Theory of Trotter Error with Commutator Scaling

Andrew M. Childs,[1,2,3] Yuan Su,[1,2,3] Minh C. Tran,[3,4] Nathan Wiebe,[5,6,7] and Shuchen Zhu[8]

[1]Department of Computer Science, University of Maryland, College Park, Maryland 20742, USA
[2]Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742, USA
[3]Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, Maryland 20742, USA
[4]Joint Quantum Institute, University of Maryland, College Park, Maryland 20742, USA
[5]Department of Physics, University of Washington, Seattle, Washington 98195, USA
[6]Pacific Northwest National Laboratory, Richland, Washington 99354, USA
[7]Google Inc., Venice, California 90291, USA
[8]Department of Computer Science, Georgetown University, Washington, DC 20057, USA

(Received 21 January 2020; revised 7 July 2020; accepted 4 November 2020; published 1 February 2021)

The Lie-Trotter formula, together with its higher-order generalizations, provides a direct approach to decomposing the exponential of a sum of operators. Despite significant effort, the error scaling of such product formulas remains poorly understood. We develop a theory of Trotter error that overcomes the limitations of prior approaches based on truncating the Baker-Campbell-Hausdorff expansion. Our analysis directly exploits the commutativity of operator summands, producing tighter error bounds for both real- and imaginary-time evolutions. Whereas previous work achieves similar goals for systems with geometric locality or Lie-algebraic structure, our approach holds, in general. We give a host of improved algorithms for digital quantum simulation and quantum Monte Carlo methods, including simulations of second-quantized plane-wave electronic structure, k-local Hamiltonians, rapidly decaying power-law interactions, clustered Hamiltonians, the transverse field Ising model, and quantum ferromagnets, nearly matching or even outperforming the best previous results. We obtain further speedups using the fact that product formulas can preserve the locality of the simulated system. Specifically, we show that local observables can be simulated with complexity independent of the system size for power-law interacting systems, which implies a Lieb-Robinson bound as a by-product. Our analysis reproduces known tight bounds for first- and second-order formulas. Our higher-order bound overestimates the complexity of simulating a one-dimensional Heisenberg model with an even-odd ordering of terms by only a factor of 5, and it is close to tight for power-law interactions and other orderings of terms. This result suggests that our theory can accurately characterize Trotter error in terms of both asymptotic scaling and constant prefactor.

DOI: 10.1103/PhysRevX.11.011020 Subject Areas: Quantum Information

I. INTRODUCTION

Product formulas provide a convenient approach to decomposing the evolution of a sum of operators. The Lie product formula was introduced in the study of Lie groups in the late 1800s; later developments considered more general operators and higher-order approximations. Originally studied in the context of pure mathematics, product formulas have found numerous applications in other areas, such as applied mathematics (under the name "splitting method" or "symplectic integrators"), physics (under the name "Trotterization"), and theoretical computer science.

This paper considers the application of product formulas to simulating quantum systems. It has been known for over two decades that these formulas are useful for digital quantum simulation and quantum Monte Carlo methods. However, their error scaling is poorly understood, and existing bounds can be several orders of magnitude larger than those observed in practice, even for simulating relatively small systems.

We develop a theory of Trotter error that directly exploits the commutativity of operator summands to give tighter bounds. Whereas previous work achieves similar goals for systems with geometric locality or Lie-algebraic structure, our theory has no such restrictions. We present a host of examples in which product formulas can nearly match or even outperform state-of-the-art simulation results. We accompany our analysis with numerical calculation, which suggests that the bounds also have nearly tight constant prefactors.

We hope this work will motivate further studies of the product-formula approach, which has been deemphasized

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

in recent years in favor of more advanced simulation algorithms that are easier to analyze but harder to implement. Indeed, despite the sophistication of these "post-Trotter methods" and their optimality in certain general models, our work shows that they can be provably outperformed by product formulas for simulating many quantum systems.

A. Simulating quantum systems by product formulas

Simulating the dynamics of quantum systems is one of the most promising applications of digital quantum computers. Classical computers apparently require exponential time to simulate typical quantum dynamics. This intractability led Feynman [1] and others to propose the idea of quantum computers. In 1996, Lloyd gave the first explicit quantum algorithm for simulating k-local Hamiltonians [2]. Subsequent work considered the broader class of sparse Hamiltonians [3–8] and developed techniques for simulating particular physical systems [9–15], with potential applications to developing new pharmaceuticals, catalysts, and materials. The study of quantum simulation has also inspired the design of various quantum algorithms for other problems [16–20].

Lloyd's approach to quantum simulation is based on product formulas. Specifically, let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a k-local Hamiltonian [i.e., each H_γ acts nontrivially on $k = \mathcal{O}(1)$ qubits]. Assuming H is time independent, evolution under H for time t is described by the unitary operation e^{-itH}. When t is small, this evolution can be well approximated by the Lie-Trotter formula $\mathcal{S}_1(t) = e^{-itH_\Gamma} \cdots e^{-itH_1}$, where each e^{-itH_γ} can be efficiently implemented on a quantum computer. To simulate for a longer time, we may divide the evolution into r Trotter steps and simulate each step with Trotter error of at most ϵ/r. We choose the Trotter number r to be sufficiently large so that the entire simulation achieves an error of at most ϵ. The Lie-Trotter formula only provides a first-order approximation to the evolution, but higher-order approximations are also known from the work of Suzuki and others [21,22]. While many previous works focused on the performance of specific formulas, the theory we develop holds for any formula; we use the term "product formula" to emphasize this generality. A quantum simulation algorithm using product formulas does not require ancilla qubits, making this approach advantageous for near-term experimental demonstration.

Recent studies have provided alternative simulation algorithms beyond the product-formula approach (sometimes called "post-Trotter methods"). Some of these algorithms have logarithmic dependence on the allowed error [5–7,23–25], an exponential improvement over product formulas. However, this dependence generally does not lead to an exponential reduction in time complexity for practical applications of quantum simulation. In practice, the simulation accuracy is often chosen to be constant. Then, the error dependence only enters as a constant prefactor, which may not significantly affect the overall gate complexity. The reduction in complexity is more significant when quantum simulation is used as a subroutine in another quantum algorithm (such as phase estimation) since this may require high-precision simulation to ensure reliable behavior. However, this logarithmic error dependence typically replaces a factor that scales polynomially with time or the system size by another that scales logarithmically, giving only a polynomial reduction in the complexity. Furthermore, the constant-factor overhead and extra space requirements of post-Trotter methods may make them uncompetitive with the product-formula approach in practice.

Product formulas and their generalizations [26–29] can perform significantly better when the operator summands commute or nearly commute—a unique feature that does not seem to hold for other quantum simulation algorithms [5–7,23–25,30]. This effect has been observed numerically in previous studies of quantum simulations of condensed matter systems [31] and quantum chemistry [32–34]. An intuitive explanation of this phenomenon comes from truncating the Baker-Campbell-Hausdorff (BCH) expansion. However, the intuition that the lowest-order terms of the BCH expansion are dominant is surprisingly difficult to justify (and sometimes is not even valid [9,35]). Thus, previous work established loose Trotter error bounds, sometimes suggesting poor performance. Our results rigorously demonstrate that for many systems, such arguments do not accurately reflect the true performance of product formulas.

Product-formula decompositions directly translate terms of the Hamiltonian into elementary simulation steps, making them well suited to preserve certain properties such as the locality of the simulated system. We show that this property can be used to further reduce the simulation cost when the goal is to simulate local observables as opposed to the full dynamics [36,37].

Besides digital quantum simulation, product formulas can also be applied to quantum Monte Carlo methods, in which the goal is to classically compute certain properties of the Hamiltonian, such as the partition function, the free energy, or the ground energy. Our results can also be applied to improve the efficiency of previous applications of quantum Monte Carlo methods for systems such as the transverse field Ising model [38] and quantum ferromagnets [39].

B. Previous analyses of Trotter error

We now briefly summarize prior approaches to analyzing Trotter error for simulating quantum systems, and we discuss their limitations.

The original work of Lloyd [2] analyzes product formulas by truncating the Taylor expansion (or the BCH expansion). Recall that the Lie-Trotter formula $\mathcal{S}_1(t)$

provides a first-order approximation to the evolution, so $\mathcal{S}_1(t) = e^{-itH} + O(t^2)$. To better analyze the Trotter error, Lloyd dropped all higher-order terms in the Taylor expansion and focused only on the terms of lowest order, t^2. This approach is intuitive and has been employed by subsequent works to give a rough estimation of Trotter error. The drawback of this analysis is that it implicitly assumes that the high-order terms are dominated by the lowest-order term. However, this case does not necessarily hold for many systems, such as nearest-neighbor lattice Hamiltonians [35] and chemical Hamiltonians [9], when the time step t is fixed.

This issue was addressed in the seminal work of Berry, Ahokas, Cleve, and Sanders by using a tail bound of the Taylor expansion [4]. This approach gave, for the first time, a concrete bound on the Trotter error for high-order Suzuki formulas. For a Hamiltonian $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ containing Γ summands, their bound scales with $\Gamma \max_\gamma \|H_\gamma\|$, although it is not hard to improve this [26] to $\sum_{\gamma=1}^{\Gamma} \|H_\gamma\|$ [28,40]. Regardless of which scaling to use, this worst-case analysis does not exploit the commutativity of Hamiltonian summands, and the rseulting complexity is worse than many post-Trotter methods.

Error bounds that exploit the commutativity of summands are known for low-order formulas, such as the Lie-Trotter formula [40,41] and the second-order Suzuki formula [9,40,42,43]. These bounds are tight in the sense that they match the lowest-order term of the BCH expansion up to an application of the triangle inequality. However, it is unclear whether they can be generalized, say, to the fourth- or the sixth-order case, which are still reasonably simple and can provide a significant advantage in practice [31].

Instead, previous works made compromises to obtain improved analyses of higher-order formulas. Somma gave an improved bound by representing the Trotter error as an infinite series of nested commutators [44]. This approach is advantageous when the simulated system has an underlying Lie-algebraic structure with small structure factors, such as for a quantum harmonic oscillator and certain nonquadratic potentials. However, this approach reduces to the worst-case analysis of Berry, Ahokas, Cleve, and Sanders for other systems.

An alternative approach of Thalhammer represented the error of a pth-order product formula using commutators of order up to q for $q \geq p$ [45], with the $(q+1)$st-order remainder further bounded by some tail bound. This analysis was bottlenecked by the use of a tail bound. The special case where $q = p + 1$ was studied in Ref. [31], and the result was applied to estimate the quantum resource for simulating a one-dimensional Heisenberg model, which only offers a modest improvement over the worst-case analysis.

In recent work [35], Childs and Su gave a Trotter error bound in which only the lowest-order error appears, avoiding manipulation of infinite series or use of tail bounds. As an immediate application, they showed that product formulas can nearly optimally simulate lattice systems with geometrically local interactions, justifying an earlier claim of Jordan, Lee, and Preskill [13] in the context of simulating quantum field theory.

Their improvement is based on a representation of the Trotter error introduced by Descombes and Thalhammer [42], which streamlines the previous analysis [45]. In this approach, the Trotter error is represented using commutators nested with conjugations of matrix exponentials. For Hamiltonians with nearest-neighbor interactions, Ref. [35] gave an argument based on locality to cancel the majority of the Trotter error. However, this approach reduces to the worst-case scenario for systems lacking geometric locality. In contrast, our representation of the Trotter error does not have this restriction and results in speedups for simulating even strongly long-range interacting systems (see Table I).

For other related studies of Trotter error in the context of numerical analysis, we refer the reader to Refs. [45–50] and the references therein.

C. Trotter error with commutator scaling

We give a new bound on the Trotter error that depends on nested commutators of the operator summands. This bound is formally stated in Sec. III D and previewed here.

Theorem: Trotter error with commutator scaling. Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator consisting of Γ summands, and let $t \geq 0$. Let $\mathcal{S}(t)$ be a pth-order Υ-stage product formula as in Sec. II C. Define $\tilde{\alpha}_{\mathrm{comm}} = \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\|$, where $\|\cdot\|$ is the spectral norm. Then, the additive error $\mathcal{A}(t)$ and the multiplicative error $\mathcal{M}(t)$, defined, respectively, by $\mathcal{S}(t) = e^{tH} + \mathcal{A}(t)$ and $\mathcal{S}(t) = e^{tH}(I + \mathcal{M}(t))$, can be asymptotically bounded sa

$$\|\mathcal{A}(t)\|, \|\mathcal{M}(t)\| = \mathcal{O}(\tilde{\alpha}_{\mathrm{comm}} t^{p+1} e^{4t\Upsilon \sum_{\gamma=1}^{\Gamma} \|H_\gamma\|}). \quad (1)$$

Furthermore, if the H_γ are anti-Hermitian, corresponding to physical Hamiltonians, we have

$$\|\mathcal{A}(t)\|, \|\mathcal{M}(t)\| = \mathcal{O}(\tilde{\alpha}_{\mathrm{comm}} t^{p+1}). \quad (2)$$

We emphasize that this theorem does not follow from truncating the BCH series. Although the $(p+1)$st-order term of the BCH series is also a linear combination of nested commutators similar to $\tilde{\alpha}_{\mathrm{comm}}$, such a term can be dominated by a higher-order term when t is fixed, as is the case for nearest-neighbor lattice systems (see Supplementary Sec. I of Ref. [35]) and quantum chemistry (see Appendix B of Ref. [9]). Truncating the BCH series ignores significant, potentially dominant error contributions and thus does not accurately characterize the Trotter error.

TABLE I. Comparison of our results and the best previous results for simulating quantum dynamics, simulating local observables, and quantum Monte Carlo simulation.

Application	System	Best previous result	New result
Simulating quantum dynamics	Electronic structure	$\tilde{\mathcal{O}}(n^2t)$ (interaction picture)	$n^{2+o(1)}t^{1+o(1)}$
	k-local Hamiltonians	$\tilde{\mathcal{O}}(n^k\|H\|_1t)$ (qubitization)	$n^k\|\|\|H\|\|\|_1\|H\|_1^{o(1)}t^{1+o(1)}$
	$1/x^\alpha$ $(\alpha < d)$	$\tilde{O}(n^{4-\alpha/d}t)$ (qubitization)	$n^{3-\alpha/d+o(1)}t^{1+o(1)}$
	$1/x^\alpha$ $(d \le \alpha \le 2d)$	$\tilde{O}(n^3t)$ (qubitization)	$n^{2+o(1)}t^{1+o(1)}$
	$1/x^\alpha$ $(\alpha > 2d)$	$\tilde{\mathcal{O}}((nt)^{1+2d/(\alpha-d)})$ (Lieb-Robinson bound)	$(nt)^{1+d/(\alpha-d)+o(1)}$
	Clustered Hamiltonians	$2^{\mathcal{O}(h_B^2t^2cc(g)/\epsilon)}$	$2^{\mathcal{O}(h_B^{o(1)}t^{1+o(1)}cc(g)/\epsilon^{o(1)})}$
Simulating local observables	$1/x^\alpha(\alpha > 2d)$	…	$t^{[1+d(\alpha-d/\alpha-2d)][1+(d/\alpha-d)]+o(1)}$
Monte Carlo simulation	Transverse field Ising model	$\tilde{\mathcal{O}}(n^{59}j^{21}\epsilon^{-9})$	$\tilde{\mathcal{O}}(n^{45}j^{14}\epsilon^{-2}+n^{38}j^{21}\epsilon^{-9})$
	Quantum ferromagnets	$\tilde{\mathcal{O}}(n^{115}(1+\beta^{46})/\epsilon^{25})$	$\tilde{\mathcal{O}}(n^{92}(1+\beta^{46})/\epsilon^{25})$

The above expression for our asymptotic error bound is succinct and easy to evaluate. In Sec. IV, we compute $\tilde{\alpha}_{\mathrm{comm}}$ for various examples, including the second-quantized electronic-structure Hamiltonians, k-local Hamiltonians, rapidly decaying power-law interactions, and clustered Hamiltonians. We further study the tightness of the prefactor of our bound in Sec. V and give a numerical implementation for one-dimensional Heisenberg models with either nearest-neighbor interactions or power-law interactions.

Although the definition of a specific product formula depends on the ordering of operator summands, our asymptotic bound does not. As an immediate consequence, the asymptotic speedups we obtain in Sec. IV A hold irrespective of how we order the operator summands in the simulation. For the special case of nearest-neighbor lattice models, this answers a previous question of Ref. [35] regarding the "ordering robustness" of higher-order formulas. However, the ordering becomes important if our goal is to simulate local observables or to get error bounds with tight constant prefactors, as we further discuss in Secs. IV B and Sec. V, respectively.

As mentioned in Sec. I B, prior Trotter error analyses typically produce loose bounds and are only effective in special cases. Our approach overcomes those limitations in the following respects:

(i) Our bound only contains a finite number of error terms, in contrast to the bound in Ref. [44].
(ii) Our bound involves pure nested commutators without conjugations of matrix exponentials or using tail bounds, overcoming the drawbacks of Refs. [35,45,47].
(iii) Our bound reduces to the worst-case analysis of Ref. [4] by further bounding terms with the triangle inequality.
(iv) For Hamiltonians with two summands, our bound encompasses the tight analyses [9,40–43] of the Lie-Trotter formula and the second-order Suzuki formula as special cases.

D. Overview of results

The commutator scaling of the Trotter error uncovers a host of examples where product formulas can nearly match or even outperform the state-of-the-art results in digital quantum simulation. These examples include (i) a simulation of second-quantized plane-wave electronic structure with n spin orbitals for time t with gate complexity $n^{2+o(1)}t^{1+o(1)}$, whereas the state-of-the-art approach performs simulation in the interaction picture [25] with cost $\tilde{\mathcal{O}}(n^2t)$ and likely large overhead; (ii) a simulation of n-qubit k-local Hamiltonians H with complexity $n^k|||H|||_1\|H\|_1^{o(1)}t^{1+o(1)}$ that almost scales with the induced 1-norm $|||H|||_1$ [51], implying an improved simulation of d-dimensional power-law interactions that decay with distance x as $1/x^\alpha$ for $\alpha \le 2d$, whereas the fastest previous approach uses the qubitization algorithm [24] with cost $\tilde{\mathcal{O}}(n^k\|H\|_1t)$; (iii) a simulation of d-dimensional power-law interactions $1/x^\alpha$ (for fixed $\alpha > 2d$) with gate complexity $(nt)^{1+d/(\alpha-d)+o(1)}$, whereas the best previous algorithm decomposes the evolution based on Lieb-Robinson bounds [52] with cost $\tilde{\mathcal{O}}((nt)^{1+2d/(\alpha-d)})$; and (iv) a hybrid simulation of clustered Hamiltonians of interaction strength h_B and contraction complexity $cc(g)$ with runtime $2^{\mathcal{O}(h_B^{o(1)}t^{1+o(1)}cc(g)/\epsilon^{o(1)})}$, improving the previous result of $2^{\mathcal{O}(h_B^2t^2cc(g)/\epsilon)}$ [53]. We discuss these examples in more detail in Sec. IV A.

We show in Sec. IV B that these gate complexities can be further improved when the goal is to simulate local observables instead of the full dynamics. We illustrate this for d-dimensional lattice systems with $1/x^\alpha$ interactions ($\alpha > 2d$). Lieb-Robinson bounds for power-law interactions [52] suggest that the evolution of a local observable is mostly confined inside a light cone induced by the interactions. Simulating such an evolution by simulating the dynamics of the entire system appears redundant, especially when the system size is large. We realize this intuition and show, without using Lieb-Robinson bounds,

that the gate count for simulating the evolution of a local observable scales as $t^{[1+d(\alpha-d)/(\alpha-2d)][1+d/(\alpha-d)]+o(1)}$, which is independent of the system size n and smaller than simulating the dynamics of the entire system when $n = \Omega(t^{d(\alpha-d)/(\alpha-2d)})$. The scaling also reduces to $t^{d+1+o(1)}$—proportional to the space-time volume inside a linear light cone—in the limit $\alpha \to \infty$, which corresponds to nearest-neighbor interactions.

Our bound can also be applied to improve the performance of quantum Monte Carlo simulation. In this case, we are limited to the use of the second-order Suzuki formula, and because of imaginary-time evolution, the Trotter number scales at least linearly with the system size. Nevertheless, we are able to improve several existing classical simulations using our bound, without modifying the original algorithms. This improvement includes (i) a simulation of an n-qubit transverse field Ising model with maximum interaction strength j and precision ϵ with runtime $\tilde{\mathcal{O}}(n^{45}j^{14}\epsilon^{-2} + n^{38}j^{21}\epsilon^{-9})$, tightening the previous result of $\tilde{\mathcal{O}}(n^{59}j^{21}\epsilon^{-9})$ [38]; and (ii) a simulation of ferromagnetic quantum spin systems for (imaginary) time β and accuracy ϵ with runtime $\tilde{\mathcal{O}}(n^{92}(1+\beta^{46})/\epsilon^{25})$, improving the previous complexity of $\tilde{\mathcal{O}}(n^{115}(1+\beta^{46})/\epsilon^{25})$ [39]. These applications are further discussed in Sec. IV C. Table I compares our results against the best previous ones for simulating quantum dynamics, simulating local observables, and quantum Monte Carlo simulation.

Given the numerous applications our bound provides in the asymptotic regime, we ask whether it has a favorable constant prefactor as well. This consideration is relevant to the practical performance of product formulas, especially for near-term quantum simulation experiments. For a two-term Hamiltonian, we show that our bound reduces to the known analyses of the Lie-Trotter formula [40,41] and the second-order Suzuki formula [9,42,43]. We then bootstrap the result to analyze Hamiltonians with an arbitrary number of summands (Sec. V A). The resulting bound matches the lowest-order term of the BCH expansion up to an application of the triangle inequality, and our analysis is thus provably tight for these low-order formulas.

We further numerically implement our bound for a one-dimensional Heisenberg model with a random magnetic field. This model can be simulated to understand condensed matter phenomena, but even a simulation of modest size seems to be infeasible for current classical computers. Childs *et al.* compared different quantum simulation algorithms for this model [31] and observed that product formulas have the best empirical performance, although their provable bounds were off by orders of magnitude even for systems of modest size, making it hard to identify with confidence the most efficient approach for near-term simulation. Reference [35] claimed an improved fourth-order bound that is off by a factor of about 17. Here, we give a tight bound that overestimates by only a factor of about 5. We also give a nearly tight Trotter error bound for power-law interactions. We describe the numerical implementation of our bound in detail in Sec. V B.

Underpinning these improvements is a theory we develop concerning the types, order conditions, and representations of the Trotter error. We illustrate these concepts in Sec. III A with the simple example of the first-order Lie-Trotter formula.

Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a sum of operators, and let $\mathcal{S}(t)$ be a product formula corresponding to this decomposition. We say that $\mathcal{A}(t)$, $\mathcal{M}(t)$, and $\mathcal{E}(t)$ are the additive, multiplicative, and exponentiated Trotter error if

$$\mathcal{S}(t) = e^{tH} + \mathcal{A}(t), \tag{3}$$

$$\mathcal{S}(t) = e^{tH}(I + \mathcal{M}(t)), \tag{4}$$

$$\mathcal{S}(t) = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau (H + \mathcal{E}(\tau))\right), \tag{5}$$

respectively, where $\mathcal{T}$ exp denotes the time-ordered matrix exponential. For applications in digital quantum simulation, these three types of Trotter error are equivalent to each other. However, the multiplicative type and the exponentiated type are more versatile for analyzing quantum Monte Carlo simulation. We give a constructive definition of these error types and discuss how they are related in Sec. III B.

A pth-order product formula $\mathcal{S}(t)$ can approximate the ideal evolution to pth order, in the sense that $\mathcal{S}(t) = e^{tH} + O(t^{p+1})$. Motivated by this approximation, we say that an operator-valued function $\mathcal{F}(t)$ satisfies the pth-order condition if $\mathcal{F}(t) = O(t^{p+1})$. In Sec. III C, we give order conditions for the Trotter error and its various derived operators. One significance of order conditions is that they can be used to cancel low-order terms. In particular, if $\mathcal{F}(t)$ satisfies the pth-order condition, then all terms with order at most p vanish in the Taylor series. This result can be verified by brute-force differentiation when $\mathcal{F}(t)$ is explicitly given, but applying the correct order condition avoids such a cumbersome calculation.

We then consider representations of the Trotter error in Sec. III D. Our representation only involves finitely many error terms, each of which is given by a nested commutator of operator summands. As mentioned earlier, these features overcome the drawbacks of previous representations and motivate a host of new applications. In deriving our representation, we work in a general setting where operator summands are not necessarily anti-Hermitian, so our analysis simultaneously handles real-time evolutions for digital quantum simulation and imaginary-time evolutions for quantum Monte Carlo simulation.

Sec. II gives a summary of background material that is necessary for understanding our Trotter error theory and its applications. Sec. VI concludes the paper with a brief discussion of the results and some open questions.

II. PRELIMINARIES

In this section, we summarize the preliminaries that we use in subsequent sections of the paper. Specifically, we introduce notation and terminology in Sec. II A, including various notions of norms and common asymptotic notations. In Sec. II B, we discuss time-ordered evolutions and the properties that are relevant to our analysis. We then define general product formulas in Sec. II C and prove a Trotter error bound with 1-norm scaling. Readers who are familiar with these preliminaries may skip ahead to Sec. III for the main result of our paper.

A. Notation and terminology

Unless otherwise noted, we use lowercase Latin letters to represent scalars, such as the evolution time t, the system size n, and the order of a product formula p. We also use the Greek alphabet to denote scalars, especially when we want to write a summation like $\sum_{\gamma=1}^{\Gamma}$. We use uppercase Latin letters, such as A, to denote operators. Throughout the paper, we assume that the underlying Hilbert space is finite dimensional, and operators can be represented by complex square matrices. We expect that some of our analyses can be generalized to spaces with infinite dimensions, but we restrict ourselves to the finite-dimensional setting since this is most relevant for applications to digital quantum simulation and quantum Monte Carlo simulation. We use scripted uppercase letters, such as $\mathscr{F}(t)$, to denote operator-valued functions.

We organize scalars to form vectors h_γ and tensors $h_{\gamma_1,\ldots,\gamma_k}$. We use standard norms for tensors, including the 1-norm $\|h\|_1 \coloneqq \sum_{\gamma_1,\ldots,\gamma_k} |h_{\gamma_1,\ldots,\gamma_k}|$, the Euclidean norm (or 2-norm) $\|h\|_2 \coloneqq \sqrt{\sum_{\gamma_1,\ldots,\gamma_k} |h_{\gamma_1,\ldots,\gamma_k}|^2}$, and the ∞-norm $\|h\|_\infty \coloneqq \max_{\gamma_1,\ldots,\gamma_k} |h_{\gamma_1,\ldots,\gamma_k}|$. In case there is ambiguity, we use $\vec{h}$ to emphasize the fact that h is a vector (or a tensor, more generally).

For an operator A, we use $\|A\|$ to denote its spectral norm—the largest singular value of A. The spectral norm is also known as the operator norm. It is a matrix norm that satisfies the scaling property $\|aA\| = |a|\|A\|$, the submultiplicative property $\|AB\| \le \|A\|\|B\|$, and the triangle inequality $\|A+B\| \le \|A\| + \|B\|$. If A is unitary, then $\|A\| = 1$. We further use $A_{\gamma_1,\ldots,\gamma_k}$ to denote a tensor where each elementary object is an operator. We define a norm of $A_{\gamma_1,\ldots,\gamma_k}$ by taking the spectral norm of each elementary operator and evaluating the corresponding norm of the resulting tensor. For example, we have $\|A\|_1 \coloneqq \sum_{\gamma_1,\ldots,\gamma_k} \|A_{\gamma_1,\ldots,\gamma_k}\|$ and $\|A\|_\infty \coloneqq \max_{\gamma_1,\ldots,\gamma_k} \|A_{\gamma_1,\ldots,\gamma_k}\|$.

For a tensor $A_{\gamma_1,\ldots,\gamma_k}$, we define

$$|||A|||_1 \coloneqq \max_j \max_{\gamma_j} \sum_{\substack{\gamma_1,\ldots,\gamma_{j-1},\\ \gamma_{j+1},\ldots,\gamma_k}} \|A_{\gamma_1,\ldots,\gamma_k}\|. \tag{6}$$

We call $|||A|||_1$ the induced 1-norm of A since it can be seen as a generalization of the induced 1-norm $\max_{\gamma_2} \sum_{\gamma_1} |a_{\gamma_1,\gamma_2}|$ of a matrix a_{γ_1,γ_2} [54]. A quantum simulation algorithm with induced 1-norm scaling runs faster than a 1-norm scaled algorithm because

$$|||A|||_1 \le \|A\|_1. \tag{7}$$

In fact, as we will see in Sec. IV A, the gap between these two norms can be significant for many realistic systems.

We also use $\mathrm{ad}_A B$ to denote the commutator $[A,B]$. Using a standard Lie-algebraic identity, we have $e^A B e^{-A} = e^{\mathrm{ad}_A} B$.

Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be functions of real variables. We write $f = O(g)$ if there exist $c, t_0 > 0$ such that $|f(\tau)| \le c|g(\tau)|$ whenever $|\tau| \le t_0$. Note that we consider the limit when the variable τ approaches zero as opposed to infinity, which is different from the usual setting of algorithmic analysis. For that purpose, we write $f = \mathcal{O}(g)$ if there exist $c, t_1 > 0$ such that $|f(\tau)| \le c|g(\tau)|$ for all $|\tau| \ge t_1$. When there is no ambiguity, we will use $f = \mathcal{O}(g)$ to also represent the case where $|f(\tau)| \le c|g(\tau)|$ holds for all $\tau \in \mathbb{R}$. We then extend the definition of $\mathcal{O}$ to functions of positive integers and multivariate functions. For example, we use $f(n,t,1/\epsilon) = \mathcal{O}((nt)^2/\epsilon)$ to mean that $|f(n,t,1/\epsilon)| \le c(n|t|)^2/\epsilon$ for some c, n_0, t_0, $\epsilon_0 > 0$ and all $|t| \ge t_0$, $0 < \epsilon < \epsilon_0$, and integers $n \ge n_0$. If $\mathscr{F}(\tau)$ is an operator-valued function, we first compute its spectral norm and analyze the asymptotic scaling of $\|\mathscr{F}(\tau)\|$. We write $f = \Omega(g)$ if $g = \mathcal{O}(f)$, and $f = \Theta(g)$ if both $f = \mathcal{O}(g)$ and $f = \Omega(g)$. We use $\tilde{\mathcal{O}}$ to suppress logarithmic factors in the asymptotic expression and $o(1)$ to represent a positive number that approaches zero as some parameter grows.

Finally, we use $\overleftarrow{\prod}$, $\prod_{\gamma=1}^{\Gamma}$ to denote a product where the elements have increasing indices from right to left and $\overrightarrow{\prod}$, $\prod_{\gamma=\Gamma}^{1}$ vice versa. Under this convention,

$$\prod_{\gamma=1}^{\Gamma} A_\gamma = \overleftarrow{\prod_{\gamma}} A_\gamma = A_\Gamma \cdots A_2 A_1, \tag{8}$$

$$\prod_{\gamma=\Gamma}^{1} A_\gamma = \overrightarrow{\prod_{\gamma}} A_\gamma = A_1 A_2 \cdots A_\Gamma. \tag{9}$$

We let a summation be zero if its lower limit exceeds its upper limit.

B. Time-ordered evolutions

Let $\mathcal{H}(\tau)$ be an operator-valued function defined for $0 \le \tau \le t$. We say that $\mathcal{U}(\tau)$ is the time-ordered evolution generated by $\mathscr{H}(\tau)$ if $\mathcal{U}(0) = I$ and $(\mathrm{d}/\mathrm{d}\tau)\mathcal{U}(\tau) = \mathscr{H}(\tau)\mathcal{U}(\tau)$ for $0 \le \tau \le t$. In the case where $\mathscr{H}(\tau)$ is anti-Hermitian, the function $\mathcal{U}(\tau)$ represents the evolution of a quantum system under Hamiltonian $i\mathscr{H}(\tau)$. We do not impose any restrictions on the Hermiticity of $\mathscr{H}(\tau)$ in the

development of our theory, so our analysis can be applied to not only real-time but also imaginary-time evolutions. Throughout this paper, we assume that operator-valued functions are continuous, which guarantees the existence and uniqueness of their generated evolutions (see p. 12 of Ref. [55]). We then formally represent the time-ordered evolution $\mathcal{U}(t)$ by $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau \mathscr{H}(\tau))$, where $\mathcal{T}\exp$ denotes the time-ordered exponential. In the special case where $\mathscr{H}(\tau) = H$ is constant, the generated evolution is given by an ordinary matrix exponential $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau \mathscr{H}(\tau)) = e^{tH}$.

In a similar way, we define the time-ordered evolution $\mathcal{T}\exp(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau))$ on an arbitrary interval $t_1 \le \tau \le t_2$. Its determinant satisfies (see p. 9 of Ref. [55])

$$\det\left(\mathcal{T}\exp\left(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau)\right)\right) = e^{\int_{t_1}^{t_2} \mathrm{d}\tau \mathrm{Tr}(\mathscr{H}(\tau))} \neq 0,$$

so the inverse operator $\mathcal{T}\exp^{-1}(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau))$ exists; we denote it by $\mathcal{T}\exp(\int_{t_2}^{t_1} \mathrm{d}\tau \mathscr{H}(\tau))$. We have thus defined $\mathcal{T}\exp(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau))$ for every pair of t_1 and t_2 in the domain of $\mathscr{H}(\tau)$ [56]. In the Appendix A, we list a few identities of time-ordered exponentials that are useful in our analysis.

C. Product formulas

Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a time-independent operator consisting of Γ summands so that the evolution generated by H is $e^{t\sum_{\gamma=1}^{\Gamma} H_\gamma}$. Product formulas provide a convenient way of decomposing such an evolution into a product of exponentials of individual H_γ. Examples of product formulas include the first-order Lie-Trotter formula

$$\mathcal{S}_1(t) := e^{tH_\Gamma} \cdots e^{tH_1} \tag{10}$$

and higher-order Suzuki formulas [21] defined recursively via

$$\begin{aligned} \mathcal{S}_2(t) &:= e^{(t/2)H_1} \cdots e^{(t/2)H_\Gamma} e^{(t/2)H_\Gamma} \cdots e^{(t/2)H_1}, \\ \mathcal{S}_{2k}(t) &:= \mathcal{S}_{2k-2}^2(u_k t)\mathcal{S}_{2k-2}((1-4u_k)t)\mathcal{S}_{2k-2}^2(u_k t), \end{aligned} \tag{11}$$

where $u_k := 1/(4 - 4^{1/(2k-1)})$. It is a challenge in practice to find the formula with the best performance for simulating a specific physical system [31]. However, we address a different question, developing a theory of Trotter error that holds for a general product formula. For in-depth studies of these formulas, especially in the context of numerical analysis, we refer the reader to Refs. [45–50] and the references therein.

Specifically, we consider a product formula of the form

$$\mathcal{S}(t) := \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\Gamma} e^{t a_{(\upsilon,\gamma)} H_{\pi_\upsilon(\gamma)}}, \tag{12}$$

where the coefficients $a_{(\upsilon,\gamma)}$ are real numbers. The parameter Υ denotes the number of *stages* of the formula; for the Suzuki formula $\mathcal{S}_{2k}(t)$, we have $\Upsilon = 2 \times 5^{k-1}$. The permutation π_υ controls the ordering of operator summands within stage υ of the formula. For Suzuki's constructions, we alternately reverse the ordering of summands between neighboring stages, but other formulas may use general permutations. Throughout this paper, we fix Υ, π_υ and assume that the coefficients $a_{(\upsilon,\gamma)}$ are uniformly bounded by 1 in absolute value. We then consider the performance of the product formula with respect to the input operator summands H_γ (for $\gamma = 1, \ldots, \Gamma$) and the evolution time t.

Product formulas provide a good approximation to the ideal evolution when the time t is small. Specifically, a pth-order formula $\mathcal{S}(t)$ satisfies

$$\mathcal{S}(t) = e^{tH} + O(t^{p+1}). \tag{13}$$

This asymptotic analysis gives the correct error scaling with respect to t; however, the dependence on the H_γ is ignored, so it does not provide a full characterization of the Trotter error. This issue was addressed in the work of Berry, Ahokas, Cleve, and Sanders [4], who gave a concrete error bound for product formulas with dependence on both t and H_γ. Their original bound depends on the ∞-norm $\Gamma \max_\gamma \|H_\gamma\|$, although it is not hard to improve this to the 1-norm scaling $\sum_{\gamma=1}^{\Gamma} \|H_\gamma\|$.

Lemma 1: (Trotter error with 1-norm scaling). Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator consisting of Γ summands and $t \ge 0$. Let $\mathcal{S}(t) = \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\Gamma} e^{t a_{(\upsilon,\gamma)} H_{\pi_\upsilon(\gamma)}}$ be a pth-order product formula. Then,

$$\|\mathcal{S}(t) - e^{tH}\| = \mathcal{O}\left(\left(\sum_{\gamma=1}^{\Gamma} \|H_\gamma\| t\right)^{p+1} e^{t\Upsilon \sum_{\gamma=1}^{\Gamma} \|H_\gamma\|}\right). \tag{14}$$

Furthermore, if H_γ are anti-Hermitian,

$$\|\mathcal{S}(t) - e^{tH}\| = \mathcal{O}\left(\left(\sum_{\gamma=1}^{\Gamma} \|H_\gamma\| t\right)^{p+1}\right). \tag{15}$$

We provide a proof of Lemma 1 in Appendix B. For real-time evolutions, this lemma improves a multiplicative factor of $e^{t\Upsilon \sum_{\gamma=1}^{\Gamma} \|H_\gamma\|}$ over the best previous analysis [see Eq. (13) in Ref. [28]].

The above bound on the Trotter error works well for small t. To simulate anti-Hermitian H_γ for a large time, we divide the evolution into r steps and apply the product formula within each step. The overall simulation has error

$$\|\mathcal{S}^r(t/r) - e^{tH}\| = \mathcal{O}\left(\frac{(\sum_{\gamma=1}^{\Gamma} \|H_\gamma\| t)^{p+1}}{r^p}\right). \tag{16}$$

To simulate with accuracy ϵ, it suffices to choose

$$r = \mathcal{O}\left(\frac{(\sum_{\gamma=1}^{\Gamma} \|H_\gamma\| t)^{1+1/p}}{\epsilon^{1/p}}\right). \tag{17}$$

We therefore arrive at the following corollary

Corollary 2: (Trotter number with 1-norm scaling) Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator consisting of Γ summands with H_γ anti-Hermitian and $t \geq 0$. Let $\mathcal{S}(t)$ be a pth-order product formula. Then, we have $\|\mathcal{S}^r(t/r) - e^{tH}\| = \mathcal{O}(\epsilon)$ provided

$$r = \mathcal{O}\left(\frac{(\sum_{\gamma=1}^{\Gamma} \|H_\gamma\| t)^{1+1/p}}{\epsilon^{1/p}}\right). \tag{18}$$

Note that the above analysis only uses information about the norms of the summands. In the extreme case where all H_γ commute, the Trotter error becomes zero, but the above bound can be arbitrarily large. This contradiction suggests that the analysis can be significantly improved by leveraging information about commutation of the H_γ. Unfortunately, despite extensive efforts, dramatic improvements to the Trotter error bound are only known for certain low-order formulas [9,40–43] and special systems [35,44].

To explain the limitations of prior approaches, it is instructive to examine a general bound developed by Descombes and Thalhammer [42,45],

$$\|\mathcal{S}(t) - e^{tH}\| \leq a_{p+1} t^{p+1} + \cdots + a_q t^q + a_{q+1} t^{q+1}, \tag{19}$$

where $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ is a sum of anti-Hermitian operators, $\mathcal{S}(t)$ is a pth-order formula, $q \geq p$ is a positive integer, and $t \geq 0$, suggesting a choice of

$$r = \max\left\{\mathcal{O}\left(\frac{a_{p+1}^{1/p} t^{1+1/p}}{\epsilon^{1/p}}\right), \ldots, \mathcal{O}\left(\frac{a_q^{1/(q-1)} t^{1+1/(q-1)}}{\epsilon^{1/(q-1)}}\right), \mathcal{O}\left(\frac{a_{q+1}^{1/q} t^{1+1/q}}{\epsilon^{1/q}}\right)\right\} \tag{20}$$

to simulate with accuracy ϵ. Here, all the leading coefficients $a_{p+1}, \ldots, a_q$ depend on nested commutators of H_γ, but a_{q+1} is determined by commutators interlaced with matrix exponentials, which is technically challenging to evaluate except for geometrically local systems. Consequently, a bound on a_{q+1} must be used, resulting in a 1-norm scaling similar to that of Lemma 1 and a loose Trotter error estimate for simulating general quantum systems.

We develop a theory of Trotter error that directly exploits the commutativity of operator summands. The resulting bound naturally reduces to the previous bounds for low-order formulas and special systems, but our analysis uncovers a host of new speedups for product formulas that were previously unknown. The central concepts of this theory are the types, order conditions, and representations of the Trotter error, which we explain in Sec. III.

III. THEORY

We now develop a theory for analyzing the Trotter error. We explain the core ideas of this theory in Sec. III A using the simple example of the first-order Lie-Trotter formula. We then discuss the analysis of a general formula. In particular, we study various types of Trotter error in Sec. III B and compute their order conditions in Sec. III C. We then derive explicit representations of the Trotter error in Sec. III D, establishing the commutator scaling of the Trotter error in Theorem 6. We focus on the asymptotic error scaling here and discuss potential applications and constant-prefactor improvements of our results in Secs. IV and Sec. V, respectively.

A. Example of the Lie-Trotter formula

In this section, we use the example of the first-order Lie-Trotter formula to illustrate the general theory we develop for analyzing the Trotter error. For simplicity, consider an operator $H = A + B$ with two summands. The ideal evolution generated by H is given by $e^{tH} = e^{t(A+B)}$. To decompose this evolution, we may use the Lie-Trotter formula $\mathcal{S}_1(t) = e^{tB} e^{tA}$. This formula is first-order accurate, so we have $\mathcal{S}_1(t) = e^{tH} + O(t^2)$.

A key observation here is that the error of a product formula can have various *types*. Specifically, we consider three types of Trotter error: additive error, multiplicative error, and error that appears in the exponent. Note that $\mathcal{S}_1(t)$ satisfies the differential equation $(\mathrm{d}/\mathrm{d}t)\mathcal{S}_1(t) = H\mathcal{S}_1(t) + [e^{tB}, A]e^{tA}$ with the initial condition $\mathcal{S}_1(0) = I$. By the variation-of-parameters formula (Lemma A.1),

$$\mathcal{S}_1(t) = e^{tH} + \int_0^t \mathrm{d}\tau e^{(t-\tau)H}[e^{\tau B}, A]e^{\tau A}, \tag{21}$$

so we get the additive error $\mathcal{A}_1(t) = \int_0^t \mathrm{d}\tau e^{(t-\tau)H}[e^{\tau B}, A]e^{\tau A}$ of the Lie-Trotter formula. For error with the exponentiated type, we differentiate $\mathcal{S}_1(t)$ to get $(\mathrm{d}/\mathrm{d}t)\mathcal{S}_1(t) = (B + e^{t\mathrm{ad}_B} A)\mathcal{S}_1(t)$. Applying the fundamental theorem of time-ordered evolution (Lemma A.3), we have

$$\mathcal{S}_1(t) = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau (B + e^{\tau \mathrm{ad}_B} A)\right), \tag{22}$$

so $\mathcal{E}_1(\tau) = e^{\tau \mathrm{ad}_B} A - A$ is the error of the Lie-Trotter formula that appears in the exponent. To obtain the multiplicative error, we switch to the interaction picture using Lemma A.2:

$$\mathcal{S}_1(t) = e^{tH}\mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau e^{-\tau \mathrm{ad}_H}(e^{\tau \mathrm{ad}_B}A - A)\right),$$

so $\mathscr{M}_1(t) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau e^{-\tau \mathrm{ad}_H}(e^{\tau \mathrm{ad}_B}A - A)) - I$ is the multiplicative Trotter error.

These three types of Trotter error are equivalent for analyzing the complexity of digital quantum simulation (Sec. IV A) and simulating local observables (Sec. IV B), whereas the multiplicative error and the exponentiated error are more versatile when applied to quantum Monte Carlo simulation (Sec. IV C). We compute the error operators for a general product formula in Sec. III B.

Since product formulas provide a good approximation to the ideal evolution for small t, we expect all three error operators $\mathscr{A}_1(t)$, $\mathcal{E}_1(t)$, and $\mathscr{M}_1(t)$ to converge to zero in the limit $t \to 0$. The rates of convergence are what we call order conditions. More precisely,

$$\begin{aligned}
\mathscr{A}_1(t) &= \int_0^t \mathrm{d}\tau\, e^{(t-\tau)H}[e^{\tau B}, A]e^{\tau A} = O(t^2),\\
\mathcal{E}_1(t) &= e^{t\mathrm{ad}_B}A - A = O(t),\\
\mathscr{M}_1(t) &= \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau e^{-\tau \mathrm{ad}_H}(e^{\tau \mathrm{ad}_B}A - A)\right) - I\\
&= O(t^2).
\end{aligned}$$

For the Lie-Trotter formula, these conditions can be verified by direct calculation, although such an approach becomes inefficient, in general. Instead, we describe an indirect approach in Sec. III C to compute order conditions for a general product formula.

