta}\right)}{\log\left(b|\mathcal{K}|\right)}\right). \tag{A12}$$

Under the same assumption but choosing $\mathcal{K}$ to be the set of arbitrary two-qubit gates,

$$\min\{g \mid \forall \text{ discrete diagonal unitary } U, \exists \text{ } b\text{-qubit circuit}$$

$$V \text{ with } g \text{ gates from set } \mathcal{K}, \|U - V\| \leq \delta\}$$

$$= \Omega\left(\frac{2^\mu \log\left(\frac{1}{\delta}\right)}{\log b}\right). \tag{A13}$$

2. Reduction from the coefficient oracle

Recall that in the block-encoding method, we use the following oracle of Hamiltonian coefficients:

$$O_\beta|u, v, 0\rangle = |u, v, \beta_{u,v}\rangle. \tag{A14}$$

Here, u and v each take n values and $\beta_{u,v}$ are m-bit binary strings. To ensure the efficiency of our method, we assume that the gate complexity of implementing oracle O_β is $\mathcal{O}(\text{polylog}(nt/\epsilon))$. Implicitly, this implies the existence of certain structural properties of $\beta_{u,v}$: we now show that if $\beta_{u,v}$ are allowed to depend arbitrarily on u and v, then one needs at least approximately n^2 gates to implement this oracle, resulting in an inefficient block-encoding approach.

The key ingredient of the proof is a reduction from the application of coefficient oracle to the approximate synthesis of discrete diagonal unitaries. This is done as follows:

$$|u, v, 0\rangle \overset{O_\beta}{\mapsto} |u, v, \beta_{u,v}\rangle$$

$$\mapsto e^{i\widetilde{\beta}_{u,v}/2^m}|u, v, \beta_{u,v}\rangle$$

$$\overset{O_\beta^\dagger}{\mapsto} e^{i\widetilde{\beta}_{u,v}/2^m}|u, v, 0\rangle, \tag{A15}$$

where in the second step we introduce the approximate phase factors $\widetilde{\beta}_{u,v}$ by applying $\Theta\left(\log\left(1/\epsilon\right)\right)$ rotations $|0\rangle\langle 0| + e^{i/2^\ell}|1\rangle\langle 1|$ ($\ell = 1, \ldots, \Theta\left(\log\left(1/\epsilon\right)\right)$). We thus implement an arbitrary discrete diagonal unitary on registers $|u, v\rangle$ of total length $\mu = 2 \log n$ using a circuit of two oracle calls (each with accuracy ϵ) and $\Theta\left(\log\left(1/\epsilon\right)\right)$ single-qubit gates (which can be compiled with respect to a fixed gate set like Clifford+T with only logarithmic overhead).

In summary, suppose O_β can be implemented to accuracy ϵ using a b-qubit circuit of g gates chosen from

a gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$. Then, we can implement the entire reduction to accuracy 3ϵ using a b-qubit circuit of $2g + \mathcal{O}(\text{polylog}(1/\epsilon))$ gates chosen from $\mathcal{K}' = \mathcal{K} \cup (\text{Clifford}+T)$. But our lower bound indicates that

$$\Omega\left(\frac{n^2 \log\left(\frac{1}{\epsilon}\right)}{\log\left(b|\mathcal{K}'|\right)}\right) = \Omega\left(\frac{n^2 \log\left(\frac{1}{\epsilon}\right)}{\log\left(b|\mathcal{K}|\right)}\right) \qquad \text{(A16)}$$

gates are required for the synthesis task. We have thus proved the following.

Corollary 2: (Approximate synthesis of coefficient oracles). *Consider the set of coefficient oracles*

$$\{O_\beta \mid O_\beta |u, v, 0\rangle = |u, v, \beta_{u,v}\rangle\},$$

where u and v each takes n values and $\beta_{u,v}$ are m-bit strings. Given accuracy $\epsilon = \mathcal{O}(1)$ and $\epsilon = \Omega\left(1/2^m\right)$, the number of qubits $b \geq m + 2\log n$, and two-qubit gate set $\mathcal{K}$ of finite size $|\mathcal{K}|$,

$$\min\left\{g \mid \forall \text{oracle } O_\beta, \exists \text{ circuit } V \text{ on } b \text{ qubits with}\right.$$
$$\left. g \text{ gates from set } \mathcal{K}, \|O_\beta - V\| \leq \epsilon\right\}$$
$$= \Omega\left(\frac{n^2 \log\left(\frac{1}{\epsilon}\right)}{\log\left(b|\mathcal{K}|\right)} - \text{polylog}\left(\frac{1}{\epsilon}\right)\right). \qquad \text{(A17)}$$

Under the same assumption but choosing $\mathcal{K}$ to be the set of arbitrary two-qubit gates,

$$\min\left\{g \mid \forall \text{oracle } O_\beta, \exists \text{ circuit } V \text{ on } b \text{ qubits with}\right.$$
$$\left. g \text{ gates from set } \mathcal{K}, \|O_\beta - V\| \leq \epsilon\right\}$$
$$= \Omega\left(\frac{n^2 \log\left(\frac{1}{\epsilon}\right)}{\log b} - \text{polylog}\left(\frac{1}{\epsilon}\right)\right). \qquad \text{(A18)}$$

We see that implementations of the coefficient oracle O_β will generally cost approximately n^2 gates as well, even when approximate circuit synthesis is allowed. Thus, one needs structural properties of the Hamiltonian not only to realize faster Trotter steps but to implement faster coefficient oracles as well.

APPENDIX B: TROTTER ERROR WITH FERMIONIC INDUCED 1-NORM SCALING

In this Appendix, we prove the fermonic induced 1-norm scaling of Trotter error claimed in Theorem 4, which is in turn used in the simulation of real-space electronic structure Hamiltonians. Specifically, we consider the Hamiltonian

$$H = T + V := \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} v_{l,m} N_l N_m, \qquad \text{(B1)}$$

where $A_j^\dagger$, A_k are the fermionic creation and annihilation operators, N_l are the occupation-number operators, τ, v are coefficient matrices, and the summations are over n spin orbitals. Let $S_p(t)$ be a pth-order product formula that splits e^{-itH} into products of e^{-itT} and e^{-itV}. Our goal is to bound the Trotter error $\|S_p(t) - e^{-iH}\|_{\mathcal{W}_\eta}$ restricted to the η-electron subspace $\mathcal{W}_\eta$ (spanned by basis states with Hamming weight η).

To describe our proof, we introduce some additional notations and terminologies. For $\gamma = 0, 1$, we let

$$H^{(\gamma)} = \sum_{j,k} \mu_{j,k}^{(\gamma)} H_{j,k}^{(\gamma)}, \qquad \mu^{(0)} = v, \qquad \mu^{(1)} = \tau,$$

$$H_{j,k}^{(0)} = N_j N_k, \qquad H_{j,k}^{(1)} = A_j^\dagger A_k. \qquad \text{(B2)}$$

In other words, $H^{(0)} = V$, $H^{(1)} = T$. Then, we have the commutator scaling of Trotter error from Lemma 2:

$$\left\|S_p(t) - e^{-itH}\right\|_{\mathcal{W}_\eta} = \mathcal{O}\left(\max_{\gamma \in \{0,1\}^{p+1}} \max_{|\psi_\eta\rangle \in \mathcal{W}_\eta} \left|\langle\psi_\eta| \left[H^{(\gamma_{p+1})}, \ldots \left[H^{(\gamma_2)}, H^{(\gamma_1)}\right]\right] |\psi_\eta\rangle\right| t^{p+1}\right), \qquad \text{(B3)}$$

where we take the expectation value since nested commutators of Hermitian operators are necessarily anti-Hermitian. Note the following commutation rules for fermionic operators [127]:

$$\left[A_j^\dagger A_k, A_{j_x}^\dagger\right] = \delta_{k,j_x} A_j^\dagger, \quad \left[A_j^\dagger A_k, A_{k_y}\right] = -\delta_{k_y,j} A_k, \quad \left[A_j^\dagger A_k, N_{l_z}\right] = \delta_{k,l_z} A_j^\dagger A_{l_z} - \delta_{j,l_z} A_{l_z}^\dagger A_k, \qquad \text{(B4)}$$

$$\left[N_l N_m, A_{j_x}^\dagger\right] = \delta_{m,j_x} N_l A_{j_x}^\dagger + \delta_{l,j_x} A_{j_x}^\dagger N_m, \quad \left[N_l N_m, A_{k_x}\right] = -\delta_{m,k_x} N_l A_{k_x} - \delta_{l,k_x} A_{k_x} N_m, \qquad \text{(B5)}$$

where $\delta_{j,k} = 1$ if and only if $j = k$. Thus, the nested commutator $\left[H_{j_{p+1},k_{p+1}}^{(\gamma_{p+1})}, \ldots \left[H_{j_2,k_2}^{(\gamma_2)}, H_{j_1,k_1}^{(\gamma_1)}\right]\right]$ can be written as a linear combination of products of elementary fermionic operators. Following Ref. [87, Proposition 11], we call each summand P from the nested commutator a *fermionic path* and write $P \triangleright \left(H_{j_{p+1},k_{p+1}}^{(\gamma_{p+1})}, \ldots, H_{j_1,k_1}^{(\gamma_1)}\right)$. Then, we have the following bound:

$$\left| \langle \psi_\eta | \left[H^{(\gamma_{p+1})}, \ldots \left[H^{(\gamma_2)}, H^{(\gamma_1)} \right] \right] | \psi_\eta \rangle \right|$$

$$\leq \max_{\|c\|_1 = \eta} \sum_{j_{p+1}, k_{p+1}} \cdots \sum_{j_1, k_1} \sum_{P \triangleright \left(H_{j_{p+1}, k_{p+1}}^{(\gamma_{p+1})}, \ldots, H_{j_1, k_1}^{(\gamma_1)} \right)} \left| \mu_{j_{p+1}, k_{p+1}}^{(\gamma_{p+1})} \right| \cdots \left| \mu_{j_1, k_1}^{(\gamma_1)} \right| \frac{\|P|c\rangle\| + \|P^\dagger|c\rangle\|}{2}, \tag{B6}$$

where the maximization is taken over all n-bit strings c with Hamming weight η, and $P|c\rangle$ $(P^\dagger|c\rangle)$ either gives another basis state or gives the zero vector.

Previous work [87] further proceeds to upper bound the coefficients by the max norm $\|\tau\|_{\max}$ and $\|\nu\|_{\max}$ and counts the number of fermionic paths with nonzero contribution to the summation. This results in a bound with the max norm scaling asymptotically worse than our result. To get our improved Theorem 4, we need a more careful treatment of the summation order as well as the action of fermionic operators on a fixed basis state.

We now present a proof by induction. The base case of $p = 0$ is easy to check. For any n-bit string c with Hamming weight η, we have

$$\sum_{j,k} |\tau_{j,k}| \frac{\left\| A_j^\dagger A_k |c\rangle \right\| + \left\| A_k^\dagger A_j |c\rangle \right\|}{2} = \mathcal{O} \left(\|\|\tau\|\|_1 \eta \right),$$

$$\sum_{l,m} |\nu_{l,m}| \frac{\|N_l N_m |c\rangle\| + \|N_m N_l |c\rangle\|}{2} = \mathcal{O} \left(\|\|\nu\|\|_{1,[\eta]} \eta \right). \tag{B7}$$

This is because in order to make a nonzero contribution, the right-most fermionic operators only have η possible choices. This motivates us to choose the summation ordering so that the indices of operators on the left are summed over first:

$$\sum_k \sum_j |\tau_{j,k}| \left\| A_j^\dagger A_k |c\rangle \right\|, \qquad \sum_j \sum_k |\tau_{j,k}| \left\| A_k^\dagger A_j |c\rangle \right\|,$$

$$\sum_m \sum_l |\nu_{l,m}| \|N_l N_m |c\rangle\|, \qquad \sum_l \sum_m |\nu_{l,m}| \|N_m N_l |c\rangle\|. \tag{B8}$$

Fixing one choice for the right-most operator, the remaining quantity can be upper bounded by either the induced 1-norm $\|\|\tau\|\|_1$ or the restricted induced 1-norm $\|\|\nu\|\|_{1,[\eta]}$, respectively. This proves the desired bound in the base case.

For the inductive step, assume that given any fermionic path $P \triangleright \left(H_{j_p, k_p}^{(\gamma_p)}, \ldots, H_{j_1, k_1}^{(\gamma_1)} \right)$, we can always reorder the summation to yield the desired bound. Now take the next layer of commutator $[T, P]$ or $[V, P]$. By Eqs. (B4) and (B5) and a separate induction, $P = \prod_{q=1}^{p+2} P_\upsilon$ is the product of at

most $p + 2$ elementary fermionic operators (which include $A_j^\dagger, A_k, N_l$), and at any point in the product the number of creation operators is always smaller than the number of annihilation operators. Then, the commutation may be performed sequentially as

$$[T, P] = \sum_{q=1}^{p+2} P_{p+2} \cdots P_{q+1} \left[T, P_q \right] P_{q-1} \cdots P_1,$$

$$[V, P] = \sum_{q=1}^{p+2} P_{p+2} \cdots P_{q+1} \left[V, P_q \right] P_{q-1} \cdots P_1. \tag{B9}$$

Depending on $P_q = A_{j_x}^\dagger, A_{k_y}$, or N_{l_z}, we divide the analysis into five subcases:

(i) $\left[A_j^\dagger A_k, A_{j_x}^\dagger \right] = \delta_{k,j_x} A_j^\dagger$: we keep the original sum over j_x while introducing a new sum over j **on the right**. The summation range of j is set to be over **all n spin orbitals**. The number of electrons increases by 1 due to $A_j^\dagger$ with the new location determined by j (as opposed to j_x). Asymptotically, this does not affect the summation of other existing indices and increases the norm by a factor of $\|\|\tau\|\|_1$ through the index j.

(ii) $\left[A_j^\dagger A_k, A_{k_y} \right] = -\delta_{k_y, j} A_k$: we keep the original sum over k_y while introducing a new sum over k **on the right**. The summation range of k is set to be over **all n spin orbitals**. The number of electrons decreases by 1 due to A_k with the new location determined by k (as opposed to k_y). Asymptotically, this does not affect the summation of other existing indices and increases the norm by a factor of $\|\|\tau\|\|_1$ through the index k.