Finally, we consider representations of the Trotter error that leverage the commutativity of operator summands. We discuss how to represent $\mathscr{M}_1(t)$ in detail, although it is straightforward to extend the analysis to $\mathscr{A}_1(t)$ and $\mathcal{E}_1(t)$ as well. To this end, we first consider the term $e^{-\tau \mathrm{ad}_H}e^{\tau \mathrm{ad}_B}A$, which contains two layers of conjugations of matrix exponentials. We apply the fundamental theorem of calculus to the first layer of conjugation and obtain

$$e^{\tau \mathrm{ad}_B}A = A + \int_0^\tau \mathrm{d}\tau_2 e^{\tau_2 \mathrm{ad}_B}\mathrm{ad}_B A. \tag{23}$$

Substituting into the expression of $\mathscr{M}_1(t)$, we get

$$\mathscr{M}_1(t) = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau \int_0^\tau \mathrm{d}\tau_2 e^{-\tau \mathrm{ad}_H}e^{\tau_2 \mathrm{ad}_B}\mathrm{ad}_B A\right) - I, \tag{24}$$

which implies, through Corollary A.5, that $\|\mathscr{M}_1(t)\| = \mathcal{O}(\|[B,A]\|t^2)$ when A, B are anti-Hermitian and $t \geq 0$. In the above derivation, it is important that we only expand the first layer of conjugation of exponentials, that we apply the fundamental theorem of calculus only once, and that we can cancel the terms $e^{-\tau H}Ae^{\tau H}$ in pairs. The validity of such an approach, in general, is guaranteed by the appropriate order condition, which we explain in detail in Sec. III D.

B. Error types

In this section, we discuss error types of a general product formula. In particular, we give explicit expressions for three different types of Trotter error: the additive error, the multiplicative error, and error that appears in the exponent of a time-ordered exponential (the "exponentiated" error). These types are equivalent for analyzing the complexity of simulating quantum dynamics and local observables, but the latter two types are more versatile for quantum Monte Carlo simulation.

Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator with Γ summands. The ideal evolution under H for time t is given by $e^{tH} = e^{t\sum_{\gamma=1}^{\Gamma} H_\gamma}$, which we approximate by a general product formula $\mathcal{S}(t) = \prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma} e^{ta_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}$. For convenience, we use the lexicographic order on a pair of tuples (υ,γ) and (υ',γ'), defined as follows: We write $(\upsilon,\gamma) \succeq (\upsilon',\gamma')$ if $\upsilon > \upsilon'$, or if $\upsilon = \upsilon'$ and $\gamma \geq \gamma'$. We have $(\upsilon,\gamma) \succ (\upsilon',\gamma')$ if both $(\upsilon,\gamma) \succeq (\upsilon',\gamma')$ and $(\upsilon,\gamma) \neq (\upsilon',\gamma')$ hold. Notations $(\upsilon,\gamma) \preccurlyeq (\upsilon',\gamma')$ and $(\upsilon,\gamma) \prec (\upsilon',\gamma')$ are defined in a similar way, except that we reverse the directions of all the inequalities.

For the additive Trotter error, we seek an operator-valued function $\mathscr{A}(t)$ such that $\mathcal{S}(t) = e^{tH} + \mathscr{A}(t)$. This can be achieved by constructing the differential equation $(\mathrm{d}/\mathrm{d}t)\mathcal{S}(t) = H\mathcal{S}(t) + \mathcal{R}(t)$ with initial condition $\mathcal{S}(0) = I$, followed by the use of the variation-of-parameters formula (Lemma A.1). For the exponentiated type of Trotter error, we aim to construct an operator-valued function $\mathcal{E}(t)$ such that $\mathcal{S}(t) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau(H + \mathcal{E}(\tau)))$. We find $\mathcal{E}(t)$ by differentiating the product formula $\mathcal{S}(t)$ and applying the fundamental theorem of time-ordered evolution (Lemma A.3). Finally, we obtain the multiplicative error by switching to the interaction picture using Lemma A.2. The derivation follows from a similar analysis as in Sec. III A and is detailed in Appendix C.

Theorem 3: (Types of Trotter error). Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator with Γ summands. The evolution under H for time $t \in \mathbb{R}$ is given by $e^{tH} = e^{t\sum_{\gamma=1}^{\Gamma} H_\gamma}$, which we decompose using the product formula $\mathcal{S}(t) = \prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma} e^{ta_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}$. Then, the following statements holds:

(1) Trotter error can be expressed in the additive form $\mathcal{S}(t) = e^{tH} + \int_0^t \mathrm{d}\tau e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau)$, where

$$\mathcal{T}(\tau) = \sum_{(v,\gamma)} \overrightarrow{\prod_{\substack{(v',\gamma') \\ \prec (v,\gamma)}}} e^{-\tau a_{(v',\gamma')} \mathrm{ad}_{H_{\pi_{v'}(\gamma')}}} (a_{(v,\gamma)} H_{\pi_v(\gamma)}) - \overrightarrow{\prod_{(v',\gamma')}} e^{-\tau a_{(v',\gamma')} \mathrm{ad}_{H_{\pi_{v'}(\gamma')}}} H. \tag{25}$$

(2) Trotter error can be expressed in the exponentiated form $\mathcal{S}(t) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau(H + \mathcal{E}(\tau)))$, where

$$\mathcal{E}(\tau) = \sum_{(v,\gamma)} \overleftarrow{\prod_{\substack{(v',\gamma') \\ \succ (v,\gamma)}}} e^{\tau a_{(v',\gamma')} \mathrm{ad}_{H_{\pi_{v'}(\gamma')}}} (a_{(v,\gamma)} H_{\pi_v(\gamma)}) - H. \tag{26}$$

(3) Trotter error can be expressed in the multiplicative form $\mathcal{S}(t) = e^{tH}(I + \mathcal{M}(t))$, where

$$\mathcal{M}(t) = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau\, e^{-\tau \mathrm{ad}_H}\mathcal{E}(\tau)\right) - I, \tag{27}$$

with $\mathcal{E}(\tau)$ as above.

Note that the error operators $\mathcal{T}(\tau)$ and $\mathcal{E}(\tau)$ both consist of conjugations of matrix exponentials of the form $e^{\tau A_s}\cdots e^{\tau A_2} e^{\tau A_1} B e^{-\tau A_1} e^{-\tau A_2}\cdots e^{-\tau A_s}$. To bound the Trotter error, it thus suffices to analyze such conjugations of matrix exponentials. The previous work of Somma [44] expanded them into infinite series of nested commutators, which is favorable for systems with appropriate Lie-algebraic structures. An alternative approach of Childs and Su [35] represented them as commutators nested with conjugations of matrix exponentials, which provides a tight analysis for geometrically local systems. Unfortunately, both approaches can be loose, in general. Instead, we apply order conditions (Sec. III C) and derive a new representation of the Trotter error (Sec. III D) that provides a tight analysis for general systems.

C. Order conditions

In this section, we study the order conditions of the Trotter error. By order condition, we mean the rate at which a continuous operator-valued function $\mathcal{F}(\tau)$, defined for $\tau \in \mathbb{R}$, approaches zero in the limit $\tau \to 0$. Formally, we write $\mathcal{F}(\tau) = O(\tau^p)$ with non-negative integer p if there exist constants c, $t_0 > 0$, independent of τ, such that $\|\mathcal{F}(\tau)\| \le c|\tau|^p$ whenever $|\tau| \le t_0$.

Order conditions arise naturally in the analysis of the Trotter error [21,57–59]. Indeed, a pth-order product formula $\mathcal{S}(t)$ has a Taylor expansion that agrees with the ideal evolution e^{tH} up to order t^p, which implies the order condition $\mathcal{S}(t) = e^{tH} + O(t^{p+1})$ by definition. Our approach is to use this relation in the reverse direction: Given a smooth operator-valued function $\mathcal{F}(\tau)$ satisfying the order condition $\mathcal{F}(\tau) = O(\tau^p)$, we conclude that $\mathcal{F}(\tau)$ has a Taylor expansion where terms with order τ^{p-1} or lower vanish. We make this argument more precise in Appendix D.

We can determine the order condition of an operator-valued function through either direct calculation or indirect derivation. To illustrate this, we consider decomposing $e^{tH} = e^{t(A+B)}$ using the first-order Lie-Trotter formula $\mathcal{S}_1(t) = e^{tB}e^{tA}$. We see from Sec. III A that this decomposition has the additive Trotter error

$$\mathcal{A}_1(t) = \int_0^t \mathrm{d}\tau\, e^{(t-\tau)H}(\mathcal{S}_1'(\tau) - H\mathcal{S}_1(\tau)) = \int_0^t \mathrm{d}\tau\, e^{(t-\tau)H}[e^{\tau B}, A]e^{\tau A}. \tag{28}$$

We know that $\mathcal{A}_1(t)$ has the order condition $\mathcal{A}_1(t) = O(t^2)$, which follows directly from the fact that $\mathcal{A}_1(0) = \mathcal{A}_1'(0) = 0$. On the other hand, an indirect argument would proceed as follows. We use the known order condition $\mathcal{S}_1(t) = e^{tH} + O(t^2)$ to conclude that $\mathcal{S}_1'(\tau) - H\mathcal{S}_1(\tau) = O(\tau)$. Multiplying the matrix exponential $e^{(t-\tau)H} = O(1)$ does not change the order condition, so we still have $e^{(t-\tau)H}(\mathcal{S}_1'(\tau) - H\mathcal{S}_1(\tau)) = O(\tau)$. A final integration of $\int_0^t \mathrm{d}\tau$ then gives the desired condition $\mathcal{A}_1(t) = O(t^2)$.

Although we obtain the same order condition through two different analyses, the direct approach becomes inefficient for analyzing the Trotter error of a general high-order product formula. Instead, we use an indirect analysis to prove the following theorem on the order conditions of the Trotter error (see Appendix D for proof details).

Theorem 4: (Order conditions of Trotter error). Let H be an operator, and let $\mathcal{S}(\tau)$, $\mathcal{T}(\tau)$, $\mathcal{E}(\tau)$, and $\mathcal{M}(\tau)$ be infinitely differentiable operator-valued functions defined for $\tau \in \mathbb{R}$, such that

$$\mathcal{S}(t) = e^{tH} + \int_0^t \mathrm{d}\tau\, e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau), = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau(H + \mathcal{E}(\tau))\right), = e^{tH}(I + \mathcal{M}(t)). \tag{29}$$

For any non-negative integer p, the following conditions are equivalent:

(1) $\mathcal{S}(t) = e^{tH} + O(t^{p+1})$,
(2) $\mathcal{T}(\tau) = O(\tau^p)$,
(3) $\mathcal{E}(\tau) = O(\tau^p)$, and
(4) $\mathcal{M}(t) = O(t^{p+1})$.

In Sec. III D, we apply these conditions to cancel low-order Trotter error terms and represent higher-order ones as nested commutators of operator summands.

D. Error representations

For a product formula with a certain error type and order condition, we now represent its error in terms of nested commutators of the operator summands. In particular, we give upper bounds on the additive and the multiplicative errors of pth-order product formulas in Theorem 6.

Consider an operator $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ with Γ summands. The ideal evolution generated by H is $e^{tH} = e^{t\sum_{\gamma=1}^{\Gamma} H_\gamma}$, which we decompose using a pth-order product formula $\mathcal{S}(t) = \prod_{v=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma} e^{ta_{(v,\gamma)}H_{\pi_v(\gamma)}}$. We know from Theorem 3 that the Trotter error can be expressed in the additive form $\mathcal{S}(t) = e^{tH} + \int_0^t \mathrm{d}\tau e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau)$, the multiplicative form $\mathcal{S}(t) = e^{tH}(I + \mathscr{M}(t))$, where $\mathscr{M}(t) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau e^{-\tau H}\mathcal{E}(\tau)e^{\tau H}) - I$, and the exponentiated form $\mathcal{S}(t) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau(H + \mathcal{E}(\tau)))$. Furthermore, both $\mathcal{T}(\tau)$ and $\mathcal{E}(\tau)$ consist of conjugations of matrix exponentials and have order condition $\mathcal{T}(\tau), \mathcal{E}(\tau) \in O(\tau^p)$ (Theorem 4).

We first consider the representation of a single conjugation of matrix exponentials,

$$e^{\tau A_{\mathcal{S}}}\cdots e^{\tau A_2}e^{\tau A_1}Be^{-\tau A_1}e^{-\tau A_2}\cdots e^{-\tau A_s}, \tag{30}$$

where $A_1, A_2, \ldots, A_s, B$ are operators and $\tau \in \mathbb{R}$. Our goal is to expand this conjugation into a finite series in the time variable τ. We only keep track of those terms with order $O(\tau^p)$ because terms corresponding to $1, \tau, \ldots, \tau^{p-1}$ will vanish in the final representation of the Trotter error due to the order condition. As mentioned before, such a conjugation was previously analyzed based on a naive application of Taylor's theorem [35] and an infinite-series expansion [44]. However, those results do not represent the Trotter error as a finite number of commutators of operator summands, and they only apply to special systems such as those with geometrical locality or suitable Lie-algebraic structure. Our new representation overcomes these limitations.

Theorem 5: (Commutator expansion of a conjugation of matrix exponentials) Let $A_1, A_2, \ldots, A_s$ and B be operators. Then, the conjugation $e^{\tau A_s}\cdots e^{\tau A_2}e^{\tau A_1}Be^{-\tau A_1}e^{-\tau A_2}\cdots e^{-\tau A_s}$ $(\tau \in \mathbb{R})$ has the expansion

$$\begin{aligned} &e^{\tau A_s}\cdots e^{\tau A_2}e^{\tau A_1}Be^{-\tau A_1}e^{-\tau A_2}\cdots e^{-\tau A_s} \\ &\quad = C_0 + C_1\tau + \cdots + C_{p-1}\tau^{p-1} + \mathcal{C}(\tau). \end{aligned} \tag{31}$$

Here, $C_0, \ldots, C_{p-1}$ are operators independent of τ. The operator-valued function $\mathcal{C}(\tau)$ is given by

$$\begin{aligned} \mathcal{C}(\tau) \coloneqq &\sum_{k=1}^{s}\sum_{\substack{q_1+\cdots+q_k=p \\ q_k\neq 0}} e^{\tau\mathrm{ad}_{A_s}}\cdots e^{\tau\mathrm{ad}_{A_{k+1}}} \\ &\cdot\int_0^\tau \mathrm{d}\tau_2 e^{\tau_2\mathrm{ad}_{A_k}}\mathrm{ad}_{A_k}^{q_k}\cdots\mathrm{ad}_{A_1}^{q_1}B\frac{(\tau-\tau_2)^{q_k-1}\tau^{q_1+\cdots+q_{k-1}}}{(q_k-1)!q_{k-1}!\cdots q_1!}. \end{aligned}$$

Furthermore, we have the spectral-norm bound

$$\|\mathcal{C}(\tau)\| \le \alpha_{\mathrm{comm}}(A_s, \ldots, A_1, B)\frac{|\tau|^p}{p!}e^{2|\tau|\sum_{k=1}^{s}\|A_k\|}$$

for general operators and

$$\|\mathcal{C}(\tau)\| \le \alpha_{\mathrm{comm}}(A_s, \ldots, A_1, B)\frac{|\tau|^p}{p!} \tag{32}$$

when A_k $(k = 1, \ldots, s)$ are anti-Hermitian, where

$$\begin{aligned} &\alpha_{\mathrm{comm}}(A_s, \ldots, A_1, B) \\ &\quad\coloneqq \sum_{q_1+\cdots+q_s=p}\binom{p}{q_1\cdots q_s}\|\mathrm{ad}_{A_s}^{q_s}\cdots\mathrm{ad}_{A_1}^{q_1}(B)\|. \end{aligned} \tag{33}$$

Proof.—We begin with the innermost layer $e^{\tau\mathrm{ad}_{A_1}}B$. Applying Taylor's theorem to order $p-1$ with the integral form of the remainder, we have

$$\begin{aligned} e^{\tau\mathrm{ad}_{A_1}}B = &B + \tau\mathrm{ad}_{A_1}B + \cdots + \frac{\tau^{p-1}}{(p-1)!}\mathrm{ad}_{A_1}^{p-1}B \\ &+ \int_0^\tau \mathrm{d}\tau_2\frac{\tau_2^{p-1}}{(p-1)!}e^{(\tau-\tau_2)\mathrm{ad}_{A_1}}\mathrm{ad}_{A_1}^{p}B. \end{aligned} \tag{34}$$

By the multiplication rule and the integration rule of Proposition D.3, the last term has order

$$\int_0^\tau \mathrm{d}\tau_2\frac{\tau_2^{p-1}}{(p-1)!}e^{(\tau-\tau_2)\mathrm{ad}_{A_1}}\mathrm{ad}_{A_1}^{p}B = O(\tau^p). \tag{35}$$

This term cannot be canceled by the order condition, and we keep it in our expansion. The remaining terms corresponding to $1, \tau, \ldots, \tau^{p-1}$ are substituted back to the original conjugation of matrix exponentials.

We now consider the next layer of conjugation. We apply Taylor's theorem to the operators $e^{\tau\mathrm{ad}_{A_2}}B$, $e^{\tau\mathrm{ad}_{A_2}}\mathrm{ad}_{A_1}B, \ldots, e^{\tau\mathrm{ad}_{A_2}}\mathrm{ad}_{A_1}^{p-1}B$ to order $p-1, p-2, \ldots, 0$, respectively, obtaining

$$\begin{aligned} e^{\tau\mathrm{ad}_{A_2}}\mathrm{ad}_{A_1}^{q_1}B = &\mathrm{ad}_{A_1}^{q_1}B + \cdots + \frac{\tau^{p-2}}{(p-2)!}\mathrm{ad}_{A_2}^{p-1-q_1}\mathrm{ad}_{A_1}^{q_1}B \\ &+ \int_0^\tau \mathrm{d}\tau_2\frac{\tau_2^{p-1-q_1}}{(p-1-q_1)!}e^{(\tau-\tau_2)\mathrm{ad}_{A_2}}\mathrm{ad}_{A_2}^{p-q_1}\mathrm{ad}_{A_1}^{q_1}B, \end{aligned} \tag{36}$$

for $q_1 = 0, 1, \ldots, p-1$. Combined with the result from the first layer, the Taylor remainders in the above equation have order

$$\int_0^\tau \mathrm{d}\tau_2 \frac{\tau_2^{p-1}}{(p-1-q_1)!q_1!} e^{(\tau-\tau_2)\mathrm{ad}_{A_2}} \mathrm{ad}_{A_2}^{p-q_1} \mathrm{ad}_{A_1}^{q_1} B = O(\tau^p), \tag{37}$$

for all q_1. We keep these terms in our expansion and substitute the remaining ones back to the original conjugation of matrix exponentials.

We repeat this analysis for all the remaining layers of the conjugation of matrix exponentials. In doing so, we keep track of those terms with order $O(\tau^p)$, obtaining

$$\begin{aligned} e^{\tau A_s}\cdots e^{\tau A_2} e^{\tau A_1} B e^{-\tau A_1} e^{-\tau A_2}\cdots e^{-\tau A_s} &= C_0 + C_1\tau + \cdots + C_{p-1}\tau^{p-1} \\ &\quad + \sum_{k=1}^{s} \sum_{\substack{q_1+\cdots+q_k=p \\ q_k\neq 0}} e^{\tau \mathrm{ad}_{A_s}}\cdots e^{\tau \mathrm{ad}_{A_{k+1}}} \cdot \int_0^\tau \mathrm{d}\tau_2 \frac{(\tau-\tau_2)^{q_k-1}\tau^{q_1+\cdots+q_{k-1}}}{(q_k-1)!q_{k-1}!\cdots q_1!} \\ &\quad \cdot e^{\tau_2 \mathrm{ad}_{A_k}} \mathrm{ad}_{A_k}^{q_k}\cdots \mathrm{ad}_{A_1}^{q_1} B \end{aligned} \tag{38}$$

for some operators $C_0, C_1, \ldots, C_{p-1}$. Because of the order condition, the terms of order $1, \tau, \ldots, \tau^{p-1}$ will vanish in our final representation of the Trotter error.

We now bound the spectral norm of those terms with order $O(\tau^p)$. By the triangle inequality, we have an upper bound of

$$\begin{aligned} &\sum_{k=1}^{s} \sum_{\substack{q_1+\cdots+q_k=p \\ q_k\neq 0}} \int_0^{|\tau|} \mathrm{d}\tau_2 \frac{(|\tau|-\tau_2)^{q_k-1}|\tau|^{q_1+\cdots+q_{k-1}}}{(q_k-1)!q_{k-1}!\cdots q_1!} \|\mathrm{ad}_{A_k}^{q_k}\cdots \mathrm{ad}_{A_1}^{q_1}(B)\| e^{2|\tau|\sum_{l=1}^{s}\|A_l\|} \\ &= \sum_{k=1}^{s} \sum_{\substack{q_1+\cdots+q_k=p \\ q_k\neq 0}} \binom{p}{q_1\cdots q_k} \frac{|\tau|^p}{p!} \|\mathrm{ad}_{A_k}^{q_k}\cdots \mathrm{ad}_{A_1}^{q_1}(B)\| e^{2|\tau|\sum_{l=1}^{s}\|A_l\|} \\ &= \sum_{q_1+\cdots+q_s=p} \binom{p}{q_1\cdots q_s} \frac{|\tau|^p}{p!} \|\mathrm{ad}_{A_s}^{q_s}\cdots \mathrm{ad}_{A_1}^{q_1}(B)\| e^{2|\tau|\sum_{l=1}^{s}\|A_l\|} \\ &= \alpha_{\mathrm{comm}}(A_s,\ldots,A_1,B) \frac{|\tau|^p}{p!} e^{2|\tau|\sum_{l=1}^{s}\|A_l\|}. \end{aligned} \tag{39}$$

This bound holds for arbitrary operators $A_1, A_2, \ldots, A_s$. When these operators are anti-Hermitian, we can tighten the above analysis by evaluating the spectral norm of a matrix exponential as 1. We have therefore arrived at the theorem. ∎

We apply Theorem 5 to expand every conjugation of matrix exponentials of the error operators $\mathcal{T}(\tau)$ and $\mathcal{E}(\tau)$ into a finite series in τ. After taking the linear combination, we obtain

$$\begin{aligned} \mathcal{T}(\tau) &= T_0 + T_1\tau + \cdots + T_{p-1}\tau^{p-1} + \mathcal{T}_p(\tau), \\ \mathcal{E}(\tau) &= E_0 + E_1\tau + \cdots + E_{p-1}\tau^{p-1} + \mathcal{E}_p(\tau). \end{aligned} \tag{40}$$

The operator-valued functions $\mathcal{T}_p(\tau)$ and $\mathcal{E}_p(\tau)$ have order condition $O(\tau^p)$, whereas $T_0, \ldots, T_{p-1}$ and $E_0, \ldots, E_{p-1}$ are independent of τ. By Lemma D.1 and the order condition $\mathcal{T}(\tau), \mathcal{E}(\tau) \in O(\tau^p)$, we have

$$T_0 = \cdots = T_{p-1} = E_0 = \cdots = E_{p-1} = 0, \tag{41}$$

or equivalently,

$$\mathcal{T}(\tau) = \mathcal{T}_p(\tau), \qquad \mathcal{E}(\tau) = \mathcal{E}_p(\tau). \tag{42}$$

We then bound the spectral norm of $\mathcal{T}_p(\tau)$ and $\mathcal{E}_p(\tau)$ using Theorem 5. This argument establishes the commutator scaling of the Trotter error. We state the result below and leave the calculation details to Appendix E.