(iii) $\left[A_j^\dagger A_k, N_{l_z} \right] = \delta_{k,l_z} A_j^\dagger A_{l_z} - \delta_{j,l_z} A_{l_z}^\dagger A_k$: we keep the original sum over l_z while introducing a new sum over j (k) **on the right**, respectively. The summation range of j (k) is set to be over **all n spin orbitals**. The number of electrons remains the same but the occupation changes with the locations determined by j (k) and l_z. Asymptotically, this does not affect the summation of other existing indices and increases the norm by a factor of $\|\|\tau\|\|_1$ through the index j (k).

(iv) $\left[N_l N_m, A_{j_x}^\dagger\right] = \delta_{m,j_x} N_l A_{j_x}^\dagger + \delta_{l,j_x} A_{j_x}^\dagger N_m = \delta_{m,j_x} N_l A_{j_x}^\dagger + \delta_{l,j_x} N_m A_{j_x}^\dagger - \delta_{l,j_x} \delta_{m,j_x} A_{j_x}^\dagger$: we keep the original sum over j_x while introducing a new sum over l (m) **on the right**, respectively. The summation range of l (m) is set to be over **all occupied spin orbitals**. Both the number and the locations of electrons remain the same. Asymptotically, this does not affect the summation of other existing indices and increases the norm by a factor of $\||v\||_{1,[\eta]}$ through the index l (m).

(v) $\left[N_l N_m, A_{k_x}\right] = -\delta_{m,k_x} N_l A_{k_x} - \delta_{l,k_x} A_{k_x} N_m = -\delta_{m,k_x} N_l A_{k_x} - \delta_{l,k_x} N_m A_{k_x} - \delta_{l,k_x} \delta_{m,k_x} A_{k_x}$: we keep the original sum over k_x while introducing a new sum over l (m) **on the right**, respectively. The summation range of l (m) is set to be over **all occupied spin orbitals**. Both the number and the locations of electrons remain the same. Asymptotically, this does not affect the summation of other existing indices and increases the norm by a factor of $\||v\||_{1,[\eta]}$ through the index l (m).

Note that in the fourth and fifth cases, we implicitly use the fact that the number of occupied modes is no larger than η. The proof of Theorem 4 is now complete.

APPENDIX C: GENERALIZATION

1. Higher spatial dimensionality

In the main body of the paper, we have presented faster methods to perform Trotter steps for simulating power-law interactions in one spatial dimension. Here, we briefly describe how to generalize those methods to higher spatial dimensionality.

Specifically, consider an n-qubit d-dimensional square lattice $\mathcal{L} \subseteq \mathbb{Z}^d$. Without loss of generality, we assume $n^{1/d}$ is a power of 2 and take $\mathcal{L} = \{1, \ldots, n^{1/d}\}^d$. Then, the Hamiltonian has the form $H = \sum_{j,k \in \mathcal{L}} H_{j,k}$. For power-law interactions, we have the norm condition $\|H_{j,k}\| \leq 1/\|j-k\|^\alpha$ for constant $\alpha > 0$, where $\|j-k\|$ is simply the Euclidean distance between d-dimensional vectors j and k. Just like in the 1D case, we can expand the Hamiltonian with respect to tensor products of Pauli operators and hereafter assume the Hamiltonian takes the form

$$H = \sum_{\substack{j,k \in \mathcal{L} \\ j \neq k}} \beta_{j,k} X_j Y_k. \tag{C1}$$

The commutator norm Eq. (28) corresponding to such decompositions has the scaling [44, Theorem H.2]

$$\beta_{\text{comm}}{}^{o(1)} = \begin{cases} n^{o(1)}, & \alpha \geq d, \\ n^{1-\frac{\alpha}{d}+o(1)}, & 0 < \alpha < d, \end{cases} \tag{C2}$$

which implies

$$r = \begin{cases} \dfrac{n^{o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}, & \alpha \geq d, \\[3mm] \dfrac{n^{1-\alpha/d+o(1)} t^{1+o(1)}}{\epsilon^{o(1)}}, & 0 < \alpha < d. \end{cases} \tag{C3}$$

Depending on the order $j_s < k_s$ or $j_s > k_s$ between each pair of the corresponding coordinates, we can further divide the Hamiltonian into 2^d subterms. The remaining degenerate cases where some pair of indices collide $j_s = k_s$ can be merged accordingly. We assume that

$$H = \sum_{1 \leq j_1 < k_1 \leq n^{1/d}} \cdots \sum_{1 \leq j_d < k_d \leq n^{1/d}} \beta_{j,k} X_j Y_k \tag{C4}$$

in the following discussion, and a similar analysis holds for other cases with some notation changes.

We now perform a recursive decomposition along each spatial dimension of the system. So for the first spatial dimension, we have the term

$$\sum_{\substack{1 \leq j_1 < n^{1/d}/2 \\ n^{1/d}/2 \leq k_1 \leq n^{1/d}}} \cdots \sum_{1 \leq j_d < k_d \leq n^{1/d}} \beta_{j,k} X_j Y_k \tag{C5}$$

in the first layer and terms

$$\sum_{\substack{1 \leq j_1 < n^{1/d}/4 \\ n^{1/d}/4 \leq k_1 < n^{1/d}/2}} \cdots \sum_{1 \leq j_d < k_d \leq n^{1/d}} \beta_{j,k} X_j Y_k,$$

$$\sum_{\substack{n^{1/d}/2 \leq j_1 < 3n^{1/d}/4 \\ 3n^{1/d}/4 \leq k_1 \leq n^{1/d}}} \cdots \sum_{1 \leq j_d < k_d \leq n^{1/d}} \beta_{j,k} X_j Y_k \tag{C6}$$

in the second layer and so on. Because of the master theorem Lemma 1, we may restrict to a single term in the decomposition, say the one in the first layer. We then perform another recursive decomposition along the second spatial dimension, and once again restrict to only a specific term in the decomposition using Lemma 1. We iterate through all the d spatial dimensions. See Fig. 5 for an illustration of this strategy in the 2D case.

For the block-encoding method in Sec. III, the cost of circuit implementation depends on the 1-norm of Hamiltonian coefficients, which determines the number of steps in the qubitization algorithm. Specifically, the cost of implementing decomposed Hamiltonian terms should now be revised to
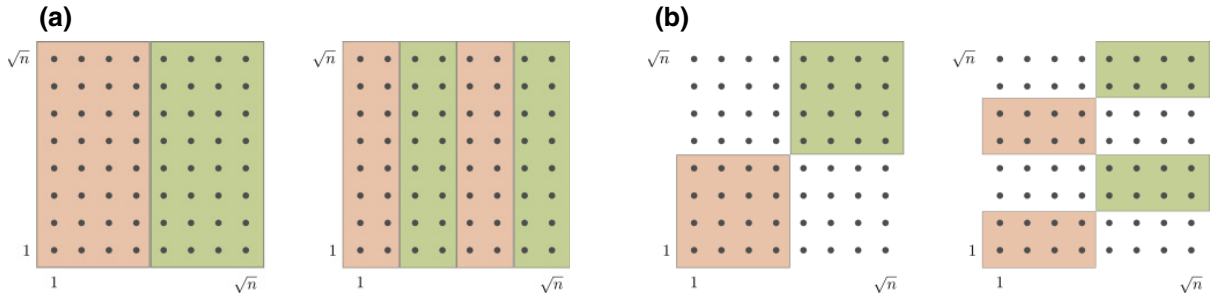
(a)

(b)

FIG. 5.   Illustration of the 2D recursive decomposition. Decomposed terms act across two disjoint cubes. The colors indicate types of the operator support following the convention in Fig. 1. Subfigure (a) represents terms in the first two layers of the decomposition of the first spatial dimension. Subfigure (b) recursively decomposes the first layer from (a) in the second spatial dimension.

$$\text{cost}(n) = \begin{cases} \mathcal{O}\left(\left(\frac{t}{r} + 1\right) n \operatorname{polylog}\left(\frac{nt}{\epsilon}\right)\right), & \alpha \geq 2d, \\ \mathcal{O}\left(\left(n^{2-\frac{\alpha}{d}} \frac{t}{r} + 1\right) n \operatorname{polylog}\left(\frac{nt}{\epsilon}\right)\right), & 0 < \alpha < 2d, \end{cases} \tag{C7}$$

which implies that the entire simulation has a complexity of

$$\begin{aligned} & nt\left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq 2d, \\ & \mathcal{O}\left(n^{3-\alpha/d} t \operatorname{polylog}\left(\frac{nt}{\epsilon}\right)\right) + nt\left(\frac{nt}{\epsilon}\right)^{o(1)}, & d \leq \alpha < 2d, \\ & \mathcal{O}\left(n^{3-\alpha/d} t \operatorname{polylog}\left(\frac{nt}{\epsilon}\right)\right) + n^{2-\frac{\alpha}{d}} t\left(\frac{nt}{\epsilon}\right)^{o(1)}, & 0 < \alpha < d. \end{aligned} \tag{C8}$$

This simulation can be further improved for $\alpha < 2d$ using the average-cost simulation technique described in Sec. IV. Specifically, for the two cubes sharing a common vertex, we further divide each spatial dimension into m intervals of equal length; see Fig. 6 for an illustration of this decomposition in 2D. Then, considering evolution time t/r, we have the modified cost function

$$\text{cost}(n) = \tilde{\mathcal{O}}\left(\left(n^{2-\alpha/d} \frac{t}{r} + m^{2d}\right) \frac{n}{m^d}\right), \tag{C9}$$

which implies that the entire simulation has cost

$$\begin{aligned} & \min\left\{n^{2-\alpha/2d} t\left(\frac{nt}{\epsilon}\right)^{o(1)}, \ \mathcal{O}\left(n^{3-\alpha/d} t \operatorname{polylog}\left(\frac{nt}{\epsilon}\right)\right) + nt\left(\frac{nt}{\epsilon}\right)^{o(1)}\right\}, & d \leq \alpha < 2d, \\ & \min\left\{n^{5/2-\frac{\alpha}{d}} t\left(\frac{nt}{\epsilon}\right)^{o(1)}, \ \mathcal{O}\left(n^{3-\frac{\alpha}{d}} t \operatorname{polylog}\left(\frac{nt}{\epsilon}\right)\right) + n^{2-\frac{\alpha}{d}} t\left(\frac{nt}{\epsilon}\right)^{o(1)}\right\}, & 0 < \alpha < d. \end{aligned} \tag{C10}$$

The generalization of the low-rank method in Sec. V is similar and straightforward, by using higher-dimensional cubes instead of 1D intervals to perform the recursive decomposition. Unlike the block-encoding method, here we do not have the issue of normalization factor; the circuit implementation is just a direct translation of classical computations. Moreover, for Coulomb interactions in higher spatial dimensions, one can show using the multipole expansion that the decomposed
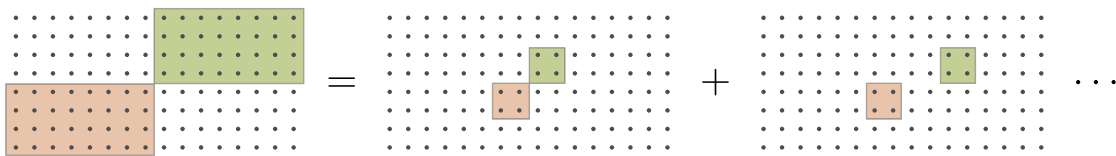

FIG. 6.   Illustration of an average-cost block encoding for power-law Hamiltonians in 2D.

terms can also be approximated with a matrix with rank logarithmic in the input parameters. The resulting implementation has a complexity similar to that of the 1D case:

$$
\begin{cases}
\rho n t \left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha \geq d, \\
\rho n^{2-\frac{\alpha}{d}} t \left(\frac{nt}{\epsilon}\right)^{o(1)}, & \alpha < d.
\end{cases}
\tag{C11}
$$

We refer the reader to Ref. [79] for a more detailed discussion of the decomposition.

2. κ-local Hamiltonians

We have focused on the simulation of two-local Hamiltonians in the main body of the paper. In this subsection, we discuss how to extend our approach to handle more general κ-local Hamiltonians for constant κ.

Specifically, the Hamiltonian takes the form $H = \sum_{1 \leq j_1 < j_2 < \cdots < j_\kappa \leq n} H_{j_1, j_2, \ldots, j_\kappa}$, where operator $H_{j_1, j_2, \ldots, j_\kappa}$ acts nontrivially only on qubits $j_1, \ldots, j_\kappa$. Similar to Eq. (48), we can rewrite the Hamiltonian using tensor products of Pauli operators and decompose the evolution accordingly using product formulas. This introduces a Trotter error no larger than the naive decomposition where all Hamiltonian terms are decomposed. Thus without loss of generality, we consider

$$
H = \sum_{1 \leq j_1 < j_2 < \cdots < j_\kappa \leq n} \beta_{j_1, j_2, \ldots, j_\kappa} P_{j_1}^{(\sigma_1)} P_{j_2}^{(\sigma_2)} \cdots P_{j_\kappa}^{(\sigma_\kappa)},
\tag{C12}
$$

where $P_{j_b}^{(\sigma_b)}$ is a Pauli operator on qubit j labeled by $\sigma_b = x, y, z$.

To describe the recursive decomposition, we introduce the abbreviation

$$
H_{[j,k]} := \sum_{j \leq u_1 < \cdots < u_\kappa \leq k} \beta_{u_1, \ldots, u_\kappa} P_{u_1}^{(\sigma_1)} \cdots P_{u_\kappa}^{(\sigma_\kappa)},
$$

$$
H_{[j_1,k_1]:[j_2,k_2]:\cdots:[j_\kappa,k_\kappa], \kappa_1:\kappa_2:\cdots:\kappa_\kappa} := \sum_{j_1 \leq u_1 < \cdots < u_{\kappa_1} \leq k_1} \sum_{j_2 \leq v_1 < \cdots < v_{\kappa_2} \leq k_2} \cdots \sum_{j_\kappa \leq w_1 < \cdots < w_{\kappa_\kappa} \leq k_\kappa}
$$

$$
\beta_{u_1, \ldots, u_{\kappa_1}, v_1, \ldots, v_{\kappa_2}, \ldots, w_1, \ldots, w_{\kappa_\kappa}} P_{u_1}^{(\sigma_1)} \cdots P_{w_{\kappa_\kappa}}^{(\sigma_\kappa)}
\tag{C13}
$$

to represent terms within a specific interval and across κ disjoint intervals of sites, where $j_1 \leq k_1 < j_2 \leq k_2 < \ldots$ and $\kappa_1, \ldots, \kappa_\kappa$ are non-negative integers summing to $\kappa_1 + \cdots + \kappa_\kappa = \kappa$. Unlike the two-local case, we now need additional parameters $\kappa_1, \ldots, \kappa_\kappa$ to record the number of indices within each interval. With that, the decomposition used in the block-encoding method in Sec. III should be modified to

$$
H_{[1,n]} = \sum_{\substack{\kappa_1 + \cdots + \kappa_\kappa = \kappa \\ \forall \kappa_b < \kappa}} H_{[1,n/\kappa]:[n/\kappa+1,2n/\kappa]:\cdots:[(\kappa-1)n/\kappa,n], \kappa_1:\kappa_2:\cdots:\kappa_\kappa} + \sum_{b=0}^{\kappa-1} H_{[b(n/\kappa)+1,(b+1)n/\kappa]}.
\tag{C14}
$$