Theorem 6: (Trotter error with commutator scaling). Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator consisting of Γ summands and $t \geq 0$. Let $\mathcal{S}(t) = \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\Gamma} e^{t a_{(\upsilon,\gamma)} H_{\pi_\upsilon(\gamma)}}$ be a pth-order product formula. Define $\tilde{\alpha}_{\mathrm{comm}} = \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\cdots]\|$. Then, the additive Trotter error and the multiplicative Trotter error, defined, respectively, by $\mathcal{S}(t) = e^{tH} + \mathcal{A}(t)$ and $\mathcal{S}(t) = e^{tH}(I + \mathcal{M}(t))$, can be asymptotically bounded as

$$\|\mathscr{A}(t)\|, \|\mathscr{M}(t)\| = \mathcal{O}(\tilde{\alpha}_{\mathrm{comm}} t^{p+1} e^{4t\Upsilon \sum_{\gamma=1}^{\Gamma} \|H_\gamma\|}). \quad (43)$$

Furthermore, if H_γ are anti-Hermitian,

$$\|\mathscr{A}(t)\|, \|\mathscr{M}(t)\| = \mathcal{O}(\tilde{\alpha}_{\mathrm{comm}} t^{p+1}). \quad (44)$$

Corollary 7. (Trotter number with commutator scaling) Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator consisting of Γ summands with H_γ anti-Hermitian and $t \geq 0$. Let $\mathcal{S}(t) = \prod_{v=1}^{\Upsilon} \prod_{\gamma=1}^{\Gamma} e^{t a_{(v,\gamma)} H_{\pi_v(\gamma)}}$ be a pth-order product formula. Define $\tilde{\alpha}_{\mathrm{comm}} = \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\|$. Then, we have $\|\mathcal{S}^r(t/r) - e^{tH}\| = \mathcal{O}(\epsilon)$, provided that

$$r = \mathcal{O}\left(\frac{\tilde{\alpha}_{\mathrm{comm}}^{1/p} t^{1+1/p}}{\epsilon^{1/p}}\right). \quad (45)$$

For any $\delta > 0$, we can choose p sufficiently large so that $1/p < \delta$. For this choice of p, we have $r = \mathcal{O}(\tilde{\alpha}_{\mathrm{comm}}^{\delta} t^{1+\delta}/\epsilon^{\delta})$. Therefore, the Trotter number scales as $r = \tilde{\alpha}_{\mathrm{comm}}^{o(1)} t^{1+o(1)}$ if we simulate with constant accuracy. To obtain the asymptotic complexity of the product-formula algorithm, it thus suffices to compute the quantity $\tilde{\alpha}_{\mathrm{comm}} = \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\|$, which can often be done by induction. We illustrate this process by presenting a host of applications of our bound to simulating quantum dynamics (Sec. IV A), local observables (Sec. IV B), and quantum Monte Carlo methods (Sec. IV C).

Note that we did not evaluate the constant prefactor of our bound in Theorem 6. Indeed, our proof involves inequality zooming that suffices to establish the correct asymptotic scaling but is likely loose in practice. For practical implementation, it is better to use Theorem 5, which gives a concrete expression for the error operator. A general methodology to obtain error bounds with small constant factors is described in Appendix M. In Sec. V A, we show that our bound reduces to previous bounds for the Lie-Trotter formula [40,41] and the second-order Suzuki formula [9,42,43,46], which are known to be tight up to an application of the triangle inequality. We further provide numerical evidence in Sec. V B suggesting that our bound has a small prefactor for higher-order formulas as well.

IV. APPLICATIONS

Our main result (Theorem 6) on the commutator scaling of the Trotter error uncovers a host of speedups of the product-formula approach. In this section, we provide improved product-formula algorithms for digital quantum simulation (Sec. IV A), simulating local observables (Sec. IV B), and quantum Monte Carlo methods (Sec. IV C). We show that these results can nearly match or even outperform the best previous results for simulating quantum systems.

A. Applications to digital quantum simulation

We now present applications of our bound to digital quantum simulation, including simulations of second-quantized electronic structure, k-local Hamiltonians, rapidly decaying long-range and quasilocal interactions, and clustered Hamiltonians. Throughout this section, we let H be Hermitian and $t \geq 0$ be non-negative, and we consider the real-time evolution e^{-itH}.

Second-quantized electronic structure. Simulating electronic-structure Hamiltonians is one of the most widely studied applications of digital quantum simulation. An efficient solution of this problem could help design and engineer new pharmaceuticals, catalysts, and materials [11]. Recent studies have focused on solving this problem using more advanced simulation algorithms. Here, we demonstrate the power of product formulas for simulating electronic-structure Hamiltonians.

We consider the second-quantized representation of the electronic-structure problem. In the plane-wave dual basis, the electronic-structure Hamiltonian has the form [see Eq. (8) of Ref. [11]]

$$\begin{aligned} H = & \underbrace{\frac{1}{2n} \sum_{j,k,\nu} \kappa_\nu^2 \cos[\kappa_\nu \cdot r_{k-j}] A_j^\dagger A_k}_{T} \\ & \underbrace{- \frac{4\pi}{\omega} \sum_{j,\iota,\nu \neq 0} \frac{\zeta_\iota \cos[\kappa_\nu \cdot (\tilde{r}_\iota - r_j)]}{\kappa_\nu^2} N_j}_{U} \\ & \underbrace{+ \frac{2\pi}{\omega} \sum_{\substack{j \neq k \\ \nu \neq 0}} \frac{\cos[\kappa_\nu \cdot r_{j-k}]}{\kappa_\nu^2} N_j N_k}_{V}, \end{aligned} \quad (46)$$

where j, k range over all n orbitals and ω is the volume of the computational cell. Following the assumptions of Refs. [11,25], we consider the constant density case where $n/\omega = \mathcal{O}(1)$. Here, $\kappa_\nu = 2\pi\nu/\omega^{1/3}$ are n vectors of plane-wave frequencies, where ν are three-dimensional vectors of integers with elements in $[-n^{1/3}, n^{1/3}]$, r_j are the positions of electrons, ζ_ι are nuclear charges such that $\sum_\iota |\zeta_\iota| = \mathcal{O}(n)$, and $\tilde{r}_\iota$ are the nuclear coordinates. The operators $A_j^\dagger$ and A_k are electronic creation and annihilation operators, and $N_j = A_j^\dagger A_j$ are the number operators. The potential terms U and V are already diagonalized in the plane-wave dual basis. To further diagonalize the kinetic term T, we may switch to the plane-wave basis, which is accomplished by the fermionic fast Fourier transform (FFFT) [see Eq. (10) in Ref. [11]]. We have

$$H = \mathrm{FFFT}^\dagger \underbrace{\left(\frac{1}{2}\sum_\nu \kappa_\nu^2 N_\nu\right)}_{\tilde{T}} \mathrm{FFFT} + U + V. \quad (47)$$

To simulate the dynamics of such a Hamiltonian for time t, the current fastest algorithms are qubitization [24,60] with $\tilde{\mathcal{O}}(n^3 t)$ gate complexity and a small prefactor, and the interaction-picture algorithm [25] with complexity $\tilde{\mathcal{O}}(n^2 t)$ and a large prefactor. We show that higher-order product formulas can perform the same simulation with gate complexity $n^{2+o(1)}t^{1+o(1)}$. For the special case of the second-order Suzuki formula, this confirms a recent observation of Kivlichan *et al.* from numerical calculation [43].

Using the plane-wave basis for the kinetic operator and the plane-wave dual basis for the potential operators, we find that all terms in $\tilde{T}$ and $U+V$ commute with each other, respectively. Then, we can decompose $e^{-it\tilde{T}}$ and $e^{-it(U+V)}$ into a product of elementary matrix exponentials without introducing additional error, giving the product formula

$$\begin{aligned} &e^{-ita_{(\Upsilon,2)}T} e^{-ita_{(\Upsilon,1)}(U+V)} \cdots e^{-ita_{(1,2)}T} e^{-ita_{(1,1)}(U+V)} \\ &\quad = \mathrm{FFFT}^\dagger e^{-ita_{(\Upsilon,2)}\tilde{T}} \mathrm{FFFT} e^{-ita_{(\Upsilon,1)}(U+V)} \\ &\qquad \cdots \mathrm{FFFT}^\dagger e^{-ita_{(1,2)}\tilde{T}} \mathrm{FFFT} e^{-ita_{(1,1)}(U+V)}. \end{aligned} \quad (48)$$

For practical implementation, we need to further exponentiate spin operators using a fermionic encoding, such as the Jordan-Wigner encoding. However, these implementation details do not affect the analysis of the Trotter error and will thus be ignored in our discussion. The fermionic fast Fourier transform and the exponentiation of $\tilde{T}$, U, and V can all be implemented using the Jordan-Wigner encoding with complexity $\tilde{O}(n)$ [25,61].

We compute the norm of $[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\cdots]$, $H_\gamma \in \{T, U, V\}$ by induction. We show in Appendix F that

$$\begin{aligned} \tilde{\alpha}_{\mathrm{comm}} &= \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\cdots]\| \\ &= \mathcal{O}(n^{p+1}). \end{aligned} \quad (49)$$

Theorem 6 and Corollary 7 then imply that a Trotter number of $r = \mathcal{O}((nt)^{1+1/p}/\epsilon^{1/p})$ suffices to simulate with accuracy ϵ. Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step as in Refs. [25,61], we have the gate complexity

$$n^{2+o(1)}t^{1+o(1)} \quad (50)$$

for simulating plane-wave electronic structure in the second quantization.

k-local Hamiltonians. A Hamiltonian is k-local if it can be expressed as a linear combination of terms, each of which acts nontrivially on at most $k = \mathcal{O}(1)$ qubits. Such Hamiltonians, especially 2-local ones, are ubiquitous in physics. The first explicit quantum simulation algorithm by Lloyd was specifically developed for simulating k-local Hamiltonians [2], and later work provided more advanced approaches based on the linear-combination-of-unitary technique [5–7,23–25]. Here, we give an improved product-formula algorithm that can be advantageous over previous simulation methods.

We consider a k-local Hamiltonian acting on n qubits,

$$H = \sum_{j_1,\ldots,j_k} H_{j_1,\ldots,j_k}, \quad (51)$$

where each $H_{j_1,\ldots,j_k}$ acts nontrivially only on qubits $j_1, \ldots, j_k$. We say $H_{j_1,\ldots,j_k}$ has support $\{j_1, \ldots, j_k\}$, denoting

$$\mathcal{S}(H_{j_1,\ldots,j_k}) \coloneqq \{j_1, \ldots, j_k\}. \quad (52)$$

We may assume that the summands are unitaries up to scaling and can be implemented with constant cost; otherwise, we expand them further with respect to the Pauli operators. The fastest previous approach to simulating a general k-local Hamiltonian is the qubitization algorithm by Low and Chuang [24], which has gate complexity $\tilde{\mathcal{O}}(n^k \|H\|_1 t)$ where $\|H\|_1 = \sum_{j_1,\ldots,j_k} \|H_{j_1,\ldots,j_k}\|$.

To compare with the product-formula algorithm, we need to analyze the nested commutators $[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\cdots]$, where each H_γ is some local operator $H_{j_1,\ldots,j_k}$. In order for this commutator to be nonzero, every operator must have support that overlaps with the support of operators from the inner layers. Using this idea, we estimate that

$$\begin{aligned} \tilde{\alpha}_{\mathrm{comm}} &= \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\cdots]\| \\ &= \mathcal{O}(|||H|||_1^p \|H\|_1), \end{aligned} \quad (53)$$

where $|||H|||_1$ is the induced one-norm defined in Eq. (6). Theorem 6 and Corollary 7 then imply that a Trotter number of $r = \mathcal{O}(|||H|||_1 \|H\|_1^{1/p} t^{1+1/p}/\epsilon^{1/p})$ suffices to simulate with accuracy ϵ. Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step with $\Theta(n^k)$ gates, we have the total gate complexity

$$n^k |||H|||_1 \|H\|_1^{o(1)} t^{1+o(1)} \quad (54)$$

for simulating a k-local Hamiltonian H. See Appendix G for more details.

We know from Sec. II A that the norm inequality $|||H|||_1 \leq \|H\|_1$ always holds. In fact, the gap between these two norms can be significant for many k-local Hamiltonians. As an example, we consider n-qubit power-law interactions $H = \sum_{\vec{i},\vec{j}\in\Lambda} H_{\vec{i},\vec{j}}$ with exponent

α [52], where $\Lambda \subseteq \mathbb{R}^d$ is a d-dimensional square lattice, $H_{\vec{i},\vec{j}}$ is an operator supported on two sites $\vec{i}, \vec{j} \in \Lambda$, and

$$\|H_{\vec{i},\vec{j}}\| \leq \begin{cases} 1 & \text{if } \vec{i} = \vec{j} \\ \frac{1}{\|\vec{i}-\vec{j}\|_2^\alpha} & \text{if } \vec{i} \neq \vec{j}. \end{cases} \tag{55}$$

Examples of such systems include those that interact via the Coulomb interactions ($\alpha = 1$), the dipole-dipole interactions ($\alpha = 3$), and the van der Waals interactions ($\alpha = 6$). It is straightforward to upper bound the induced 1-norm,

$$|||H|||_1 = \begin{cases} \mathcal{O}(n^{1-\alpha/d}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(\log n) & \text{for } \alpha = d \\ \mathcal{O}(1) & \text{for } \alpha > d, \end{cases} \tag{56}$$

whereas the 1-norm scales like

$$\|H\|_1 = \begin{cases} \mathcal{O}(n^{2-\alpha/d}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(n \log n) & \text{for } \alpha = d \\ \mathcal{O}(n) & \text{for } \alpha > d. \end{cases} \tag{57}$$

Thus, the product-formula algorithm has gate complexity

$$g_\alpha = \begin{cases} n^{3-\frac{\alpha}{d}+o(1)} t^{1+o(1)} & \text{for } 0 \leq \alpha < d \\ n^{2+o(1)} t^{1+o(1)} & \text{for } \alpha \geq d, \end{cases} \tag{58}$$

which has better n dependence than the qubitization approach [24]. We give further calculation details in Appendix H.

Rapidly decaying power-law and quasilocal interactions. We now consider d-dimensional power-law interactions $1/x^\alpha$ with exponent $\alpha > 2d$ and interactions that decay exponentially with distance. Although these Hamiltonians can be simulated using algorithms for k-local Hamiltonians, more efficient methods exist that exploit the locality of the systems [52]. We show that product formulas can also leverage locality to provide an even faster simulation.

We first consider an n-qubit d-dimensional power-law Hamiltonian $H = \sum_{i,j\in\Lambda} H_{\vec{i},\vec{j}}$ with exponent $\alpha > 2d$. Such a Hamiltonian represents a rapidly decaying long-range system that becomes nearest-neighbor interacting in the limit $\alpha \to \infty$. For $\alpha > 2d$, the state-of-the-art simulation algorithm decomposes the evolution based on the Lieb-Robinson bound with gate complexity $\tilde{\mathcal{O}}((nt)^{1+2d/(\alpha-d)})$ [52]. We give an improved approach using product formulas which has gate complexity $(nt)^{1+d/(\alpha-d)+o(1)}$.

The idea of our approach is to simulate a truncated Hamiltonian $\tilde{H} = \sum_{\|\vec{i}-\vec{j}\|_2 \leq \ell} H_{\vec{i},\vec{j}}$ by taking only the terms $H_{\vec{i},\vec{j}}$ where $\|\vec{i} - \vec{j}\|_2$ is not more than ℓ, a parameter that we determine later. The resulting $\tilde{H}$ is a 2-local Hamiltonian with 1-norm $\|\tilde{H}\|_1 = \mathcal{O}(n)$ and induced 1-norm $|||\tilde{H}|||_1 = \mathcal{O}(1)$. Theorem 6 and Corollary 7 then imply that a Trotter number of $r = \mathcal{O}(n^{1/p} t^{1+1/p}/\epsilon^{1/p})$ suffices to simulate with accuracy ϵ. Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step with $\mathcal{O}(n\ell^d)$ gates, we have the total gate complexity $\ell^d (nt)^{1+o(1)}$ for simulating $\tilde{H}$.

We know from Corollary A.5 that the approximation of $\exp(-iHt)$ by $\exp(-i\tilde{H}t)$ has error

$$\|e^{-itH} - e^{-it\tilde{H}}\| = \mathcal{O}(\|H - \tilde{H}\| t), \tag{59}$$

where $\|H - \tilde{H}\| = \mathcal{O}(n/\ell^{\alpha-d})$ for all $\alpha > 2d$. To make this at most $\mathcal{O}(\epsilon)$, we choose the cutoff $\ell = \Theta((nt/\epsilon)^{1/(\alpha-d)})$. Note that we require $nt \geq \epsilon$ and $t \leq \epsilon n^{\alpha/d-2}$ so that $n^{1/d} \geq \ell \geq 1$. This choice implies the gate complexity

$$(nt)^{1+d/(\alpha-d)+o(1)}, \tag{60}$$

which is better than the state-of-the-art algorithm based on Lieb-Robinson bounds [52]. We leave the calculation details to Appendix H.

We also consider interactions that decay exponentially with the distance x as $e^{-\beta x}$:

$$\|H_{\vec{i},\vec{j}}\| \leq e^{-\beta\|\vec{i}-\vec{j}\|_2}, \tag{61}$$

where $\beta > 0$ is a constant. Although such interactions are technically long range, their fast decay makes them quasilocal for most applications in physics. Our approach to simulating such a quasilocal system is similar to that for the rapidly decaying power-law Hamiltonian, except we choose the cutoff $\ell = \Theta(\log(nt/\epsilon))$, giving a product-formula algorithm with gate complexity

$$(nt)^{1+o(1)}. \tag{62}$$

See Appendix H for further details.

Our result for quasilocal systems is asymptotically the same as a recent result for nearest-neighbor Hamiltonians [35]. For rapidly decaying power-law systems, we reproduce the nearest-neighbor case [35] in the limit $\alpha \to \infty$.

Clustered Hamiltonians. We now consider the application of our theory to simulating clustered Hamiltonians [53]. Such systems appear naturally in the study of classical fragmentation methods and quantum mechanics or molecular mechanics methods for simulating large molecules. Peng, Harrow, Ozols, and Wu recently proposed a hybrid simulator for clustered Hamiltonians [53]. Here, we show that the performance of their simulator can be significantly improved using our Trotter error bound.

Let H be a Hamiltonian acting on n qubits. Following the same setting as in Ref. [53], we assume that each term in H acts on at most two qubits with spectral norm of at most one, and each qubit is interacted with at most a constant

number d' of other qubits. We further assume that the qubits are grouped into multiple parties and write

$$H = A + B = \sum_l H_l^{(1)} + \sum_l H_l^{(2)}, \tag{63}$$

where terms in A act on qubits within a single party, terms in B act between two different parties, and $\|H_l^{(1)}\|, \|H_l^{(2)}\| \le 1$ for all l.

The key step in the approach of Peng *et al.* is to group the terms within each party in A and simulate the resulting Hamiltonian. This step is accomplished by applying product formulas to the decomposition

$$H = A + \sum_l H_l^{(2)}. \tag{64}$$

Using the first-order Lie-Trotter formula, Ref. [53] chooses the Trotter number

$$r = \mathcal{O}\left(\frac{h_B^2 t^2}{\epsilon}\right) \tag{65}$$

to ensure that the error of the decomposition is at most ϵ, where $h_B = \sum_l \|H_l^{(2)}\|$ is the interaction strength. Here, we use Theorem 6 and Corollary 7 to show that it suffices to take

$$r = \mathcal{O}\left(\frac{d'^{(1+p/2)} h_B^{(1/p)} t^{1+(1/p)}}{\epsilon^{(1/p)}}\right) = \mathcal{O}\left(\frac{h_B^{1/p} t^{1+1/p}}{\epsilon^{1/p}}\right) \tag{66}$$

using a pth-order product formula

$$\mathcal{S}(t) = e^{-ita_\Upsilon A} \prod_l e^{-ita_{(\Upsilon,l)} H_l^{(2)}} \cdots e^{-ita_1 A} \prod_l e^{-ita_{(1,l)} H_l^{(2)}}. \tag{67}$$

This analysis improves the analysis of Ref. [53] for the first-order formula and extends the result to higher-order cases. Details can be found in Appendix I.

The hybrid simulator of Ref. [53] has runtime $2^{\mathcal{O}(r \cdot \mathrm{cc}(g))}$, where r is the Trotter number and $\mathrm{cc}(g)$ is the contraction complexity of the interaction graph g between the parties. Our improved choice of r thus provides a dramatic improvement.

B. Applications to simulating local observables

In this section, we consider quantum simulation of local observables. Our goal is to simulate the time evolution $\mathcal{A}(t) := e^{itH} A e^{-itH}$ of an observable A, where the support $\mathcal{S}(A)$ can be enclosed in a d-ball of constant radius x_0 on a d-dimensional lattice $\Lambda \subseteq \mathbb{R}^d$. Throughout this section, we consider power-law interactions with exponent $\alpha > 2d$, and we assume $t \ge 0$.

Although a local observable can be simulated by simulating the full dynamics as in Sec. IV A, this is not the most efficient approach. Instead, we use product formulas to give an algorithm whose gate complexity is independent of the system size for a short-time evolution; this complexity is much smaller than the cost of full simulation. As a by-product, we prove a Lieb-Robinson-type bound for power-law Hamiltonians that nearly matches a recent bound of Tran *et al.* [52].

Locality of time-evolved observables. Our approach is to approximate the evolution $\mathcal{A}(t) = e^{itH} A e^{-itH}$ of the local observable A by $e^{itH_{\mathrm{lc}}} A e^{-itH_{\mathrm{lc}}}$, where H_{lc} is a Hamiltonian supported within a light cone originating from A at time 0. Although this approximation can be achieved using Lieb-Robinson bounds [52], we give a direct construction using product formulas.

Without loss of generality, we assume that the Hamiltonian H is supported on an infinite lattice [62]. The idea behind our approach is as follows. We first truncate the original Hamiltonian to obtain H_{trunc}. We group the terms of H_{trunc} into d-dimensional shells based on their distance to the observable and use a product formula $\mathcal{S}_{\mathrm{trunc}}(t)$ to approximate the evolution. Unlike in Sec. IV A, we choose a specific ordering of the summands so that the majority of the terms in $\mathcal{S}_{\mathrm{trunc}}(t)$ can be commuted through the observable to cancel their counterparts in $\mathcal{S}_{\mathrm{trunc}}^\dagger(t)$. We define the reduced product formula $\mathcal{S}_{\mathrm{reduce}}(t)$ as in Fig. 1 by collecting all the remaining terms in $\mathcal{S}_{\mathrm{trunc}}(t)$. This approach gives an accurate approximation to a short-time evolution. For larger times, we divide the evolution into r Trotter steps and apply the above approximation within each step. We reverse this procedure within the light cone to obtain $\mathcal{S}_{\mathrm{lc}}(t)$, which simulates the desired Hamiltonian H_{lc}. See Fig. 2 for a step-by-step illustration of this approach.

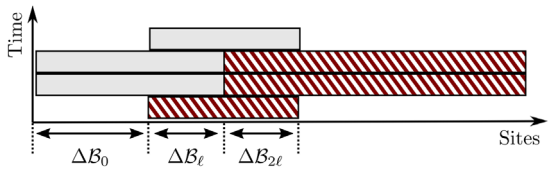

FIG. 1. Demonstration of the second-order product formula for simulating the evolution of an observable B supported on $\mathcal{S}(B) = \Delta\mathcal{B}_0$. Each rectangle represents a unitary supported on the sites covered by the width of the rectangle. The evolution unitary e^{-itH} is decomposed using the second-order product formula into $\Upsilon = 2$ stages. Each stage is a sequence of $\Gamma = 3$ matrix exponentials generated by Hamiltonian terms supported on parts of the system. Some of these unitaries (red shaded rectangles) can subsequently be commuted through B in the expression $\mathcal{S}_{\mathrm{trunc}}^\dagger(t) B \mathcal{S}_{\mathrm{trunc}}(t)$ to cancel out with their Hermitian conjugate. As a result, the time-evolved version of B can be effectively described by the remaining unitaries (light-gray rectangles).

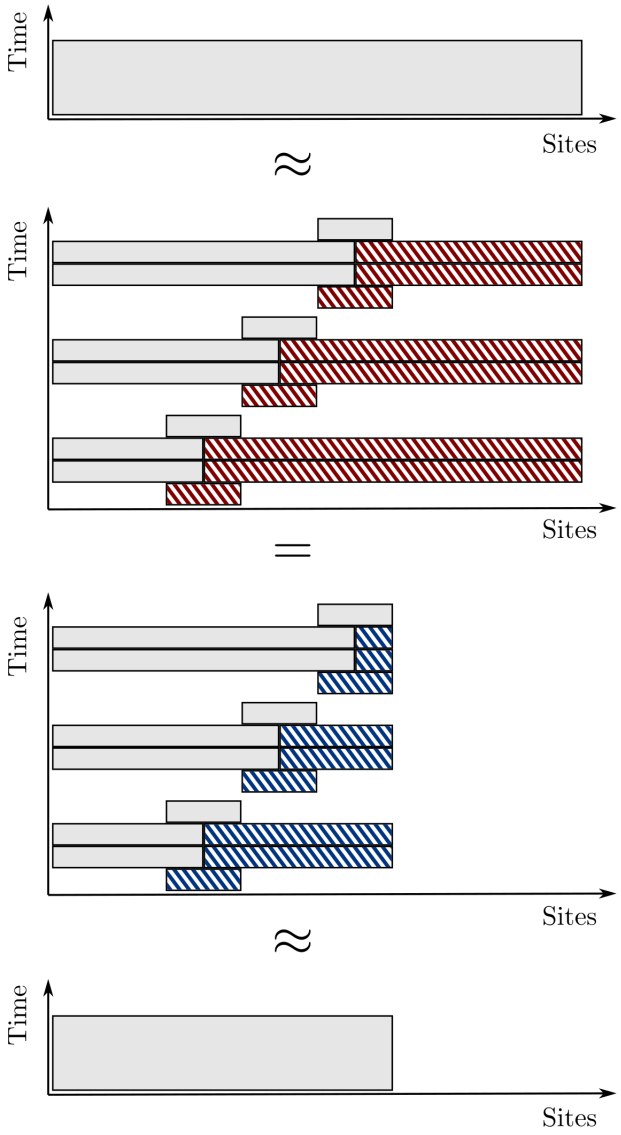

FIG. 2. Construction of the Hamiltonian H_{lc} within the light cone such that $e^{itH}Ae^{-itH} \approx e^{itH_{\mathrm{lc}}}Ae^{-itH_{\mathrm{lc}}}$. Each rectangle represents a unitary supported on the sites covered by the width of the rectangle. Specifically, the rectangle in the top panel represents the evolution e^{-itH} that we want to decompose. We divide the evolution into r steps. Within each step, we truncate the Hamiltonian to H_{trunc} and decompose its evolution using a product formula $\mathcal{S}_{\mathrm{trunc}}$. We commute certain matrix exponentials in $\mathcal{S}_{\mathrm{trunc}}$ (represented by red shaded rectangles) through the observable to cancel their counterpart, obtaining $\mathcal{S}_{\mathrm{reduce}}$ in the second panel. We reverse this procedure within the light cone to construct $\mathcal{S}_{\mathrm{lc}}$ in the third panel, which approximates $e^{-itH_{\mathrm{lc}}}$ as illustrated in the bottom panel.

We consider a general observable B, and we assume that $\mathcal{S}(B)$—the support of B—is a d-dimensional ball of radius y_0 centered on the origin. We analyze B as opposed to the original observable A so that our argument not only applies to the first Trotter step but also to later steps where A is evolved and its support is expanded. We denote by $\mathrm{dist}(\vec{i}, \mathcal{S}(B)) := \inf_{\vec{j}\in\mathcal{S}(B)} \|\vec{i}-\vec{j}\|_2$ the distance between $\vec{i}$ and $\mathcal{S}(B)$, by $\mathcal{B}_y := \{\vec{i} \in \Lambda : \mathrm{dist}(\vec{i}, \mathcal{S}(B)) \le y\}$ a ball of radius $y + y_0$ centered on $\mathcal{S}(B)$, and by $\Delta\mathcal{B}_{\gamma\ell} = \mathcal{B}_{\gamma\ell} \backslash \mathcal{B}_{(\gamma-1)\ell}$ the shell containing sites between distance $(\gamma-1)\ell$ and $\gamma\ell$ from $\mathcal{S}(B)$, where $\ell \ge 1$ is a parameter to be chosen later and $\gamma \in \mathbb{N}$ is a non-negative integer—with the convention that $\mathcal{B}_{-\ell} = \varnothing$ so that $\Delta\mathcal{B}_0 = \mathcal{B}_0 = \mathcal{S}(B)$. We illustrate the sets $\mathcal{B}_{\gamma\ell}$ and $\Delta\mathcal{B}_{\gamma\ell}$ for several values of γ in Fig. 1.

Starting from the power-law Hamiltonian $H = \sum_{\vec{i},\vec{j}\in\Lambda} H_{\vec{i},\vec{j}}$, we group terms based on their distance to the observable B and define $H_1 = \sum_{\vec{i},\vec{j}\in\mathcal{B}_\ell} H_{\vec{i},\vec{j}}$,

$$H_\gamma = \sum_{\vec{i},\vec{j}\in\Delta\mathcal{B}_{\gamma\ell}} H_{\vec{i},\vec{j}} + \sum_{\substack{\vec{i}\in\Delta\mathcal{B}_{(\gamma-1)\ell} \\ \vec{j}\in\Delta\mathcal{B}_{\gamma\ell}}} H_{\vec{i},\vec{j}}, \tag{68}$$

for $\gamma = 2, \ldots, \Gamma-1$, and $H_\Gamma = \sum_{\vec{i},\vec{j}\notin\mathcal{B}_{(\Gamma-2)\ell}} H_{\vec{i},\vec{j}}$ with constant Γ to be chosen later. In this construction, all H_γ with even γ commute, and all H_γ with odd γ commute. We consider the truncated Hamiltonian

$$H_{\mathrm{trunc}} = \sum_{\gamma=1}^{\Gamma} H_\gamma \tag{69}$$

instead of H, which incurs a truncation error of

$$\begin{aligned}\epsilon_1 &:= \|e^{-itH} - e^{-itH_{\mathrm{trunc}}}\| \\ &= \mathcal{O}(\|H - H_{\mathrm{trunc}}\|t) = \mathcal{O}\left(\frac{(y_0+\Gamma\ell)^{d-1}t}{\ell^{\alpha-d-1}}\right).\end{aligned}$$

See Appendix J for details of the proof.

Next, we simulate the evolution $e^{-itH_{\mathrm{trunc}}}$ using the pth-order product formula [see Eq. (12) and Fig. 1]:

$$\mathcal{S}_{\mathrm{trunc}}(t) = \prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma} e^{-ita_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}, \tag{70}$$

where we put additional constraints on the permutation π_ν:

$$\begin{aligned}&\pi_\upsilon(1,2,3,4,5,6,\ldots) \\ &\quad= \begin{cases} (2,4,6,\ldots,1,3,5,\ldots) & \text{if } \upsilon \text{ is odd} \\ (1,3,5,\ldots,2,4,6,\ldots) & \text{if } \upsilon \text{ is even.}\end{cases}\end{aligned} \tag{71}$$

Such a permutation can be realized using Suzuki's original construction [21] and taking into account that $[H_{2k}, H_{2k'}] = 0$ and $[H_{2k+1}, H_{2k'+1}] = 0$ for all k, k'. Using Theorem 6, we show in Appendix J that the error of approximating $e^{-itH_{\mathrm{trunc}}}$ by $\mathcal{S}_{\mathrm{trunc}}(t)$ is

$$\epsilon_2 = \|e^{-itH_{\mathrm{trunc}}} - \mathcal{S}_{\mathrm{trunc}}(t)\| = \mathcal{O}\left(\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\cdots]\| t^{p+1} \right) = \mathcal{O}((y_0 + \Gamma\ell)^{d-1}\ell t^{p+1}). \quad (72)$$

Note that the Hamiltonian terms H_γ (for $\gamma \geq 2$) commute with B. Therefore, the exponentials in $\mathcal{S}_{\mathrm{trunc}}(t)$ corresponding to these terms can be commuted through B to cancel with their counterpart in $\mathcal{S}_{\mathrm{trunc}}(t)^\dagger$. By choosing the constant $\Gamma = \Upsilon + 1$, we have

$$\mathcal{S}^\dagger_{\mathrm{trunc}}(t) B \mathcal{S}_{\mathrm{trunc}}(t) = \mathcal{S}^\dagger_{\mathrm{reduce}}(t) B \mathcal{S}_{\mathrm{reduce}}(t), \quad (73)$$

where

$$\mathcal{S}_{\mathrm{reduce}}(t) = \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\upsilon} e^{-ita_{(\upsilon,\pi_\upsilon^{-1}(\gamma))} H_\gamma}. \quad (74)$$

We call $\mathcal{S}_{\mathrm{reduce}}(t)$ the reduced product formula. This formula approximates the evolution e^{-itH} of local observable B with error

$$\begin{aligned} \|e^{itH} B e^{-itH} - \mathcal{S}^\dagger_{\mathrm{reduce}}(t) B \mathcal{S}_{\mathrm{reduce}}(t)\| &\leq \|e^{itH} B e^{-itH} - e^{itH_{\mathrm{trunc}}} B e^{-itH_{\mathrm{trunc}}}\| + \|e^{itH_{\mathrm{trunc}}} B e^{-itH_{\mathrm{trunc}}} - \mathcal{S}^\dagger_{\mathrm{trunc}}(t) B \mathcal{S}_{\mathrm{trunc}}(t)\| \\ &\quad + \|\mathcal{S}^\dagger_{\mathrm{trunc}}(t) B \mathcal{S}_{\mathrm{trunc}}(t) - \mathcal{S}^\dagger_{\mathrm{reduce}}(t) B \mathcal{S}_{\mathrm{reduce}}(t)\| \\ &\leq 2\|B\|(\epsilon_1 + \epsilon_2) + 0 \\ &= \mathcal{O}\left(\|B\| t (y_0 + \Gamma\ell)^{d-1} \left(\frac{1}{\ell^{\alpha-d-1}} + \ell t^p \right) \right). \end{aligned}$$

The above decomposition is accurate for a short-time evolution. For larger times, we divide the simulation into r Trotter steps and apply this decomposition within each step. We analyze the error in a similar way as above, except that B is defined by applying the reduced product formula to the observable A. Since the spectral norm is invariant under unitary transformations, we have $\|B\| = \|A\| = \mathcal{O}(1)$. Another difference is that the support of the observable is expanded by $\Gamma\ell$ after each Trotter step; i.e., we set y_0 to be x_0, $x_0 + \Gamma\ell$,…, and $x_0 + r\Gamma\ell$. Using the triangle inequality, we bound the error of the reduced product formula by

$$\mathcal{O}\left(t(x_0 + r\Gamma\ell)^{d-1} \left(\frac{1}{\ell^{\alpha-d-1}} + \ell \frac{t^p}{r^p} \right) \right). \quad (75)$$

We now apply the above procedure in the reverse direction but only to Hamiltonian terms within the light cone, incurring a truncation error of at most ϵ_1 and a Trotter error of at most ϵ_2. This method replaces $\mathcal{S}_{\mathrm{reduce}}(t)$ by $\mathcal{S}_{\mathrm{lc}}(t)$, the product formula that simulates the Hamiltonian H_{lc} whose terms have distance of at most $r\Gamma\ell$ to the local observable A. See Fig. 2 for a step-by-step illustration of this approach. We analyze the error in a similar way as above, establishing the following result on evolving local observables.

Proposition 8: (Product-formula decomposition of evolutions of local observables). Let $\Lambda \subseteq \mathbb{R}^d$ be a d-dimensional square lattice. Let H be a power-law Hamiltonian (55) with exponent $\alpha > 2d$, and let A be an observable with support enclosed in a d-dimensional ball of constant radius x_0. Construct the Hamiltonian H_{lc} as above using pth-order Υ-stage product formulas $\mathcal{S}_{\mathrm{trunc}}(t)$, $\mathcal{S}_{\mathrm{reduce}}(t)$, and $\mathcal{S}_{\mathrm{lc}}(t)$. Then, the support of H_{lc} has radius $x_0 + r\Gamma\ell$ and

$$\|e^{itH} A e^{-itH} - e^{itH_{\mathrm{lc}}} A e^{-itH_{\mathrm{lc}}}\| = \mathcal{O}\left(t(x_0 + r\Gamma\ell)^{d-1} \left(\frac{1}{\ell^{\alpha-d-1}} + \ell \frac{t^p}{r^p} \right) \right), \quad (76)$$

where the positive integer ℓ is a parameter and $\Gamma = \Upsilon + 1$ is constant.

Gate complexity of simulating local observables. We now analyze the gate complexity of simulating local observables using the decomposition in Proposition 8. Assuming the support $\mathcal{S}(A)$ has constant radius $x_0 = \mathcal{O}(1)$ and $\Gamma = \mathcal{O}(1)$, we simplify the error bound in Eq. (76) to

$$\|e^{itH} A e^{-itH} - e^{itH_{\mathrm{lc}}} A e^{-itH_{\mathrm{lc}}}\| = \mathcal{O}\left(t(r\ell)^{d-1} \left(\frac{1}{\ell^{\alpha-d-1}} + \frac{\ell t^p}{r^p} \right) \right). \quad (77)$$

To minimize the error, we choose the cutoff $\ell = \Theta((r/t)^{p/(\alpha-d)}) \geq 1$, which is larger than 1 provided $r \geq t$ (and recall that we assume $\alpha > 2d$, so, in particular, $\alpha > d$). With this choice of ℓ, the error becomes

$$\mathcal{O}\left(t r^{d-1} \left(\frac{t}{r} \right)^{\frac{p(\alpha-2d)}{\alpha-d}} \right) = \mathcal{O}\left(\frac{t^{\frac{p(\alpha-2d)+\alpha-d}{\alpha-d}}}{r^{\frac{p(\alpha-2d)-(\alpha-d)(d-1)}{\alpha-d}}} \right).$$

We then choose an appropriate Trotter number r as detailed in Appendix J and find that

$$g_\alpha = t^{(1+d\frac{\alpha-d}{\alpha-2d})(1+\frac{d}{\alpha-d})+o(1)} \quad (78)$$

suffices to simulate a local observable with constant accuracy. The gate count is independent of the system size and thus less than the cost of simulating the full dynamics (58) when the system size is $n = \Omega(t^{d(\alpha-d)/(\alpha-2d)})$. However, in contrast to the simulation of e^{-itH} where the asymptotic error scaling is robust against the reordering of Hamiltonian terms, we obtain a smaller error for simulating $\mathscr{A}(t)$ by defining product formulas with a special ordering that preserves the locality of the simulated system.

Additionally, in the limit $\alpha \to \infty$, which corresponds to nearest-neighbor interactions, we have the gate count

$$g_\infty = t^{d+1+o(1)}. \tag{79}$$

This result has a clear physical intuition: It is (nearly) proportional to the space-time volume t^{d+1} inside a linear light cone generated by the evolution.

Lieb-Robinson-type bound for power-law Hamiltonians. The Lieb-Robinson bounds—first derived for nearest-neighbor interactions [63] and subsequently generalized to power-law systems [52,64–70]—have found numerous applications in physics, including designing new algorithms for quantum simulations [52,71]. They bound the speed at which a local disturbance spreads in quantum systems. Here, we show that the decomposition of Proposition 9 constructed using product formulas also implies a Lieb-Robinson-type bound for power-law Hamiltonians.

The subject of the Lieb-Robinson bounds is usually the commutator norm

$$\mathcal{C}(t,\rho) = \|[e^{itH}Ae^{-itH}, B]\|, \tag{80}$$

where A, B are two operators whose supports have distance

$$\operatorname{dist}(\mathcal{S}(A),\mathcal{S}(B)) = \inf_{\vec{i}\in\mathcal{S}(A),\vec{j}\in\mathcal{S}(B)} \|\vec{i}-\vec{j}\|_2 = \rho, \tag{81}$$

and e^{-itH} is the time evolution unitary generated by a power-law Hamiltonian H. Our above discussion shows that $e^{itH}Ae^{-itH}$ is approximately $e^{itH_{\mathrm{lc}}}Ae^{-itH_{\mathrm{lc}}}$, which is supported on a ball of radius $x = \mathcal{O}(r\ell) = \mathcal{O}(r(r/t)^{p/(\alpha-d)})$ centered on $\mathcal{S}(A)$. By choosing $r = \Theta(\rho^{(\alpha-d)/(\alpha-d+p)}t^{p/(\alpha-d+p)})$ so that $x < \rho$, we make $e^{itH_{\mathrm{lc}}}Ae^{-itH_{\mathrm{lc}}}$ commute with B, and therefore $\mathcal{C}(t,\rho)$ is small. More precisely,

$$\begin{aligned}\mathcal{C}(t,r) &= \mathcal{O}\left(\frac{t^{\frac{p(\alpha-2d)+\alpha-d}{\alpha-d}}}{r^{\frac{p(\alpha-2d)-(\alpha-d)(d-1)}{\alpha-d}}}\right)\\ &= \mathcal{O}\left(\frac{t^{\frac{(p+1)(\alpha-d)}{\alpha-d+p}}}{\rho^{\frac{p(\alpha-2d)-(\alpha-d)(d-1)}{\alpha-d+p}}}\right) = \frac{t^{\alpha-d+o(1)}}{\rho^{\alpha-2d+o(1)}}.\end{aligned}$$

Note that we have implicitly assumed that $\rho^{(\alpha-d)/(\alpha-d+p)}t^{p/(\alpha-d+p)} \geq 1$ so that we can choose $r \geq 1$. The bound implies a light cone $t \gtrsim \rho^{(\alpha-2d)/(\alpha-d)+o(1)}$, which can be made arbitrarily close to the light cone $t \gtrsim \rho^{(\alpha-2d)/(\alpha-d)}$ of the recent bound in Ref. [52] for all values of d [72].

C. Applications to quantum Monte Carlo simulation

We now apply our result to improving the performance of quantum Monte Carlo simulation. Here, the goal is to approximate certain properties of the Hamiltonian, such as the partition function, rather than simulating the full dynamics. We consider two specific systems: the transverse field Ising model of Ref. [38] and the ferromagnetic quantum spin systems of Ref. [39]. For both simulations, the ideal evolution is decomposed using the second-order Suzuki formula, and we show that such a decomposition can be made more efficient using our tightened analysis.

Transverse field Ising model. Consider the n-qubit transverse field Ising model $H = -A - B$, where

$$A = \sum_{1\leq u<v\leq n} j_{u,v}Z_uZ_v, \qquad B = \sum_{1\leq u\leq n} h_uX_u. \tag{82}$$

Here, X_u and Z_u are Pauli operators acting on the uth qubit, and $j_{u,v} \geq 0$ and $h_u \geq 0$ are non-negative coefficients. Define $j := \max\{j_{u,v}, h_u\}$ to be the maximum norm of the interactions. Our goal is to approximate the partition function

$$\mathcal{Z} = \operatorname{Tr}(e^{-H}), \tag{83}$$

up to a multiplicative error $0 < \epsilon < 1$.

Reference [38] solves this problem with an efficient classical algorithm. A key step in their algorithm is a decomposition of the evolution operator using the second-order Suzuki formula, so

$$\mathcal{Z}' = \operatorname{Tr}[(e^{\frac{1}{2r}A}e^{\frac{1}{r}B}e^{\frac{1}{2r}A})^r] \lessapprox (1+\epsilon)\operatorname{Tr}(e^{-H}) = (1+\epsilon)\mathcal{Z}. \tag{84}$$

However, their original analysis does not exploit the commutativity relation between A and B, and it can be improved by the techniques developed here.

Note that this is different from the usual setting of digital quantum simulation. As the matrix exponentials in the product formula are no longer unitary, we introduce an additional multiplicative factor when we apply Theorem 6. In addition, we need to estimate the multiplicative error of the Trotter decomposition as opposed to the additive error analyzed earlier for digital quantum simulation.

Let A and B be Hermitian matrices, and consider the evolution $e^{t(A+B)}$ with $t \geq 0$. Our goal is to choose r sufficiently large so that we can approximate the eigenvalues as

$$\lambda_i((e^{\frac{t}{2r}A}e^{\frac{t}{r}B}e^{\frac{t}{2r}A})^r) \approx \lambda_i(e^{t(A+B)}), \tag{85}$$

up to a small multiplicative error, where $\lambda_i(\cdot)$ denotes the ith smallest eigenvalue. We define

$$U \coloneqq e^{\frac{t}{r}(A+B)}, \tag{86}$$

$$V \coloneqq e^{\frac{t}{2r}A}e^{\frac{t}{r}B}e^{\frac{t}{2r}A}, \tag{87}$$

$$W \coloneqq \mathcal{T}\exp\left(\int_0^{\frac{t}{r}} \mathrm{d}\tau\, e^{-\tau \mathrm{ad}_{A+B}}\left(e^{\frac{\tau}{2}\mathrm{ad}_A}B - B + e^{\frac{\tau}{2}\mathrm{ad}_A}e^{\tau \mathrm{ad}_B}\frac{A}{2} - \frac{A}{2}\right)\right). \tag{88}$$

Then, both U and V are positive-semidefinite operators, and we know from Theorem 3 that $V = UW$. In Appendix K 1, we show that if r is a power of 2,

$$\lambda_i(V^r) \le \lambda_i(U^r)\|W\|^r, \tag{89}$$

where

$$\|W\|^r \le \exp\left(\left(\frac{t^3}{24r^2}\|\mathrm{ad}_A^2 B\| + \frac{t^3}{4r^2}\|\mathrm{ad}_B^2 A\|\right)e^{\frac{4t}{r}(\|A\|+\|B\|)}\right). \tag{90}$$

Equation (89) says that the eigenvalues of V^r are upper bounded by the eigenvalues of U^r multiplied by a factor of $\|W\|^r$. To ensure that the factor is close to 1, we require

$$r \ge \max\left\{4t(\|A\| + \|B\|), \sqrt{\frac{t^3}{3\epsilon}\|\mathrm{ad}_A^2 B\|}, \sqrt{\frac{2t^3}{\epsilon}\|\mathrm{ad}_B^2 A\|}\right\}, \tag{91}$$

in addition to being a power of 2. The first condition in Eq. (91) is so that $e^{4(t/r)(\|A\|+\|B\|)} \le e < 4$, while the last two conditions make sure that both $(t^3/24r^2)\|\mathrm{ad}_A^2 B\|$ and $(t^3/4r^2)\|\mathrm{ad}_B^2 A\|$ are bounded by $\epsilon/8$. Therefore, we have $\|W^r\| \le e^{\epsilon}$, which implies

$$\mathcal{Z}' = \sum_i \lambda_i(V^r) \le \sum_i \lambda_i(U^r)e^{\epsilon} \approx (1+\epsilon)\sum_i \lambda_i(U^r) = (1+\epsilon)\mathcal{Z}, \tag{92}$$

assuming $\epsilon \ll 1$. Following similar arguments, we can show that this choice of r also gives a lower bound of $\mathcal{Z}'$ with $\mathcal{Z}' \ge (1-\epsilon)\mathcal{Z}$. Therefore, we have approximated the partition function up to a multiplicative error ϵ.