We assume that κ is also a power of 2, but the general case can be handled by redefining the boundary terms. Now, terms in the second line can be unwrapped under the same recursion, whereas terms in the first line correspond to degenerate cases with locality parameter $< \kappa$ [there are at most $\binom{2\kappa-1}{\kappa-1} = \mathcal{O}(1)$ cases] that can be handled by an outer induction on κ. We may modify the decompositions in Secs. IV and V in a similar way. However, the Hamiltonian coefficients now have a more complicated tensor structure, so further studies are required to understand the complexity of implementing Trotter steps using such methods.

3. General fermionic Hamiltonians

Fermionic Hamiltonians that are geometrically local can be mapped to spin models using a locality-preserving encoding such as Ref. [128] and solved using methods from the previous subsection. However, this strategy does not work for more general fermionic models, as the number of recursively decomposed Hamiltonian terms scale exponentially with the locality parameter κ, which is no longer constant under fermionic encodings such as the Jordan-Wigner or Bravyi-Kitaev

encoding. We now discuss how to potentially apply our techniques to simulating general fermionic Hamiltonians. For presentational purpose, we examine only the one-body operator in one spatial dimension:

$$H = \sum_{j,k=1}^{n} \beta_{j,k} A_j^{\dagger} A_k, \tag{C15}$$

where $A_j^{\dagger}$ and A_k are creation and annihilation operators on fermionic modes j and k, respectively. The analysis of higher spatial dimensionality and many-body fermionic operators may proceed in a similar way as in Appendices C 1 and C 2.

We introduce the following abbreviation:

$$H_{[j,k]} := \sum_{j \le u < v \le k} \beta_{u,v} \left(A_u^{\dagger} A_v + A_v^{\dagger} A_u \right) \ (1 \le j < k \le n),$$

$$H_{[j,k]:[l,m]} := \sum_{\substack{j \le u \le k \\ l \le v \le m}} \beta_{u,v} \left(A_u^{\dagger} A_v + A_v^{\dagger} A_u \right) \ (1 \le j \le k < l \le m \le n). \tag{C16}$$

This is similar to that used in Secs. III–V, except we add the complex-conjugate terms to make the entire operator Hermitian. For the recursive decomposition with a reduced 1-norm, we have

$$H = \sum_{\ell=1}^{\log n-1} \sum_{b=0}^{2^{\ell-1}-1} H_{[2b(n/2^{\ell})+1,(2b+1)n/2^{\ell}]:[(2b+1)n/2^{\ell}+1,2(b+1)n/2^{\ell}]}. \tag{C17}$$

This decomposition shares similar features as that in Sec. III. In particular, the 1-norm of each summand is significantly smaller compared to the full Hamiltonian, suggesting that the block-encoding method can be advantageous.

We now explain how to realize an efficient block encoding. Without loss of generality, we choose $\ell = 1$, $b = 0$ and study the term $H_{[1,n/2]:[n/2+1,n]}$. This can be block encoded as follows. We define the Majorana operators

$$M_{u,0} := A_u^{\dagger} + A_u, \qquad M_{u,1} := i \left(A_u^{\dagger} - A_u \right), \tag{C18}$$

using which we rewrite

$$\begin{aligned}
H_{[1,n/2]:[n/2+1,n]} &= \sum_{\substack{1 \le u \le n/2 \\ n/2+1 \le v \le n}} \beta_{u,v} \left(A_u^{\dagger} A_v + A_v^{\dagger} A_u \right) \\
&= \sum_{\substack{1 \le u \le n/2 \\ n/2+1 \le v \le n}} \beta_{u,v} \left(\frac{M_{u,0} - iM_{u,1}}{2} \frac{M_{v,0} + iM_{v,1}}{2} + \frac{M_{v,0} - iM_{v,1}}{2} \frac{M_{u,0} + iM_{u,1}}{2} \right) \\
&= \sum_{\substack{1 \le u \le n/2 \\ n/2+1 \le v \le n}} i\beta_{u,v} \left(M_{u,0} M_{v,1} + M_{v,0} M_{u,1} \right).
\end{aligned} \tag{C19}$$

The preparation subroutine can now be implemented in a similar way as in Sec. III B. For the selection subroutine, we need to iterate over products of Majorana operators, which can be realized with cost $\mathcal{O}(n)$ (see, for example, Ref. [129, Sec. III.B]).

For the recursive decomposition described in Sec. V, we have

$$\begin{aligned}
H = \sum_{\ell=2}^{\eta} \sum_{b=0}^{2^{\ell-1}-2} &\Bigg(H_{[1+2b(n/2^{\ell}),(2b+1)n/2^{\ell}]:[1+(2b+2)n/2^{\ell},(2b+3)n/2^{\ell}]} + H_{[1+2b(n/2^{\ell}),(2b+1)n/2^{\ell}]:[1+(2b+3)n/2^{\ell},(2b+4)n/2^{\ell}]} \\
&+ H_{[1+(2b+1)n/2^{\ell},(2b+2)n/2^{\ell}]:[1+(2b+3)n/2^{\ell},(2b+4)n/2^{\ell}]} \Bigg) + \sum_{b=0}^{2^{\eta}-1} H_{[1+b(n/2^{\ell}),(b+1)n/2^{\ell}]} \\
&+ \sum_{b=0}^{2^{\eta}-2} H_{[1+b(n/2^{\ell}),(b+1)n/2^{\ell}]:[1+(b+1)n/2^{\ell},(b+2)n/2^{\ell}]}.
\end{aligned} \tag{C20}$$

Without loss of generality, we choose $\ell = 2$, $b = 0$ and study the term

$$H_{[1,n/4]:[n/2+1,3n/4]} = \sum_{\substack{1 \le u \le n/4 \\ n/2+1 \le v \le 3n/4}} \beta_{u,v} \left(A_u^\dagger A_v + A_v^\dagger A_u \right)$$

(C21)

with the coefficient matrix

$$\widetilde{\beta} = \begin{bmatrix} 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & 0 \\ \beta^\top & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(C22)

Note that although the Hamiltonian terms no long commute in the fermionic case, if $\mathrm{rank}(\beta) = \rho$, then $\mathrm{rank}(\widetilde{\beta}) = \rho$ as well, and we can still implement the exponential of $H_{[1,n/4]:[n/2+1,3n/4]}$ using fermionic Givens rotations with cost $\mathcal{O}(\rho n)$ [102, 2.1.P28] (see Ref. [130] for the circuit implementation). Of course, the circuit complexity will depend on the target Hamiltonian as well as the decomposition scheme, which become more complicated to analyze for general fermionic operators. We leave a detailed study of these generalizations as a subject for future work.