We now specialize our result to the transverse field Ising Hamiltonian with $t = 1$. We find that

$$\|A\| = \mathcal{O}(n^2 j), \qquad \|B\| = \mathcal{O}(nj), \qquad \|\mathrm{ad}_A^2 B\| = \mathcal{O}(n^3 j^3), \qquad \|\mathrm{ad}_B^2 A\| = \mathcal{O}(n^2 j^3), \tag{93}$$

which implies

$$r = \mathcal{O}(n^2 j + n^{3/2} j^{3/2}\epsilon^{-1/2}). \tag{94}$$

Using Ref. [38] (p. 17), this analysis gives a fully polynomial randomized approximation scheme (FPRAS) with running time

$$\tilde{\mathcal{O}}(n^{17} r^{14}\epsilon^{-2}) = \tilde{\mathcal{O}}(n^{45} j^{14}\epsilon^{-2} + n^{38} j^{21}\epsilon^{-9}), \tag{95}$$

improving over the previous complexity of

$$\tilde{\mathcal{O}}(n^{59} j^{21}\epsilon^{-9}). \tag{96}$$

Quantum ferromagnets. We now apply our technique to improve the Monte Carlo simulation of ferromagnetic quantum spin systems [39]. Such systems are described by the n-qubit Hamiltonian

$$H = \sum_{1\le u<v\le n}(-b_{uv}X_uX_v + c_{uv}Y_uY_v) + \sum_{u=1}^n d_u(I + Z_u),$$

where $0 \le b_{uv} \le 1$, $-b_{uv} \le c_{uv} \le b_{uv}$, and $-1 \le d_{uv} \le 1$. It will be convenient to rewrite these Hamiltonians using the coefficients $p_{uv} = (b_{uv} - c_{uv})/2$ and $q_{uv} = (b_{uv} + c_{uv})/2$ as

$$H = \sum_{1\le u<v\le n} p_{uv}(-X_uX_v - Y_uY_v) + \sum_{1\le u<v\le n} q_{uv}(-X_uX_v + Y_uY_v) + \sum_{u=1}^n d_u(I + Z_u).$$

Since $|c_{uv}| \le b_{uv} \le 1$, we have $p_{uv}, q_{uv} \in [0,1]$.

Our goal is to approximate the partition function

$$\mathcal{Z}(\beta, H) = \mathrm{Tr}[e^{-\beta H}], \tag{97}$$

for $\beta > 0$. Following the setting of Ref. [39], we restrict ourselves to the n-qubit matchgate set

$$\left\{f_u(e^{\pm t}), g_{uv}(t), h_{uv}(t) | u, v = 1, \ldots, n, u \ne v, 0 < t < \frac{1}{2}\right\}, \tag{98}$$

where

$$f(e^{\pm t}) = \begin{bmatrix} e^{\pm t} & 0 \\ 0 & 1 \end{bmatrix}, \qquad g(t) = \begin{bmatrix} 1+t^2 & 0 & 0 & t \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ t & 0 & 0 & 1 \end{bmatrix},$$

$$h(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1+t^2 & t & 0 \\ 0 & t & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{99}$$

and the subscripts u, v indicate the qubits on which the gates act nontrivially. The motivations for using these gates can be found in Ref. [39], and we do not repeat them here. These gates approximately implement the exponential of the Hamiltonian terms in the sense that

$$f_u(e^{\pm t}) = e^{\pm \frac{t}{2}(I+Z_u)}, \tag{100}$$

$$g_{uv}(t) = e^{-\frac{t}{2}(-X_uX_v+Y_uY_v)+O(t^2)}, \tag{101}$$

$$h_{uv}(t) = e^{-\frac{t}{2}(-X_uX_v-Y_uY_v)+O(t^2)}. \tag{102}$$

We divide the evolution into r steps and apply the second-order Suzuki formula within each step to further decompose $e^{-(\beta/r)H}$ into the elementary gates [Eq. (98)]. Here, we have two sources of error: the Trotter error and the error from using the gate set in Eq. (98). In the following analysis, we assume

$$r > 2\beta \tag{103}$$

so that we can implement the product formula using gates from Eq. (98) with parameters

$$-\frac{1}{2} < -\frac{\beta}{r}d_u < \frac{1}{2}, \quad 0 < \frac{\beta}{r}q_{uv} < \frac{1}{2}, \quad 0 < \frac{\beta}{r}p_{uv} < \frac{1}{2}. \tag{104}$$

In Appendix K 2, we use the interaction picture (Lemma A.2) to show that

$$\prod_{1\le u\le n} f_u\left(e^{-\frac{\beta}{r}d_u}\right) \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \cdot \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \cdot \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \prod_{1\le u\le n} f_u(e^{-\frac{\beta}{r}d_u}) = e^{-\frac{\beta}{r}H}U, \tag{105}$$

where the operator U has spectral norm bounded by

$$\|U\| \le \exp\left(\frac{2n^2\beta^2}{r^2}e^{\frac{4n^2\beta}{r}} + \frac{cn^4\beta^3}{r^3}e^{\frac{12n^2\beta}{r}}\right), \tag{106}$$

for some constant $c > 0$. Roughly speaking, Eq. (105) approximates the exponential $e^{-(\beta/r)H}$ by the gates in the gate set (98) up to a multiplicative factor U.

The remaining analysis proceeds in a similar way as that of the transverse field Ising model. We find that each eigenvalue of

$$\left[\prod_{1\le u\le n} f_u\left(e^{-\frac{\beta}{r}d_u}\right) \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \cdot \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \cdot \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \prod_{1\le u\le n} f_u\left(e^{-\frac{\beta}{r}d_u}\right)\right]^r$$

approximates the corresponding eigenvalue of the ideal evolution $e^{-\beta H}$ with a multiplicative factor

$$\|U\|^r \le \exp\left(\frac{2n^2\beta^2}{r}e^{\frac{4n^2\beta}{r}} + \frac{cn^4\beta^3}{r^2}e^{\frac{12n^2\beta}{r}}\right). \tag{107}$$

We first set

$$r \ge 24n^2\beta \tag{108}$$

so that

$$\|U\|^r \le \exp\left(\frac{4n^2\beta^2}{r} + \frac{2cn^4\beta^3}{r^2}\right). \tag{109}$$

We then choose

$$r \ge \max\left\{\frac{8n^2\beta^2}{\epsilon}, \frac{2\sqrt{c}n^2\beta^{3/2}}{\epsilon^{1/2}}\right\} \tag{110}$$

to ensure that the multiplicative error is at most ϵ. From Eqs. (103), (108), and (110),

$$r = \mathcal{O}\left(\frac{n^2\lceil\beta\rceil^2}{\epsilon}\right), \tag{111}$$

which gives the total gate complexity (see Supplementary p. 11 in Ref. [39])

$$j := 2n^2r = \mathcal{O}\left(\frac{(1+\beta^2)n^4}{\epsilon}\right). \tag{112}$$

The result of Theorem 2 in Ref. [39] gave a Monte Carlo simulation algorithm for the ferromagnetic quantum spin systems. To improve that result, we also need to estimate the error of partial sequence of the product formula as in Ref. [39] [Eq. (13)]. This estimation can be done in a similar way to our above analysis. The resulting randomized approximation scheme has runtime

$$\tilde{\mathcal{O}}\left(\frac{j^{23}}{\epsilon^2}\right) = \tilde{\mathcal{O}}\left(\frac{n^{92}(1+\beta^{46})}{\epsilon^{25}}\right), \tag{113}$$

which improves the runtime of the original Bravyi-Gosset algorithm

$$\tilde{\mathcal{O}}\left(\frac{n^{115}(1+\beta^{46})}{\epsilon^{25}}\right). \tag{114}$$

V. ERROR BOUNDS WITH SMALL PREFACTORS

We now derive Trotter error bounds with small prefactors. These bounds complement the above asymptotic analysis and can be used to optimize near-term implementations of quantum simulation. In Sec. V A, we show that our analysis reproduces previous tight error bounds for the first- and second-order formulas [41,42,46]. We then give numerical evidence in Sec. V B, showing that our higher-order bounds are close to tight for certain nearest-neighbor interactions and power-law Hamiltonians. Throughout this section, we let H be Hermitian and let $t \in \mathbb{R}$, and we decompose the real-time evolution e^{-itH}.

A. First- and second-order error bounds

We derive error bounds for the first-order Lie-Trotter formula and second-order Suzuki formula following the work of Refs. [40,43]. In this work, we first analyze the Trotter error of decomposing the evolution of a two-term Hamiltonian. We then bootstrap the result to analyze general Hamiltonians with an arbitrary number of operator summands. The resulting bounds are nearly tight because they match the lowest-order term of the BCH expansion up to an application of the triangle inequality [9,40–43].

Let $H = A + B$ be a two-term Hamiltonian. The evolution under H for time $t \geq 0$ is given by $e^{-itH} = e^{-it(A+B)}$, which we decompose using the first-order Lie-Trotter formula $\mathcal{S}_1(t) = e^{-itB}e^{-itA}$. We first construct the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_1(t) = -iH\mathcal{S}_1(t) + e^{-itB}(e^{it\mathrm{ad}_B}(iA) - iA)e^{-itA}, \tag{115}$$

with initial condition $\mathcal{S}_1(0) = I$. Using the variation-of-parameters formula (Lemma A.1),

$$\mathcal{S}_1(t) = e^{-itH} + \int_0^t \mathrm{d}\tau_1\, e^{-i(t-\tau_1)H} \cdot e^{-i\tau_1 B}(e^{i\tau_1\mathrm{ad}_B}(iA) - iA)e^{-i\tau_1 A}. \tag{116}$$

Using Theorem 4 or by direct calculation, we find the order condition $e^{i\tau_1\mathrm{ad}_B}(iA) - iA = O(\tau_1)$, which implies

$$e^{i\tau_1\mathrm{ad}_B}(iA) - iA = \int_0^{\tau_1} \mathrm{d}\tau_2\, e^{i\tau_2\mathrm{ad}_B}\mathrm{ad}_{iB}(iA).$$

Altogether, we have the representation

$$\mathcal{S}_1(t) = e^{-itH} + \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2\, e^{-i(t-\tau_1)H} \cdot e^{-i\tau_1 B}e^{i\tau_2\mathrm{ad}_B}\mathrm{ad}_{iB}(iA)e^{-i\tau_1 A}, \tag{117}$$

and the error bound for $t \geq 0$,

$$\|\mathcal{S}_1(t) - e^{-itH}\| \leq \frac{t^2}{2}\|[B,A]\|. \tag{118}$$

We bootstrap this bound to analyze a general Hamiltonian $H = \sum_\gamma^\Gamma H_\gamma$. By the triangle inequality,

$$\begin{aligned}
&\left\|\prod_{\gamma=1}^{\Gamma} e^{-itH_\gamma} - e^{-it\sum_{\gamma=1}^{\Gamma}H_\gamma}\right\| \\
&\leq \sum_{\gamma_1=1}^{\Gamma}\left\|e^{-it\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}}\prod_{\gamma_2=1}^{\gamma_1}e^{-itH_{\gamma_2}} - e^{-it\sum_{\gamma_2=\gamma_1}^{\Gamma}H_{\gamma_2}}\prod_{\gamma_2=1}^{\gamma_1-1}e^{-itH_{\gamma_2}}\right\| \\
&= \sum_{\gamma_1=1}^{\Gamma}\left\|e^{-it\sum_{\gamma_2=\gamma_1+}^{\Gamma}H_{\gamma_2}}e^{-itH_{\gamma_1}} - e^{-it\sum_{\gamma_2=\gamma_1}^{\Gamma}H_{\gamma_2}}\right\| \\
&\leq \frac{t^2}{2}\sum_{\gamma_1=1}^{\Gamma}\left\|\left[\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}, H_{\gamma_1}\right]\right\|.
\end{aligned} \tag{119}$$

Thus, we obtain the following proposition.

Proposition 9: (Tight error bound for the first-order Lie-Trotter formula). Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a Hamiltonian consisting of Γ summands and $t \geq 0$. Let $\mathcal{S}_1(t) = \prod_{\gamma=1}^{\Gamma} e^{-itH_\gamma}$ be the first-order Lie-Trotter formula. Then, the additive Trotter error can be bounded as

$$\|\mathcal{S}_1(t) - e^{-itH}\| \leq \frac{t^2}{2}\sum_{\gamma_1=1}^{\Gamma}\left\|\left[\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}, H_{\gamma_1}\right]\right\|. \tag{120}$$

A generalization of this analysis gives an error bound for the second-order Suzuki formula with a small prefactor (see Appendix L for a proof).

Proposition 10: (Tight error bound for the second-order Suzuki formula). Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be a Hamiltonian consisting of Γ summands and $t \geq 0$. Let $\mathcal{S}_2(t) = \prod_{\gamma=\Gamma}^{1} e^{-i(t/2)H_\gamma}\prod_{\gamma=1}^{\Gamma} e^{-i(t/2)H_\gamma}$ be the second-order Suzuki formula. Then, the additive Trotter error can be bounded as

$$\begin{aligned}
&\|\mathcal{S}_2(t) - e^{-itH}\| \\
&\leq \frac{t^3}{12}\sum_{\gamma_1=1}^{\Gamma}\left\|\left[\sum_{\gamma_3=\gamma_1+1}^{\Gamma}H_{\gamma_3}, \left[\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}, H_{\gamma_1}\right]\right]\right\| \\
&\quad + \frac{t^3}{24}\sum_{\gamma_1=1}^{\Gamma}\left\|\left[H_{\gamma_1}, \left[H_{\gamma_1}, \sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}\right]\right]\right\|.
\end{aligned} \tag{121}$$

B. Higher-order error bounds

We now consider error bounds for higher-order product formulas. Compared to the first- and second-order cases, these formulas are harder to analyze due to their more complicated definitions. Nevertheless, higher-order formulas have better asymptotic scaling and can be surprisingly efficient even for simulating small systems, as observed in Ref. [31].

We have analyzed the error of higher-order product formulas in Sec. III. That analysis is sufficient to establish the commutator scaling in Theorem 6, but the resulting bounds have large prefactors. Here, we propose heuristic strategies to tighten the analysis and numerically benchmark our bounds for nearest-neighbor and power-law Hamiltonians. Our heuristics are specified in Appendix M.

Although we do not have a rigorous proof of the tightness of our higher-order bounds, numerical evidence suggests that they are close to tight for various systems. We first consider simulating a one-dimensional Heisenberg model with a random magnetic field

$$H = \sum_{j=1}^{n-1} (\vec{\Sigma}_j \cdot \vec{\Sigma}_{j+1} + h_j Z_j), \tag{122}$$

where $\vec{\Sigma}_j \coloneqq (X_j, Y_j, Z_j)$ is the vector of Pauli operators acting on the jth qubit and the coefficients $h_j \in [-1, 1]$ are chosen uniformly at random. This system can be simulated to understand the transition between the many-body localized phase and the thermalized phase in condensed matter physics, although a classical simulation is only feasible when the system size is small [74].

We classify the summands of the Hamiltonian into two groups and set

$$A = \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} (\vec{\Sigma}_{2j-1} \cdot \vec{\Sigma}_{2j} + h_{2j-1} Z_{2j-1}), \tag{123}$$

$$B = \sum_{j=1}^{\lceil \frac{n}{2} \rceil - 1} (\vec{\Sigma}_{2j} \cdot \vec{\Sigma}_{2j+1} + h_{2j} Z_{2j}). \tag{124}$$

Here, all the summands in A (and B) commute with each other, so we can further decompose exponentials like $e^{-ita_k A}$ (and $e^{-itb_k B}$) without introducing error, giving a product formula with summands ordered in an *even-odd* pattern [35]. We also consider grouping Hamiltonian summands as

$$H_1 = \sum_{j=1}^{n-1} X_j X_{j+1}, \qquad H_2 = \sum_{j=1}^{n-1} Y_j Y_{j+1},$$
$$H_3 = \sum_{j=1}^{n-1} (Z_j Z_{j+1} + h_j Z_j), \tag{125}$$

which we call the *X-Y-Z* ordering [31]. Similar to the even-odd ordering, the summands in H_1, H_2, and H_3 commute with each other, respectively, so the corresponding exponentials can also be decomposed without error. Note that our asymptotic bounds in Theorem 6 and Corollary 7 hold irrespective of the ordering of Hamiltonian summands, but the prefactors will depend on the choice of ordering. Our choice here maximizes the commutativity of the Hamiltonian.

Up to a difference on the boundary condition, Ref. [31] estimates the resource requirements of simulating the Heisenberg model using various quantum algorithms. They find that product formulas, especially the fourth-order and the sixth-order formulas, can outperform more recent quantum algorithms for simulating small instances of Eq. (122), although their best Trotter error bound is loose by several orders of magnitude. This looseness is alleviated in Ref. [35], which gives a fourth-order bound that overestimates the gate complexity by about a factor of 17. For a fair comparison, we numerically implement our approach to analyze the fourth-order formula $\mathcal{S}_4(t)$ as well; see Appendix M for detailed derivations.

For the even-odd ordering, we need to compute all the nested commutators of A and B. We do this by fixing one term $\vec{\Sigma}_{2j-1} \cdot \vec{\Sigma}_{2j} + h_{2j-1} Z_{2j-1}$ of A in the innermost layer and simplifying all the outer terms using geometrical locality. We then apply the triangle inequality to analyze the summation of terms over $j = 1, \ldots, \lfloor n/2 \rfloor$. We use a similar approach to analyze the *X-Y-Z* ordering, thus computing our error bounds for small t. To simulate for a longer time, we divide the evolution into r Trotter steps and apply our bounds within each step. We seek the smallest Trotter number r for which the estimated error is at most some desired ϵ. This number can be efficiently computed using a binary search as described in Ref. [31].

We compare our improved analysis with the best previous bounds [31,35] for simulating the Heisenberg model (122). Specifically, we consider the so-called analytic bound (see Proposition F.4 in Ref. [31], which applies to both the even-odd and the *X-Y-Z* ordering. The commutator bound of Theorem F.11 of Ref. [31] offers a slight improvement over the analytic bound; however, its numerical implementation requires extensive classical computations, so we only compare the existing result for the *X-Y-Z* ordering. Likewise, we compare the locality-based bound of Supplementary Material IV B in Ref. [35] only for the even-odd case, although it can exploit the geometrical locality of the *X-Y-Z* ordering as well.

To understand how tight our bounds are, we also include the empirical Trotter number by directly computing the error $\|(\mathcal{S}_4(t/r))^r - e^{-itH}\|$ for $n = 4, \ldots, 12$ and extrapolating the results to larger systems. We choose the evolution time $t = n$ and set the simulation accuracy $\epsilon = 10^{-3}$ as in Refs. [31,35]. For each system size, we generate five instances of Hamiltonians with random coefficients. Our results are plotted in Fig. 3.

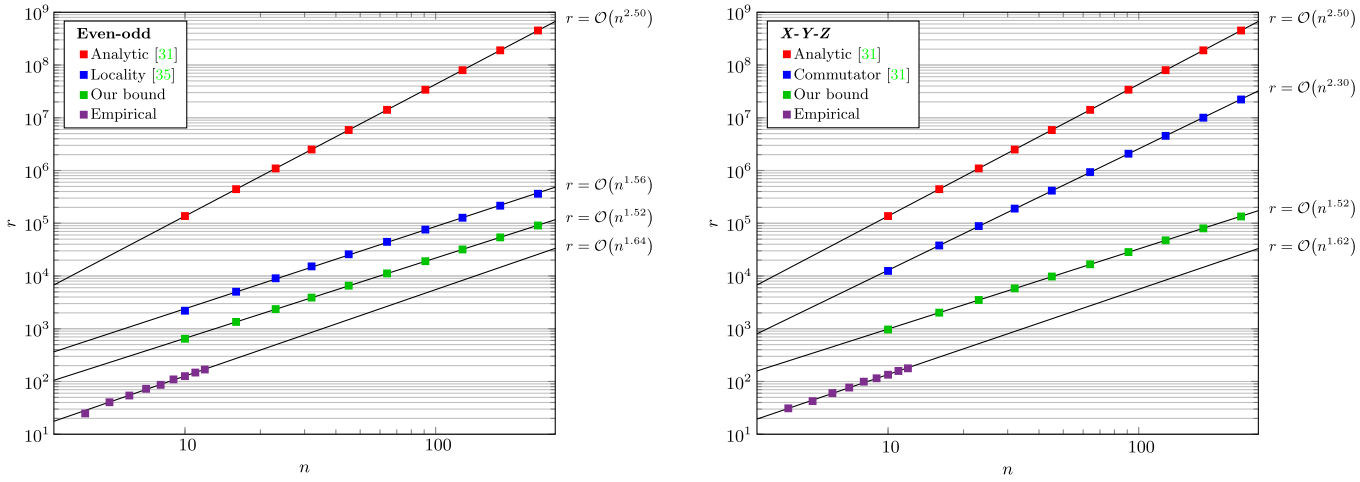

FIG. 3. Comparison of r for the Heisenberg model using the analytic bound (see Proposition F.4 in Ref. [31]), the commutator bound (see Theorem F.11 in Ref. [31]), locality-based bound (see Supplementary Material IV B in Ref. [35]), and our new bounds, Proposition M.1 and Proposition M.2. Error bars are omitted as they are negligibly small on the plot. Straight lines show power-law fits to the data. Note that the exponent for the empirical data is based on brute-force simulations of small systems and thus may not precisely capture the true asymptotic scaling due to finite-size effects.

We find that the asymptotic scaling of our new bounds matches that of the empirical result up to finite-size effects, and the prefactors are significantly tightened. At $n = 10$, the Trotter number predicted by our bounds is loose only by a factor of 5.1 for the even-odd ordering of terms and 7.2 for the X-Y-Z ordering. In comparison, the commutator bound of Ref. [31] only exploits the commutativity of the lowest-order term of the BCH series and is bottlenecked by the use of tail bounds. The previous bound [35] based on geometrical locality is also uncompetitive since it cannot directly leverage the nested commutators of the Hamiltonian terms.

In addition to nearest-neighbor interactions, we also consider the simulation of a one-dimensional Heisenberg with power-law interactions:

$$H = \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \frac{1}{|j-k|^{\alpha}} \vec{\Sigma}_j \cdot \vec{\Sigma}_k + \sum_{j=1}^{n-1} h_j Z_j, \tag{126}$$

where h_j are again chosen randomly in $[-1, 1]$ and $\alpha \geq 0$ is a constant. Similarly to the case of nearest-neighbor interactions, we use the fourth-order product formula with X-Y-Z ordering,

$$\begin{aligned} H_1 &= \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \frac{1}{|j-k|^{\alpha}} X_j X_k, \\ H_2 &= \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \frac{1}{|j-k|^{\alpha}} Y_j Y_k, \\ H_3 &= \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \left(\frac{1}{|j-k|^{\alpha}} Z_j Z_k + h_j Z_j \right), \end{aligned} \tag{127}$$

and compare the empirical Trotter number against that predicted by the best previous bound (Lemma 1) and our new bound for simulating (126) for time $t = n$ with accuracy $\epsilon = 10^{-3}$. We consider different values of α, ranging from $\alpha = 0$ (strong power-law interactions) to $\alpha = 4$ (rapidly decaying power-law interactions). We note that for $\alpha > 2$, we simulate the evolution using the truncation algorithm [with the asymptotic gate complexity given by Eq. (60)], whereas for $\alpha \leq 2$, we simply simulate the entire Hamiltonian [with the asymptotic gate complexity in Eq. (58)].

At $n = 10$, the empirical Trotter numbers are 552 ± 45 ($\alpha = 0$) and 129 ± 6 ($\alpha = 4$), where the standard deviation is obtained by averaging over five instances of the random field h_j. Meanwhile, our bound gives 5609 ± 3 ($\alpha = 0$) and 885 ± 32 ($\alpha = 4$)—about 10.2 and 6.9 times looser compared to the empirical values, respectively. Our bound for power-law interactions with small α performs slightly worse than for the nearest-neighbor interactions, partly due to the fact that the triangle inequality is invoked more often for the power-law interactions.

We note that the number of interaction terms in a long-range interacting Hamiltonian scales as n^2, making it difficult to compute our bound exactly at large n. Instead, we further upper bound the norm of the nested commutator using triangle inequalities and estimate this upper bound using a counting argument similarly to Eq. (58). In Fig. 4, we plot the empirical Trotter numbers against this counting bound for different values of n at $\alpha = 0$ and $\alpha = 4$. The figure shows that even our counting bound is tighter than the previous estimates at both values of α. We leave a thorough study of the efficient numerical implementation of our bound as a subject for future work.

In addition, we plot in Fig. 5 the scaling exponents of the Trotter numbers as functions of n at different values of

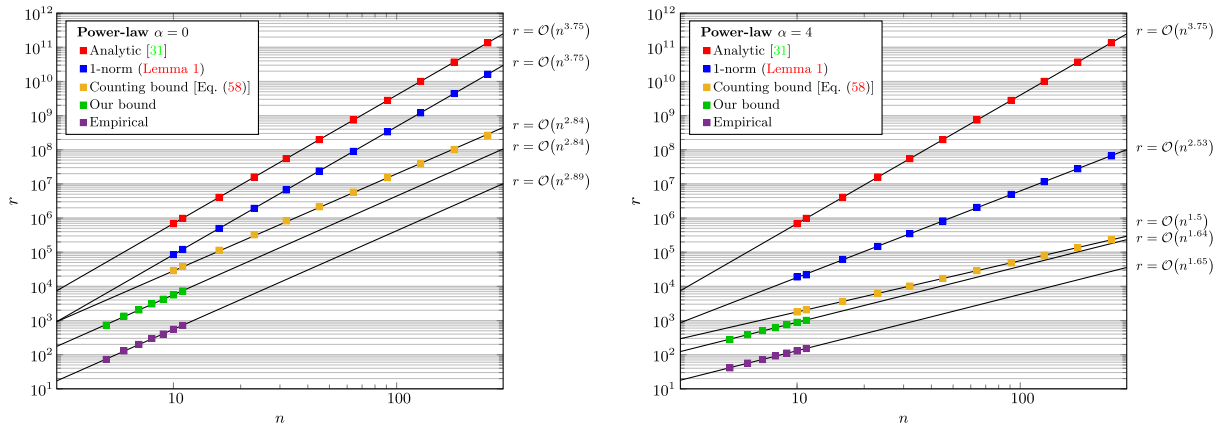

FIG. 4. Comparison of r for the power-law Heisenberg model using the analytic bound (see Proposition F.4 of Ref. [31]), 1-norm bound (Lemma 1), a bound from counting argument (58), and our bound (Proposition M.2). Error bars are omitted as they are negligibly small on the plot. Straight lines show power-law fits to the data. Note that the exponent for the empirical data is based on brute-force simulations of small systems and thus may not precisely capture the true asymptotic scaling due to finite-size effects.

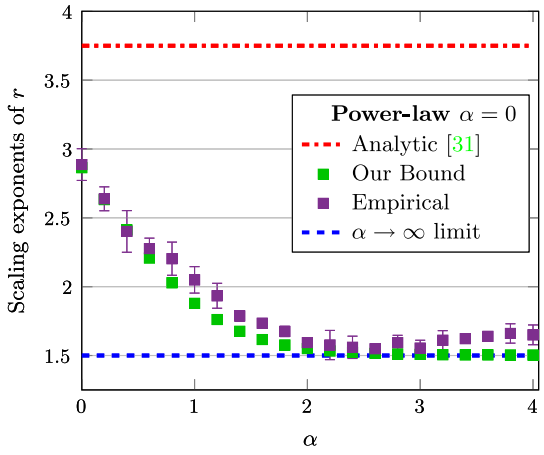

FIG. 5. Comparison of the empirical scaling exponents of the Trotter numbers (purple squares) against our bound (green squares) as functions of the system size at different values of the power-law exponent α. The error bars of the empirical values represent the standard deviation of the fitted exponents (see Fig. 4) over five instances of the random field h_j. The bound is derived from the counting argument in Eq. (53) and therefore has no standard deviation. We attribute the systematic difference between our bound and the empirical values to the fact that we only compute the empirical Trotter numbers up to $n = 11$, which may not capture the precise asymptotic scaling in the large-n limit. We also include the scaling exponent of the analytic bound in Ref. [31] (red dash-dotted line) as well as the theoretical exponent in the limit $\alpha \to \infty$, i.e., nearest-neighbor interactions (blue dashed line), for references.

$\alpha \in [0, 4]$. While the scaling exponent of the analytic bound in Ref. [31] is loose and independent of α, our bound appears to correctly capture the scaling of Trotter number at all values of α. In particular, it shows that the scaling exponent decreases with α—indicating fewer resources needed to simulate faster decaying interactions—and approaches the value for simulating nearest-neighbor interactions at large α.

VI. DISCUSSION

We have developed a general theory of Trotter error and identified a host of applications to simulating quantum dynamics, local observables, and quantum Monte Carlo methods. We work with arbitrary finite-dimensional operators as opposed to anti-Hermitian ones, which makes our theory applicable to both real- and imaginary-time evolutions. We consider Trotter error of various types, including additive error, multiplicative error, and error that appears in the exponent. For each type, we apply the correct order condition to cancel lower-order terms and represent higher-order ones as explicit nested commutators. The list of applications presented herein is not intended to be exhaustive, and we believe our techniques can uncover more speedups of the product-formula algorithm that were previously unknown.

Compared to the analysis of other simulation algorithms such as the truncated Taylor-series algorithm [23] and the qubitization approach [24], the derivation of our Trotter error theory is considerably more involved. However, the resulting error bounds are succinct and easy to evaluate. Theorem 6 shows that the Trotter error incurred by decomposing the evolution generated by $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ depends asymptotically on the quantity $\tilde{\alpha}_{\mathrm{comm}} = \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\|$, which can be computed by induction as for the second-quantized plane-wave electronic-structure, k-local Hamiltonians, rapidly decaying interacted systems, clustered Hamiltonians, the transverse field Ising model, and quantum ferromagnetic spin systems. We further show how to improve the analysis to find error bounds with small constant prefactors. Numerical simulation suggests that our higher-order error bounds are close to tight for systems with nearest-neighbor

and power-law interactions, and we hope future work can explore their tightness for other systems.

Our result shows that high-order product formulas can be advantageous for simulating many physical systems. Interestingly, we can often achieve this advantage without using a formula of very large order. For d-dimensional power-law interactions with exponent $\alpha > 2d$, we have shown that the pth-order product-formula algorithm has gate complexity $\mathcal{O}((nt)^{1+d/(\alpha-d)+1/p})$, whereas the state-of-the-art Lieb-Robinson-based approach requires $\tilde{\mathcal{O}}((nt)^{1+2d/(\alpha-d)})$ gates. Product formulas can thus scale better if $p \geq (\alpha - d)/d$, which is small for various physical systems such as the dipole-dipole interactions ($\alpha = 3$) and the van der Waals interactions ($\alpha = 6$). For other systems such as nearest-neighbor interactions and electronic-structure Hamiltonians, product formulas do not exactly match the state-of-the-art result in terms of the asymptotic scaling, but they are still advantageous for simulating systems of small sizes [31,43].

The complexity of the product-formula approach is determined by both the Trotter number (or Trotter error) and the cost per Trotter step. A naive implementation of each Trotter step exponentiates all the terms in the Hamiltonian, which has a cost that scales with the total number of terms. However, this worst-case complexity can be avoided by truncating the original Hamiltonian, as we have demonstrated in the simulation of rapidly decaying power-law Hamiltonians. Recent studies have proposed other techniques for implementing Trotter steps [9,43,75,76]. Those techniques can be applied in combination with our Trotter error analysis to further speed up the product-formula algorithm.

We have restricted our attention to the evolutions generated by time-independent operators. In the more general case, we have an operator-valued function $\mathcal{H}(\tau) = \sum_{\gamma=1}^{\Gamma} \mathcal{H}_\gamma(\tau)$, and our goal is to simulate the time-ordered evolution $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau \sum_{\gamma=1}^{\Gamma} \mathcal{H}_\gamma(\tau))$ [5,23,25,57,77–79]. Under certain smoothness assumptions, Ref. [57] shows that this evolution can be simulated using product formulas, although the analysis there does not exploit the commutativity of operator summands. We believe our approach can be extended to give improved analysis for time-dependent Hamiltonian simulation, but we leave a detailed study for future work.

Previous work considered several generalized product formulas, such as ones based on the divide-and-conquer construction [26], the randomized construction [27,29], and the linear-combination-of-unitaries construction [28]. The common underlying idea is to approximate the ideal evolution to pth order using formulas of order q_k, where $q_k \leq p$. Our theory can be applied to represent the q_kth-order Trotter error in terms of nested commutators, thus improving the previous analyses of Refs. [26–29]. This approach leads to a better understanding of these generalized formulas and justifies their potential utility in quantum simulation.

Several other questions related to our theory deserve further investigation. For example, the spectral-norm error bound computed here would be overly pessimistic if we simulate with a low-energy initial state. It would then be beneficial to change the error metric to the Euclidean distance to avoid the worst-case error propagation. Our analysis has also assumed an operator decomposition $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ given a prior, but one may instead seek an alternative decomposition to maximize the commutativity of operator summands. We focus on the error analysis within each Trotter step and apply the triangle inequality across different steps, which may be improved upon as hinted in previous work [80,81]. Finally, we have considered an idealized setting, and we hope future work could take the effect of noise into account [82].

Product formulas arguably provide the most straightforward approach to simulating quantum systems. This approach is empirically advantageous and is often the method of choice for near-term demonstration of quantum simulation. Despite their experimental success, the error scaling of product formulas was poorly understood, and prior to our work, their advantage was only rigorously analyzed for a restricted collection of systems. The theory developed here represents progress toward a precise characterization of the Trotter error, which we hope will bridge the gap between theoretical investigation and experimental realization of quantum simulation.

ACKNOWLEDGMENTS

Y. S. thanks Ryan Babbush, Fernando G. S. L. Brandão, Xun Gao, Jens Eisert, and Brian Swingle for inspiring discussions. He also received useful feedback during his visits to the Institute for Quantum Computing, the Center for Computation and Technology, and Lawrence Berkeley National Laboratory. We thank anonymous referees for their comments on an earlier draft. This work was supported in part by the U.S. Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research, Quantum Algorithms Teams and Accelerated Research in Quantum Computing programs; by the Army Research Office (MURI Grant No. W911NF-16-1-0349); and by the National Science Foundation (Grant No. CCF-1813814). In addition, Y. S. is supported by the Google Ph.D. Fellowship program. M. C. T. also acknowledges funding by the DOE BES QIS program (Award No. DE-SC0019449), the NSF PFCQC program, the DOE ASCR Quantum Testbed Pathfinder program (Award No. DE-SC0019040), AFOSR, ARL CDQI, and NSF PFC at JQI. N. W. is supported in part by a Google Quantum Research Award. S. C. Z. acknowledges support from NSF CAREER Grant No. CCF-1845125.

APPENDIX A: IDENTITIES OF TIME-ORDERED EVOLUTIONS

In this section, we discuss some useful identities of time-ordered exponentials. Let $\mathscr{H}(\tau)$ be a continuous operator-valued function defined for $\tau \in \mathbb{R}$. The time-ordered exponential generated by $\mathscr{H}(\tau)$ satisfies the differentiation rule (see p. 12 of Ref. [55])

$$\frac{\partial}{\partial t_2}\mathcal{T}\exp\left(\int_{t_1}^{t_2} \mathrm{d}\tau\, \mathscr{H}(\tau)\right) = \mathscr{H}(t_2)\mathcal{T}\exp\left(\int_{t_1}^{t_2} \mathrm{d}\tau\, \mathscr{H}(\tau)\right), \tag{A1}$$

$$\frac{\partial}{\partial t_1}\mathcal{T}\exp\left(\int_{t_1}^{t_2} \mathrm{d}\tau\, \mathscr{H}(\tau)\right) = -\mathcal{T}\exp\left(\int_{t_1}^{t_2} \mathrm{d}\tau\, \mathscr{H}(\tau)\right)\mathscr{H}(t_1), \tag{A2}$$

and the multiplicative property (see p. 11 of Ref. [55])

$$\mathcal{T}\exp\left(\int_{t_1}^{t_3} \mathrm{d}\tau \mathscr{H}(\tau)\right) = \mathcal{T}\exp\left(\int_{t_2}^{t_3} \mathrm{d}\tau \mathscr{H}(\tau)\right)\mathcal{T}\exp\left(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau)\right). \tag{A3}$$

By definition, the operator-valued function $\mathscr{U}(t) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau\, \mathscr{H}(\tau))$ satisfies the differential equation $(\mathrm{d}/\mathrm{d}\tau)\mathscr{U}(\tau) = \mathscr{H}(\tau)\mathscr{U}(\tau)$ with initial condition $\mathscr{U}(0) = I$. Generalizing this fact gives the following variation-of-parameters formula:

Lemma A.1: (Variation-of-parameters formula—see Theorem 4.9 in Ref. [83] and p. 17 in Ref. [55]). Let $\mathscr{H}(\tau)$, $\mathscr{R}(\tau)$ be continuous operator-valued functions defined for $\tau \in \mathbb{R}$. Then, the first-order differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{U}(t) = \mathscr{H}(t)\mathscr{U}(t) + \mathscr{R}(t), \quad \mathscr{U}(0) \text{ known}, \tag{A4}$$

has a unique solution given by the variation-of-parameters formula

$$\mathscr{U}(t) = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau\, \mathscr{H}(\tau)\right)\mathscr{U}(0) + \int_0^t \mathrm{d}\tau_1\, \mathcal{T}\exp\left(\int_{\tau_1}^t \mathrm{d}\tau_2\, \mathscr{H}(\tau_2)\right)\mathscr{R}(\tau_1). \tag{A5}$$

Let $\mathscr{H}(\tau) = \mathscr{A}(\tau) + \mathscr{B}(\tau)$ be a continuous operator-valued function with two summands defined for $0 \le \tau \le t$. Then, the evolution under $\mathscr{H}(\tau)$ can be seen as the evolution under the rotated operator $\mathcal{T}\exp^{-1}(\int_0^\tau \mathrm{d}\tau_2 \mathscr{A}(\tau_2))\mathscr{B}(\tau)\mathcal{T}\exp(\int_0^\tau \mathrm{d}\tau_2 \mathscr{A}(\tau_2))$, followed by another evolution under $\mathscr{A}(\tau)$ that rotates back to the original frame [25]. This is known as the "interaction-picture" representation in quantum mechanics and is formally stated in the following lemma.

Lemma A.2: (Time-ordered evolution in the interaction picture). Let $\mathscr{H}(\tau) = \mathscr{A}(\tau) + \mathscr{B}(\tau)$ be an operator-valued function defined for $\tau \in \mathbb{R}$ with continuous summands $\mathscr{A}(\tau)$ and $\mathscr{B}(\tau)$. Then,

$$\mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau\, \mathscr{H}(\tau)\right) = \mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau\, \mathscr{A}(\tau)\right) \times \mathcal{T}\exp\left[\int_0^t \mathrm{d}\tau_1\, \mathcal{T}\exp^{-1}\left(\int_0^{\tau_1} \mathrm{d}\tau_2\, \mathscr{A}(\tau_2)\right) \times \mathscr{B}(\tau_1)\mathcal{T}\exp\left(\int_0^{\tau_1} \mathrm{d}\tau_2\, \mathscr{A}(\tau_2)\right)\right]. \tag{A6}$$

Proof.—A simple calculation shows that the right-hand side of the above equation satisfies the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{U}(t) = \mathscr{H}(t)\mathscr{U}(t) \tag{A7}$$

with initial condition $\mathscr{U}(0) = I$. The lemma then follows, as $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau \mathscr{H}(\tau))$ is the unique solution to this differential equation. ∎

For any continuous $\mathscr{H}(\tau)$, the evolution $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau \mathscr{H}(\tau))$ it generates is invertible and continuously differentiable. Conversely, the following lemma asserts that any operator-valued function that is invertible and continuously differentiable is a time-ordered evolution generated by some continuous function.

Lemma A.3: (Fundamental theorem of time-ordered evolution—see p. 20 of Ref. [55]) The following statements regarding an operator-valued function $\mathscr{U}(\tau)$ $(\tau \in \mathbb{R})$ are equivalent:

(1) $\mathscr{U}(\tau)$ is invertible and continuously differentiable;
(2) $\mathscr{U}(\tau) = \mathcal{T}\exp(\int_0^\tau \mathrm{d}\tau_1 \mathscr{H}(\tau_1))\mathscr{U}(0)$ for some continuous operator-valued function $\mathscr{H}(\tau)$.

Furthermore, in the second statement, $\mathscr{H}(\tau) = [(\mathrm{d}/\mathrm{d}\tau)\mathscr{U}(\tau)]\mathscr{U}^{-1}(\tau)$ is uniquely determined.

Finally, we bound the spectral norm of a time-ordered evolution $\mathcal{T}\exp(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau))$ and the distance between two evolutions.

Lemma A.4: [Spectral-norm bound for time-ordered evolution—see p. 28 of Ref. [55]] Let $\mathscr{H}(\tau)$ be a continuous operator-valued function defined on $\mathbb{R}$. Then,

(1) $\|\mathcal{T}\exp(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau))\| \le e^{\left|\int_{t_1}^{t_2} \mathrm{d}\tau \|\mathscr{H}(\tau)\|\right|}$; and
(2) $\|\mathcal{T}\exp(\int_{t_1}^{t_2} \mathrm{d}\tau \mathscr{H}(\tau))\| = 1$ if $\mathscr{H}(\tau)$ is anti-Hermitian.

Corollary A.5. [Distance bound for time-ordered evolutions—Appendix B in Ref. [52]] Let $\mathscr{H}(\tau)$ and $\mathscr{G}(\tau)$ be continuous operator-valued functions defined on $\mathbb{R}$. Then,

(1) $\|\mathcal{T}\exp(\int_{t_1}^{t_2}\mathrm{d}\tau\mathcal{H}(\tau)) - \mathcal{T}\exp(\int_{t_1}^{t_2}\mathrm{d}\tau\mathcal{G}(\tau))\| \leq |\int_{t_1}^{t_2}\mathrm{d}\tau\|\mathcal{H}(\tau)-\mathcal{G}(\tau)\||e^{|\int_{t_1}^{t_2}\mathrm{d}\tau(\|\mathcal{H}(\tau)\|+\|\mathcal{G}(\tau)\|)|}$; and
(2) $\|\mathcal{T}\exp(\int_{t_1}^{t_2}\mathrm{d}\tau\mathcal{H}(\tau)) - \mathcal{T}\exp(\int_{t_1}^{t_2}\mathrm{d}\tau\mathcal{G}(\tau))\| \leq |\int_{t_1}^{t_2}\mathrm{d}\tau\|\mathcal{H}(\tau)-\mathcal{G}(\tau)\||$ if $\mathcal{H}(\tau)$ and $\mathcal{G}(\tau)$ are anti-Hermitian.

APPENDIX B: TROTTER ERROR WITH 1-NORM SCALING

In this section, we provide the proof for Lemma 1.

Proof.—Proof of Lemma 1. Since $\mathcal{S}(t)$ is a pth-order formula, we know from Supplementary Lemma 1 in Ref. [35] that $\mathcal{S}(0)=\mathcal{S}'(0)=\cdots=\mathcal{S}^{(p)}(0)=0$. By Taylor's theorem,

$$\mathcal{S}(t)-e^{tH}=(p+1)\int_0^1\mathrm{d}u(1-u)^p\frac{t^{p+1}}{(p+1)!}\cdot(\mathcal{S}^{(p+1)}(ut)-H^{p+1}e^{utH}), \tag{B1}$$

where

$$\mathcal{S}^{(p+1)}(ut)=\sum_{q_{(1,1)}+\cdots+q_{(\Upsilon,\Gamma)}=p+1}\binom{p+1}{q_{(1,1)}\cdots q_{(\Upsilon,\Gamma)}}\cdot\prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma}(a_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)})^{q_{(\upsilon,\gamma)}}e^{uta_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}. \tag{B2}$$

The spectral norms of $\mathcal{S}^{(p+1)}(ut)$ and $H^{p+1}e^{utH}$ can be bounded as

$$\begin{aligned}\|\mathcal{S}^{(p+1)}(ut)\|&\leq\sum_{q_{(1,1)}+\cdots+q_{(\Upsilon,\Gamma)}=p+1}\binom{p+1}{q_{(1,1)}\cdots q_{(\Upsilon,\Gamma)}}\cdot\prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma}\|H_{\pi_\upsilon(\gamma)}\|^{q_{(\upsilon,\gamma)}}e^{t\|H_{\pi_\upsilon(\gamma)}\|}\\&=\left(\Upsilon\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|\right)^{p+1}e^{t\Upsilon\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|},\end{aligned} \tag{B3}$$

$$\|H^{p+1}e^{utH}\|\leq\left(\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|\right)^{p+1}e^{t\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|}. \tag{B4}$$

Applying these bounds to the Taylor expansion, we find that

$$\begin{aligned}\|\mathcal{S}(t)-e^{tH}\|&\leq\frac{t^{p+1}}{(p+1)!}\left[\left(\Upsilon\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|\right)^{p+1}e^{t\Upsilon\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|}+\left(\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|\right)^{p+1}e^{t\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|}\right]\end{aligned} \tag{B5}$$

$$=\mathcal{O}\left(\left(\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|t\right)^{p+1}e^{t\Upsilon\sum_{\gamma=1}^{\Gamma}\|H_\gamma\|}\right). \tag{B6}$$

The special case where H_γ are anti-Hermitian can be proved in a similar way, except we directly evaluate the spectral norm of a matrix exponential to 1. ∎

APPENDIX C: ERROR TYPES

In this Appendix, we consider different types of Trotter error for a general product formula introduced in Sec. III B. In particular, we will prove Theorem 3, which gives explicit expressions for three types of Trotter error: the additive error, the multiplicative error, and the error that appears in the exponent of a time-ordered exponential. These types are equivalent for analyzing the complexity of simulating quantum dynamics and local observables, but the latter two are more versatile for quantum Monte Carlo simulation.