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[2] Y. Manin, Vychislimoe i nevychislimoe (computable and noncomputable), Moscow: Sov (1980).

[3] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Toward the first quantum simulation with quantum speedup, Proc. Nat. Acad. Sci. **115**, 9456 (2018).

[4] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[5] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).

[6] M. Motta and J. E. Rice, Emerging quantum computing algorithms for quantum chemistry, WIREs Comput. Mol. Sci. **12**, e1580 (2022).

[7] C. W. Bauer, et al., Quantum simulation for high energy physics (2022).

[8] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum Algorithm for Linear Systems of Equations, Phys. Rev. Lett. **103**, 150502 (2009).

[9] D. An and L. Lin, Quantum linear system solver based on time-optimal adiabatic quantum computing and quantum approximate optimization algorithm, ACM Trans. Quantum Comput. **3**, 1 (2022).

[10] L. Lin and Y. Tong, Near-optimal ground state preparation, Quantum **4**, 372 (2020).

[11] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, in Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (2019), p. 193.

[12] S. Chakraborty, A. Gilyén, and S. Jeffery, in 46th International Colloquium on Automata, Languages, and Programming, Vol. 132 (2019), p. 33:1.

[13] A. Anshu, A. W. Harrow, and M. Soleimanifar, Entanglement spread area law in gapped ground states, Nat. Phys. **18**, 1362 (2022).

[14] T. Kuwahara, Á. M. Alhambra, and A. Anshu, Improved Thermal Area Law and Quasilinear Time Algorithm for Quantum Gibbs States, Phys. Rev. X **11**, 011047 (2021).

[15] M. C. Tran, A. Y. Guo, Y. Su, J. R. Garrison, Z. Eldredge, M. Foss-Feig, A. M. Childs, and A. V. Gorshkov, Locality and Digital Quantum Simulation of Power-Law Interactions, Phys. Rev. X **9**, 031006 (2019).

[16] N. Abrahamsen, Y. Su, Y. Tong, and N. Wiebe, Entanglement area law for 1D gauge theories and bosonic systems (2022), ArXiv:2203.16012.

[17] H.-Y. Huang, Y. Tong, D. Fang, and Y. Su, Learning many-body Hamiltonians with Heisenberg-limited scaling (2022), ArXiv:2210.03030.

[18] S. Lloyd, Universal quantum simulators, Science **273**, 1073 (1996).

[19] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Efficient quantum algorithms for simulating sparse Hamiltonians, Commun. Math. Phys. **270**, 359 (2007).

[20] E. Campbell, Random Compiler for Fast Hamiltonian Simulation, Phys. Rev. Lett. **123**, 070503 (2019).

[21] 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).

[22] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, Quantum **3**, 163 (2019).

[23] G. H. Low, V. Kliuchnikov, and N. Wiebe, Well-conditioned multiproduct Hamiltonian simulation (2019), ArXiv:1907.11679.

[24] G. H. Low and I. L. Chuang, Hamiltonian simulation by uniform spectral amplification (2017), ArXiv:1707.05391.

[25] J. M. Martyn, Z. M. Rossi, A. K. Tan, and I. L. Chuang, Grand Unification of Quantum Algorithms, PRX Quantum **2**, 040203 (2021).

[26] G. H. Low, in Proceedings of the 51th ACM Symposium on Theory of Computing (ACM, 2019), p. 491.

[27] Y. Ge, J. Tura, and J. I. Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits, J. Math. Phys. **60**, 022202 (2019).

[28] D. Poulin and P. Wocjan, Preparing Ground States of Quantum Many-Body Systems on a Quantum Computer, Phys. Rev. Lett. **102**, 130503 (2009).

[29] P. Rall, Faster coherent quantum algorithms for phase, energy, and amplitude estimation, Quantum **5**, 566 (2021).

[30] L. Lin and Y. Tong, Heisenberg-Limited Ground-State Energy Estimation for Early Fault-Tolerant Quantum Computers, PRX Quantum **3**, 010318 (2022).

[31] Y. Dong, L. Lin, and Y. Tong, Ground-State Preparation and Energy Estimation on Early Fault-Tolerant Quantum Computers via Quantum Eigenvalue Transformation of Unitary Matrices, PRX Quantum **3**, 040305 (2022).

[32] J. Lee, D. W. Berry, C. Gidney, W. J. Huggins, J. R. McClean, N. Wiebe, and R. Babbush, Even More Efficient Quantum Computations of Chemistry Through Tensor Hypercontraction, PRX Quantum **2**, 030305 (2021).

[33] V. von Burg, G. H. Low, T. Häner, D. S. Steiger, M. Reiher, M. Roetteler, and M. Troyer, Quantum computing enhanced computational catalysis, Phys. Rev. Res. **3**, 033055 (2021).

[34] C.-F. Chen, H.-Y. Huang, R. Kueng, and J. A. Tropp, Concentration for Random Product Formulas, PRX Quantum **2**, 040305 (2021).

[35] K. Wan, M. Berta, and E. T. Campbell, Randomized Quantum Algorithm for Statistical Phase Estimation, Phys. Rev. Lett. **129**, 030503 (2022).

[36] Y. Atia and D. Aharonov, Fast-forwarding of Hamiltonians and exponentially precise measurements, Nat. Commun. **8**, 1572 (2017).

[37] S. Gu, R. D. Somma, and B. Şahinoğlu, Fast-forwarding quantum evolution, Quantum **5**, 577 (2021).

[38] E. Kökcü, T. Steckmann, Y. Wang, J. K. Freericks, E. F. Dumitrescu, and A. F. Kemper, Fixed Depth Hamiltonian Simulation via Cartan Decomposition, Phys. Rev. Lett. **129**, 070501 (2022).

[39] E. Kökcü, D. Camps, L. Bassman, J. K. Freericks, W. A. de Jong, R. Van Beeumen, and A. F. Kemper, Algebraic compression of quantum circuits for Hamiltonian evolution, Phys. Rev. A **105**, 032420 (2022).

[40] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer, The Trotter step size required for accurate quantum simulation of quantum chemistry, Quantum Inf. Comput. **15**, 361 (2015).

[41] M. Motta, E. Ye, J. R. McClean, Z. Li, A. J. Minnich, R. Babbush, and G. K.-L. Chan, Low rank representations for quantum simulation of electronic structure, npj Quantum Inf. **7**, 83 (2021).

[42] N. C. Rubin, J. Lee, and R. Babbush, Compressing many-body fermion operators under unitary constraints, J. Chem. Theory Comput. **18**, 1480 (2022).

[43] A. Anand, P. Schleich, S. Alperin-Lea, P. W. K. Jensen, S. Sim, M. Díaz-Tinoco, J. S. Kottmann, M. Degroote, A. F. Izmaylov, and A. Aspuru-Guzik, A quantum computing view on unitary coupled cluster theory, Chem. Soc. Rev. **51**, 1659 (2022).

[44] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter Error with Commutator Scaling, Phys. Rev. X **11**, 011020 (2021).

[45] S. Xu, L. Susskind, Y. Su, and B. Swingle, A sparse model of quantum holography (2020), ArXiv:2008.02303.

[46] L. Clinton, T. Cubitt, B. Flynn, F. M. Gambetta, J. Klassen, A. Montanaro, S. Piddock, R. A. Santos, and E. Sheridan, Towards near-term quantum simulation of materials (2022), ArXiv:2205.15256.

[47] R. Meister, S. C. Benjamin, and E. T. Campbell, Tailoring term truncations for electronic structure calculations using a linear combination of unitaries, Quantum **6**, 637 (2022).

[48] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, and R. Babbush, Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization, Quantum **3**, 208 (2019).

[49] A. V. Ivanov, C. Sünderhauf, N. Holzmann, T. Ellaby, R. N. Kerber, G. Jones, and J. Camps, Quantum computation for periodic solids in second quantization (2022), ArXiv:2210.02403.

[50] W. Hackbusch, A sparse matrix arithmetic based on H-matrices. Part I: Introduction to H-matrices, Computing **62**, 89 (1999).

[51] L. Grasedyck and W. Hackbusch, Construction and arithmetics of H-matrices, Computing **70**, 295 (2003).

[52] Q. T. Nguyen, B. T. Kiani, and S. Lloyd, Block-encoding dense and full-rank kernels using hierarchical matrices: applications in quantum numerical linear algebra, Quantum **6**, 876 (2022).

[53] Roughly speaking, they are implementing operators in first quantization. One may attempt to use a similar approach on the Coulomb interaction to improve the electronic structure simulation. We are not aware of a simple realization of this idea. Existing study of this problem uses a computation model stronger than the circuit model [131].

[54] I. D. Kivlichan, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, W. Sun, Z. Jiang, N. Rubin, A. Fowler, A. Aspuru-Guzik, H. Neven, and R. Babbush, Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via Trotterization, Quantum **4**, 296 (2020).

[55] A. F. Shaw, P. Lougovski, J. R. Stryker, and N. Wiebe, Quantum algorithms for simulating the lattice Schwinger model, Quantum **4**, 306 (2020).

[56] 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 **484**, 489 (2012).

[57] 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).

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

[59] 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. Photon. **9**, 326 (2015).

[60] 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 **501**, 521 (2013).

[61] T. Koffel, M. Lewenstein, and L. Tagliacozzo, Entanglement Entropy for the Long-Range Ising Chain in a Transverse Field, Phys. Rev. Lett. **109**, 267203 (2012).

[62] H. P. Büchler, E. Demler, M. Lukin, A. Micheli, N. Prokof'ev, G. Pupillo, and P. Zoller, Strongly Correlated 2D Quantum Phases with Cold Polar Molecules: Controlling the Shape of the Interaction Potential, Phys. Rev. Lett. **98**, 060404 (2007).

[63] A. V. Gorshkov, P. Rabl, G. Pupillo, A. Micheli, P. Zoller, M. D. Lukin, and H. P. Büchler, Suppression of

Inelastic Collisions between Polar Molecules with a Repulsive Shield, Phys. Rev. Lett. **101**, 073201 (2008).

[64] M. Dalmonte, G. Pupillo, and P. Zoller, One-Dimensional Quantum Liquids with Power-Law Interactions: The Luttinger Staircase, Phys. Rev. Lett. **105**, 140401 (2010).

[65] A. Micheli, G. Pupillo, H. P. Büchler, and P. Zoller, Cold polar molecules in two-dimensional traps: Tailoring interactions with external fields for novel quantum phases, Phys. Rev. A **76**, 043604 (2007).

[66] X. Deng, G. Masella, G. Pupillo, and L. Santos, Universal Algebraic Growth of Entanglement Entropy in Many-Body Localized Systems with Power-Law Interactions, Phys. Rev. Lett. **125**, 010401 (2020).

[67] The fractional-query algorithm also implements a compressed version of Trotter steps using more advanced simulation techniques [132]. That algorithm is superseded by the block-encoding method [133], so it suffers from the same normalization-factor issue pointed out here.

[68] G. H. Low and N. Wiebe, Hamiltonian simulation in the interaction picture (2018), ArXiv:1805.00675.

[69] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT press, Cambridge, Massachusetts, 2022).

[70] T. Roughgarden, *Algorithms Illuminated* (Soundlikeyourself publishing, New York, NY, 2017).

[71] R. E. Neapolitan, *Foundations of Algorithms* (Jones & Bartlett Learning, Burlington, MA, 2015).

[72] W. Kuszmaul and C. E. Leiserson, in *Symposium on Simplicity in Algorithms (SOSA)* (SIAM, 2021), p. 133.

[73] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Polynomial-time quantum algorithm for the simulation of chemical dynamics, Proc. Nat. Acad. Sci. **105**, 18681 (2008).

[74] N. C. Jones, J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. V. Meter, A. Aspuru-Guzik, and Y. Yamamoto, Faster quantum chemistry simulation on fault-tolerant quantum computers, New J. Phys. **14**, 115023 (2012).

[75] I. D. Kivlichan, N. Wiebe, R. Babbush, and A. Aspuru-Guzik, Bounding the costs of quantum simulation of many-body physics in real space, J. Phys. A: Math. Theor. **50**, 305301 (2017).

[76] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, Low-Depth Quantum Simulation of Materials, Phys. Rev. X **8**, 011044 (2018).

[77] Y. Su, D. W. Berry, N. Wiebe, N. Rubin, and R. Babbush, Fault-Tolerant Quantum Simulations of Chemistry in First Quantization, PRX Quantum **2**, 040332 (2021).

[78] H. H. S. Chan, R. Meister, T. Jones, D. P. Tew, and S. C. Benjamin, Grid-based methods for chemistry simulations on a quantum computer (2022), ArXiv:2202.05864.

[79] L. Greengard and V. Rokhlin, A fast algorithm for particle simulations, J. Comput. Phys. **73**, 325 (1987).

[80] S. Hadfield, On the representation of boolean and real functions as Hamiltonians for quantum computing, ACM Trans. Quantum Comput. **2**, 1 (2021).

[81] A. Bocharov, M. Roetteler, and K. M. Svore, Efficient Synthesis of Universal Repeat-Until-Success Quantum Circuits, Phys. Rev. Lett. **114**, 080502 (2015).

[82] A. Y. Kitaev, Quantum computations: Algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[83] S. S. Bullock and I. L. Markov, Asymptotically optimal circuits for arbitrary n-qubit diagonal computations, Quantum Inf. Comput. **4**, 27 (2004).

[84] J. Welch, D. Greenbaum, S. Mostame, and A. Aspuru-Guzik, Efficient quantum circuits for diagonal unitaries without ancillas, New J. Phys. **16**, 033040 (2014).

[85] Y. Jia and M. M. Wolf, Hay from the haystack: explicit examples of exponential quantum circuit complexity (2022), ArXiv:2205.06977.

[86] E. Knill, Approximation by quantum circuits (1995), ArXiv:quant-ph/9508006.

[87] Y. Su, H.-Y. Huang, and E. T. Campbell, Nearly tight Trotterization of interacting electrons, Quantum **5**, 495 (2021).

[88] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low, in *Proceedings of the 59th IEEE Symposium on Foundations of Computer Science* (2018), p. 350.

[89] A. M. Childs and Y. Su, Nearly Optimal Lattice Simulation by Product Formulas, Phys. Rev. Lett. **123**, 050503 (2019).

[90] S. Blanes and F. Casas, *A Concise Introduction to Geometric Numerical Integration* (Chapman and Hall/CRC, Boca Raton, FL, 2016).

[91] C.-F. Chen and F. G. Brandão, Concentration for Trotter error (2021), ArXiv:2111.05324.

[92] D. Layden, First-Order Trotter Error from a Second-Order Perspective, Phys. Rev. Lett. **128**, 210501 (2022).

[93] D. An, D. Fang, and L. Lin, Time-dependent unbounded Hamiltonian simulation with vector norm scaling, Quantum **5**, 459 (2021).

[94] D. Burgarth, N. Galke, A. Hahn, and L. van Luijk, State-dependent Trotter limits and their approximations (2022), ArXiv:2209.14787.

[95] L. Kocia, F. A. Calderon-Vargas, M. D. Grace, A. B. Magann, J. B. Larsen, A. D. Baczewski, and M. Sarovar, Digital adiabatic state preparation error scales better than you might expect (2022), ArXiv:2209.06242.

[96] Q. Zhao, Y. Zhou, A. F. Shaw, T. Li, and A. M. Childs, Hamiltonian simulation with random inputs (2021), ArXiv:2111.04773.

[97] B. Şahinoğlu and R. D. Somma, Hamiltonian simulation in the low-energy subspace, npj Quantum Inf. **7**, 119 (2021).

[98] T. Hatomura, State-dependent error bound for digital quantum simulation of driven systems, Phys. Rev. A **105**, L050601 (2022).

[99] M. Heyl, P. Hauke, and P. Zoller, Quantum localization bounds Trotter errors in digital quantum simulation, Sci. Adv. **5**, 8342 (eaau2019).

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

[101] S. Gharibian and F. Le Gall, Dequantizing the quantum singular value transformation: in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing* (2022), p. 19.

[102] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge university press, New York, NY, 2012).

[103] V. Shende, S. Bullock, and I. Markov, Synthesis of quantum-logic circuits, IEEE Trans. Comput.-Aided Design Integrated Circuits Syst. **25**, 1000 (2006).

[104] D. Poulin, A. Y. Kitaev, D. Steiger, M. Hastings, and M. Troyer, Quantum Algorithm for Spectral Measurement with Lower Gate Count, Phys. Rev. Lett. **121**, 010501 (2018).

[105] D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush, Improved techniques for preparing eigenstates of fermionic Hamiltonians, npj Quantum Inf. **4**, 22 (2018).

[106] This simplification follows from the following spectral property of V: there exists an orthonormal basis under which V is exhibited as a direct sum of 1-by-1 and 2-by-2 matrices, whose eigenvalues correspond to eigenvalues of the target Hamiltonian (which is enough for phase estimation). We can then visualize V as a collection of single-qubit unitaries acting on orthogonal subspaces; hence the name "qubitization".

[107] G. H. Low, V. Kliuchnikov, and L. Schaeffer, Trading T-gates for dirty qubits in state preparation and unitary synthesis (2018), ArXiv:1812.00954.

[108] M. Steudtner and S. Wehner, Estimating exact energies in quantum simulation without Toffoli gates, Phys. Rev. A **101**, 052329 (2020).

[109] Q. Zhao and X. Yuan, Exploiting anticommutation in Hamiltonian simulation, Quantum **5**, 534 (2021).

[110] I. Loaiza, A. M. Khah, N. Wiebe, and A. F. Izmaylov, Reducing molecular electronic Hamiltonian simulation cost for linear combination of unitaries approaches (2022), ArXiv:2208.08272.

[111] M. Giaquinta and G. Modica, *Mathematical Analysis: An Introduction to Functions of Several Variables* (Springer Science & Business Media, New York, NY, 2010).

[112] M. Giaquinta and G. Modica, *Mathematical Analysis: Foundations and Advanced Techniques for Functions of Several Variables* (Springer Science & Business Media, New York, NY, 2011).

[113] Note that H is indeed Hermitian as X_j and Y_k commute for $j < k$.

[114] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, Nearly Optimal Measurement Scheduling for Partial Tomography of Quantum States, Phys. Rev. X **10**, 031064 (2020).

[115] Alternatively, one may treat O_β as a black box and consider the query complexity.

[116] R. Babbush, D. W. Berry, J. R. McClean, and H. Neven, Quantum simulation of chemistry with sublinear scaling in basis size, npj Quantum Inf. **5**, 92 (2019).

[117] Note that we are using the same qubitization algorithm to simulate all the Hamiltonian terms, so there is a term-independent constant prefactor for all the gate complexities. This is why we have the summation rule $\sum_{\gamma=1}^{\Gamma} c_\gamma \mathcal{O}(\beta_\gamma t + \log(\Gamma/\epsilon)) = \mathcal{O}\left(\left(\sum_{\gamma=1}^{\Gamma} \beta_\gamma c_\gamma\right) t + \sum_{\gamma=1}^{\Gamma} c_\gamma \log(\Gamma/\epsilon)\right)$. This rule is used without declaration in the remainder of the paper.

[118] As before, one may instead treat O_β as a black box and consider the query complexity.

[119] Note that this logarithmic factor will be absorbed by $(nt/\epsilon)^{o(1)}$ in the final gate complexity.

[120] The scaling we report here corresponds to the (classical) fast-Fourier-transform implementation [77, Eq. (K2)] (whose analysis in turn follows from the bounds [76, Eqs. (F11)–(F13)] in second quantization). The finite-difference scheme of Ref. [76, Eq. (A7)] seems to have a different scaling for the kinetic term, which should be further studied by future work.

[121] L. Lin and Y. Tong, Low-rank representation of tensor network operators with long-range pairwise interactions, SIAM J. Sci. Comput. **43**, A164 (2021).

[122] C. R. Myers, C. J. Umrigar, J. P. Sethna, and J. D. Morgan, Fock's expansion, Kato's cusp conditions, and the exponential ansatz, Phys. Rev. A **44**, 5537 (1991).

[123] L. A. Martínez-Martínez, T.-C. Yen, and A. F. Izmaylov, Assessment of various Hamiltonian partitionings for the electronic structure problem on a quantum computer using the Trotter approximation (2022), ArXiv:2210.10189.

[124] A. Rajput, A. Roggero, and N. Wiebe, Hybridized methods for quantum simulation in the interaction picture, Quantum **6**, 780 (2022).

[125] Y. Tong, V. V. Albert, J. R. McClean, J. Preskill, and Y. Su, Provably accurate simulation of gauge theories and bosonic systems, Quantum **6**, 816 (2022).

[126] B. O'Gorman, W. J. Huggins, E. G. Rieffel, and K. B. Whaley, Generalized swap networks for near-term quantum computing (2019), ArXiv:1905.05118.

[127] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, New York, NY, 2013).

[128] F. Verstraete and J. I. Cirac, Mapping local Hamiltonians of fermions to local Hamiltonians of spins, J. Stat. Mech.: Theory Exp. **2005**, P09012 (2005).

[129] R. Babbush, D. W. Berry, and H. Neven, Quantum simulation of the Sachdev-Ye-Kitaev model by asymmetric qubitization, Phys. Rev. A **99**, 040301 (2019).

[130] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, Phys. Rev. Lett. **120**, 110501 (2018).

[131] A. M. Childs, J. Leng, T. Li, J.-P. Liu, and C. Zhang, Quantum simulation of real-space dynamics (2022), ArXiv:2203.17006.

[132] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing* (2014), p. 283, ArXiv:1312.1414.

[133] R. Kothari, *Efficient algorithms in quantum query complexity*, Ph.D. thesis, University of Waterloo (2014).