Let $H=\sum_{\gamma=1}^{\Gamma}H_\gamma$ be an operator with Γ summands. We decompose the evolution $e^{tH}=e^{t\sum_{\gamma=1}^{\Gamma}H_\gamma}$ using a general product formula $\mathcal{S}(t)=\prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma}e^{ta_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}$. We impose the lexicographical order on the tuples (υ,γ) as in Sec. III B so that

$$\mathcal{S}(t)=\overleftarrow{\prod_{(\upsilon,\gamma)}}e^{ta_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}. \tag{C1}$$

To compute the additive error, we construct the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(t)=H\mathcal{S}(t)+\mathscr{R}(t), \tag{C2}$$

with initial condition $\mathcal{S}(0)=I$, where

$$\begin{aligned}\mathscr{R}(t)\coloneqq&\sum_{(\upsilon,\gamma)}\overleftarrow{\prod_{(\upsilon',\gamma')\succ(\upsilon,\gamma)}}e^{ta_{(\upsilon',\gamma')}H_{\pi_{\upsilon'}(\gamma')}}(a_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)})\\&\cdot\overleftarrow{\prod_{(\upsilon',\gamma')\preccurlyeq(\upsilon,\gamma)}}e^{ta_{(\upsilon',\gamma')}H_{\pi_{\upsilon'}(\gamma')}}\\&-H\overleftarrow{\prod_{(\upsilon',\gamma')}}e^{ta_{(\upsilon',\gamma')}H_{\pi_{\upsilon'}(\gamma')}}.\end{aligned}$$

By the variation-of-parameters formula (Lemma A.1), $\mathcal{S}(t)-e^{tH}=\int_0^t\mathrm{d}\tau\,e^{(t-\tau)H}\mathscr{R}(\tau)$, so we obtain the additive error

$$\mathscr{A}(t)\coloneqq\int_0^t\mathrm{d}\tau\,e^{(t-\tau)H}\mathscr{R}(\tau). \tag{C3}$$

This result suffices if our purpose is to only compute the additive error operator. However, for the later discussion in Appendix E, it is convenient to further rewrite

$$\mathscr{A}(t) = \int_0^t \mathrm{d}\tau\, e^{(t-\tau)H} \mathcal{S}(\tau)\mathcal{T}(\tau), \tag{C4}$$

where

$$\begin{aligned}\mathcal{T}(\tau) \coloneqq & \sum_{(v,\gamma)} \prod_{(v',\gamma')\prec(v,\gamma)}^{\rightarrow} e^{-\tau a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}} (a_{(v,\gamma)} H_{\pi_v(\gamma)}) \\ & \cdot \prod_{(v',\gamma')\prec(v,\gamma)}^{\leftarrow} e^{\tau a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}} \\ & - \prod_{(v',\gamma')}^{\rightarrow} e^{-\tau a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}} H \prod_{(v',\gamma')}^{\leftarrow} e^{\tau a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}}. \end{aligned} \tag{C5}$$

Note that we have rewritten part of the error operator as a linear combination of conjugation of matrix exponentials. In Appendix E, we apply the correct order condition to further represent it as nested commutators of the operator summands H_γ.

For the exponentiated type of Trotter error, we aim to construct an operator-valued function $\mathcal{E}(t)$ such that

$$\mathcal{S}(t) = \mathcal{T} \exp\left(\int_0^t \mathrm{d}\tau (H + \mathcal{E}(\tau)) \right). \tag{C6}$$

To do this, we differentiate the product formula $\mathcal{S}(t)$ and obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(t) = & \sum_{(v,\gamma)} \prod_{(v',\gamma')\succ(v,\gamma)}^{\leftarrow} e^{t a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}} (a_{(v,\gamma)} H_{\pi_v(\gamma)}) \\ & \cdot \prod_{(v',\gamma')\preccurlyeq(v,\gamma)}^{\leftarrow} e^{t a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}} \\ = & \mathscr{F}(t)\mathcal{S}(t), \end{aligned} \tag{C7}$$

where

$$\begin{aligned}\mathscr{F}(t) \coloneqq & \sum_{(v,\gamma)} \prod_{(v',\gamma')\succ(v,\gamma)}^{\leftarrow} e^{t a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}} (a_{(v,\gamma)} H_{\pi_v(\gamma)}) \\ & \cdot \prod_{(v',\gamma')\succ(v,\gamma)}^{\rightarrow} e^{-t a_{(v',\gamma')} H_{\pi_{v'}(\gamma')}}. \end{aligned} \tag{C8}$$

Applying the fundamental theorem of time-ordered evolution (Lemma A.3), we have

$$\mathcal{S}(t) = \mathcal{T} \exp\left(\int_0^t \mathrm{d}\tau\, \mathscr{F}(\tau) \right), \tag{C9}$$

which gives the exponentiated error

$$\mathcal{E}(t) \coloneqq \mathscr{F}(t) - H. \tag{C10}$$

From the exponentiated type of Trotter error, we can obtain the multiplicative error by switching to the interaction picture. Specifically, we apply Lemma A.2 and get

$$\begin{aligned}\mathcal{S}(t) &= \mathcal{T} \exp\left(\int_0^t \mathrm{d}\tau (H + \mathcal{E}(\tau)) \right) \\ &= e^{tH} \mathcal{T} \exp\left(\int_0^t \mathrm{d}\tau\, e^{-\tau H} \mathcal{E}(\tau) e^{\tau H} \right). \end{aligned} \tag{C11}$$

Then, the operator-valued function

$$\mathscr{M}(t) \coloneqq \mathcal{T} \exp\left(\int_0^t \mathrm{d}\tau\, e^{-\tau H} \mathcal{E}(\tau) e^{\tau H} \right) - I \tag{C12}$$

is the multiplicative error of the product formula. We have thus established Theorem 3.

APPENDIX D: ORDER CONDITIONS

In this Appendix, we continue the discussion in Sec. III C about order conditions of the Trotter error. We show how to use these conditions to cancel low-order terms of the Taylor series. Toward the end of this section, we establish Theorem 4, which gives order conditions for the additive, multiplicative, and exponentiated Trotter error. We apply these conditions to prove the main result on the commutator scaling of the Trotter error.

Recall from Sec. III C that the order condition of an operator-valued function $\mathscr{F}(\tau)$ represents the rate at which $\mathscr{F}(\tau)$ approaches zero when $\tau \to 0$. Formally, given a continuous operator-valued function $\mathscr{F}(\tau)$ defined on $\mathbb{R}$, we write $\mathscr{F}(\tau) = O(\tau^p)$ with non-negative integer p if there exist constants $c, t_0 > 0$, independent of τ, such that $\|\mathscr{F}(\tau)\| \le c|\tau|^p$ whenever $|\tau| \le t_0$. To verify this, it suffices to check that the limit

$$\lim_{\tau\to 0} \frac{\|\mathscr{F}(\tau)\|}{|\tau|^p} \tag{D1}$$

exists.

As previously mentioned, our approach uses the order condition $\mathscr{F}(\tau) = O(\tau^p)$ to argue that terms with order $1, \tau, \ldots, \tau^{p-1}$ vanish in the Taylor series of $\mathscr{F}(\tau)$. This argument is rigorized in Lemma 6 of Ref. [35], which we restate and prove for completeness.

Lemma D.1: (Derivative condition). Any continuous operator-valued function $\mathscr{F}(\tau)$ defined on $\mathbb{R}$ satisfies the order condition

$$\mathscr{F}(\tau) = O(1). \tag{D2}$$

Furthermore, if $\mathcal{F}(\tau)$ has p continuous derivatives for some positive integer p, then the following two conditions are equivalent:

(1) $\mathcal{F}(\tau) = O(\tau^p)$; and
(2) $\mathcal{F}(0) = \mathcal{F}'(0) = \cdots = \mathcal{F}^{(p-1)}(0) = 0$.

Proof.—The continuity of $\mathcal{F}(\tau)$ at $\tau = 0$ implies $\mathcal{F}(\tau) = O(1)$ by definition. Assume that $\mathcal{F}(\tau)$, $\mathcal{F}'(\tau)$,..., $\mathcal{F}^{(p)}(\tau)$ exist and are continuous. If Condition 2 holds, we have

$$\lim_{\tau\to 0}\frac{\|\mathcal{F}(\tau)\|}{|\tau|^p} = \left\|\lim_{\tau\to 0}\frac{\mathcal{F}(\tau)}{\tau^p}\right\| = \frac{\|\mathcal{F}^{(p)}(0)\|}{p!} \tag{D3}$$

by the L'Hôpital rule. This proves that Condition 1 holds.

Given Condition 1, we have, by definition, that

$$\|\mathcal{F}(\tau)\| \le c|\tau|^p \tag{D4}$$

for some $c, t_0 > 0$ and all $|\tau| \le t_0$. Suppose, by contradiction, that Condition 2 is not true. Then, we let $0 \le j \le p-1$ be the first integer for which $\mathcal{F}^{(j)}(0) \ne 0$. We use Taylor's theorem to order j to get

$$\mathcal{F}(\tau) = \mathcal{F}^{(j)}(0)\frac{\tau^j}{j!} + \int_0^\tau \mathrm{d}\tau_2 \mathcal{F}^{(j+1)}(\tau-\tau_2)\frac{\tau_2^j}{j!}, \tag{D5}$$

which implies

$$\|\mathcal{F}(\tau)\| \ge \|\mathcal{F}^{(j)}(0)\|\frac{|\tau|^j}{j!} - \max_{|\tau_2|\le|\tau|}\|\mathcal{F}^{(j+1)}(\tau_2)\|\frac{|\tau|^{j+1}}{(j+1)!} \tag{D6}$$

by the triangle inequality. We combine the above inequalities and divide both sides by $|\tau|^j$. Taking the limit $\tau \to 0$ gives the contradiction $\|\mathcal{F}^{(j)}(0)\| \le 0$. ∎

Lemma D.1 provides a direct approach to computing order conditions for functions of real variables. This approach works for simple examples such as the power functions $f(\tau) = \tau^p = O(\tau^p)$. Another example that we use in our analysis is the integration of a monomial, like

$$\int_0^\tau \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \int_0^{\tau_1} \mathrm{d}\tau_3 \int_0^{\tau_2} \mathrm{d}\tau_4 \tau_1^3\tau_2\tau_3^4\tau_4^5. \tag{D7}$$

As the following lemma shows, we can directly evaluate such an integral and compute the order condition of the resulting power function.

Lemma D.2: (Integration of a monomial). The integration of a monomial $\tau_1^{p_1}\cdots\tau_\gamma^{p_\gamma}\cdots\tau_\Gamma^{p_\Gamma}$ is evaluated as

$$\int_0^\tau \mathrm{d}\tau_1 \cdots \int_0^{\tau_{<\gamma}} \mathrm{d}\tau_\gamma \cdots \int_0^{\tau_{<\Gamma}} \mathrm{d}\tau_\Gamma\, \tau_1^{p_1}\cdots\tau_\gamma^{p_\gamma}\cdots\tau_\Gamma^{p_\Gamma} = ct^{p_1+\cdots+p_\Gamma+\Gamma} = O(t^{p_1+\cdots+p_\Gamma+\Gamma}), \tag{D8}$$

where $\tau_{<\gamma} \in \{\tau, \tau_1, \ldots, \tau_{\gamma-1}\}$ and c is a constant that depends on non-negative integers $p_1, \ldots, p_\Gamma$.

Proof.—We prove the lemma by induction on the value of Γ. The claim trivially holds when $\Gamma = 1$. Suppose that it is true for Γ. For $\Gamma+1$, we have

$$\int_0^\tau \mathrm{d}\tau_1 \cdots \int_0^{\tau_{<\Gamma+1}} \mathrm{d}\tau_{\Gamma+1}\, \tau_1^{p_1}\cdots\tau_{\Gamma+1}^{p_{\Gamma+1}} = \int_0^\tau \mathrm{d}\tau_1 \cdots \int_0^{\tau_{<\Gamma}} \mathrm{d}\tau_\Gamma \frac{\tau_1^{q_1}\cdots\tau_\Gamma^{q_\Gamma}}{p_{\Gamma+1}+1}, \tag{D9}$$

where $q_1 + \cdots + q_\Gamma = p_1 + \cdots + p_{\Gamma+1} + 1$. The claim then follows from the inductive hypothesis. ∎

For most of our analysis, however, a direct calculation of order conditions is inefficient. In particular, a $(2k)$th-order Suzuki formula contains $2\times 5^{k-1}$ matrix exponentials, and a direct analysis becomes prohibitive when k is large. Instead, we follow standard rules of order conditions to compute them indirectly, some of which are summarized below:

Proposition D.3: [Rules of order conditions] Let $\mathcal{F}(\tau)$ and $\mathcal{G}(\tau)$ be operator-valued functions defined on $\mathbb{R}$ that are infinitely differentiable. Let p and q be non-negative integers. The following rules of order conditions hold:

(1) Addition: if $\mathcal{F}(\tau) = O(\tau^p)$ and $\mathcal{G}(\tau) = O(\tau^q)$, then $\mathcal{F}(\tau) + \mathcal{G}(\tau) = O(\tau^{\min(p,q)})$;
(2) Multiplication: if $\mathcal{F}(\tau) = O(\tau^p)$ and $\mathcal{G}(\tau) = O(\tau^q)$, then $\mathcal{F}(\tau)\mathcal{G}(\tau) = O(\tau^{p+q})$;
(3) Differentiation: $\mathcal{F}(\tau) = O(\tau^{p+1})$ if and only if $\mathcal{F}(0) = 0$ and $\mathcal{F}'(\tau) = O(\tau^p)$;
(4) Integration: $\mathcal{F}(\tau) = O(\tau^p)$ if and only if $\int_0^t \mathrm{d}\tau\mathcal{F}(\tau) = O(t^{p+1})$; and
(5) Exponentiation: $\mathcal{F}(\tau) = \mathcal{G}(\tau) + O(\tau^p)$ if and only if $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau\mathcal{F}(\tau)) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau\mathcal{G}(\tau)) + O(t^{p+1})$.

Proof.—We only prove the exponentiation rule, as the other rules follow directly from Lemma D.1. Suppose that $\mathcal{T}\exp(\int_0^t \mathrm{d}\tau\,\mathcal{F}(\tau)) = \mathcal{T}\exp(\int_0^t \mathrm{d}\tau\,\mathcal{G}(\tau)) + O(t^{p+1})$. To prove $\mathcal{F}(\tau) = \mathcal{G}(\tau) + O(\tau^p)$, it suffices to show that $\mathcal{F}^{(q)}(0) = \mathcal{G}^{(q)}(0)$ for $q = 0, \ldots, p-1$.

We prove this by induction. By the differentiation rule, we have

$$\mathcal{F}(t)\mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau\,\mathcal{F}(\tau)\right) = \mathcal{G}(t)\mathcal{T}\exp\left(\int_0^t \mathrm{d}\tau\,\mathcal{G}(\tau)\right) + O(t^p), \tag{D10}$$

so Lemma D.1 implies $\mathcal{F}(0) = \mathcal{G}(0)$. This proves the claim in the base case. Now, assume that $\mathcal{F}^{(l)}(0) = \mathcal{G}^{(l)}(0)$ holds for $l = 0, \ldots q$, where $q < p-1$. By Lemma D.1 and the general Leibniz rule,

$$\sum_{l=0}^{q+1}\binom{q+1}{l}\mathscr{F}^{q+1-l}(0)\mathcal{T}\exp^{(l)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{F}(\tau)\right) = \sum_{l=0}^{q+1}\binom{q+1}{l}\mathscr{G}^{q+1-l}(0)\mathcal{T}\exp^{(l)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{G}(\tau)\right).$$

Lemma D.1 also implies $\mathcal{T}\exp^{(l)}(\int_0^0 \mathsf{d}\tau\,\mathscr{F}(\tau)) = \mathcal{T}\exp^{(l)}(\int_0^0 \mathsf{d}\tau\,\mathscr{G}(\tau))$ for $l = 0, \ldots, q+1$. So, the above equation simplifies to

$$\mathscr{F}^{(q+1)}(0) = \mathscr{G}^{(q+1)}(0). \tag{D11}$$

This completes the inductive step.

For the reverse direction, we want to prove $\mathcal{T}\exp(\int_0^t \mathsf{d}\tau\,\mathscr{F}(\tau)) = \mathcal{T}\exp(\int_0^t \mathsf{d}\tau\,\mathscr{G}(\tau)) + O(t^{p+1})$, assuming that $\mathscr{F}(\tau) = \mathscr{G}(\tau) + O(\tau^p)$. Equivalently, we want to show that $\mathcal{T}\exp^{(q+1)}(\int_0^0 \mathsf{d}\tau\,\mathscr{F}(\tau)) = \mathcal{T}\exp^{(q+1)}(\int_0^0 \mathsf{d}\tau\,\mathscr{G}(\tau))$ for $q = 0, \ldots, p-1$, given that $\mathscr{F}^{(q)}(0) = \mathscr{G}^{(q)}(0)$. This equality can be proved by induction and by applying the Leibniz rule in a similar way as above. Specifically, the base case follows from

$$\mathcal{T}\exp^{(1)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{F}(\tau)\right) = \mathscr{F}(0) = \mathscr{G}(0) = \mathcal{T}\exp^{(1)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{G}(\tau)\right),$$

and the inductive step follows from

$$\mathcal{T}\exp^{(q+1)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{F}(\tau)\right) = \sum_{l=0}^{q}\binom{q}{l}\mathscr{F}^{(q-l)}(0)\mathcal{T}\exp^{(l)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{F}(\tau)\right) \tag{D12}$$

$$= \sum_{l=0}^{q}\binom{q}{l}\mathscr{G}^{(q-l)}(0)\mathcal{T}\exp^{(l)}\left(\int_0^0 \mathsf{d}\tau\,\mathscr{G}(\tau)\right) \tag{D13}$$

$$= \mathcal{T}\exp^{(q+1)}\left(\int_0^0 \mathsf{d}\tau\mathscr{G}(\tau)\right). \tag{D14}$$

■

We now compute order conditions for the additive, multiplicative, and exponentiated Trotter error.

Proof.—Proof of Theorem 4 Suppose that $\mathcal{T}(\tau) = O(\tau^p)$. We apply the multiplication rule of Proposition D.3 to get $e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau) = O(\tau^p)$. A further application of the integration rule gives $\mathcal{S}(t) - e^{tH} = \int_0^t \mathsf{d}\tau e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau) = O(t^{p+1})$.

Conversely, let $\mathcal{S}(t) = e^{tH} + O(t^{p+1})$. This implies $\int_0^t \mathsf{d}\tau e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau) = O(t^{p+1})$. Applying the integration rule and the multiplication rule gives $\mathcal{S}(\tau)\mathcal{T}(\tau) = O(\tau^p)$. Note that $\mathcal{S}(t) = e^{tH} + O(t^{p+1}) = I + O(t)$ implies that the operator-valued function $\mathcal{S}(t)$ is invertible for sufficiently small t, and since $(\mathsf{d}/\mathsf{d}t)\mathcal{S}^{-1}(t) = -\mathcal{S}^{-1}(t)\mathcal{S}'(t)\mathcal{S}^{-1}(t)$, the inverse function $\mathcal{S}^{-1}(t)$ is infinitely differentiable. Applying the multiplication rule gives $\mathcal{T}(\tau) = O(\tau^p)$, which establishes the equivalence of Conditions 1 and 2.

Note that $\mathcal{S}(t) = e^{tH} + O(t^{p+1})$ is equivalent to $\mathcal{T}\exp(\int_0^t \mathsf{d}\tau(H + \mathcal{E}(\tau))) = e^{tH} + O(t^{p+1})$, which is further equivalent to $H + \mathcal{E}(\tau) = H + O(\tau^p)$ by the exponentiation rule. Canceling H from both sides proves the equivalence of Conditions 1 and 3.

Finally, note that $\mathcal{S}(t) = e^{tH}(I + \mathscr{M}(t)) = e^{tH} + O(t^{p+1})$ can be simplified to $e^{tH}\mathscr{M}(t) = O(t^{p+1})$. The equivalence of Conditions 1 and 4 then follows from the multiplication rule. ■

APPENDIX E: ERROR REPRESENTATIONS

We now prove Theorem 6, which establishes the commutator scaling of the Trotter error. The proof is sketched in Sec. III D and will be detailed here. For simplicity, we only consider the additive error, although the analysis can be easily adapted to handle the multiplicative error and the exponentiated error.

Let $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be an operator that generates the evolution $e^{tH} = e^{t\sum_{\gamma=1}^{\Gamma} H_\gamma}$. Let $\mathcal{S}(t) = \prod_{\upsilon=1}^{\Upsilon}\prod_{\gamma=1}^{\Gamma} e^{t a_{(\upsilon,\gamma)} H_{\pi_\upsilon(\gamma)}}$ be a pth-order product formula as in Sec. II C. We know from Theorem 3 that the Trotter error can be expressed in an additive form as $\mathcal{S}(t) = e^{tH} + \int_0^t \mathsf{d}\tau e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau)$, where

$$\mathcal{T}(\tau) = \sum_{(\upsilon,\gamma)}\ \prod_{(\upsilon',\gamma')\prec(\upsilon,\gamma)}^{\rightarrow} e^{-\tau a_{(\upsilon',\gamma')} H_{\pi_{\upsilon'}(\gamma')}}\left(a_{(\upsilon,\gamma)} H_{\pi_\upsilon(\gamma)}\right) \cdot \prod_{(\upsilon',\gamma')\prec(\upsilon,\gamma)}^{\leftarrow} e^{\tau a_{(\upsilon',\gamma')} H_{\pi_{\upsilon'}(\gamma')}} - \prod_{(\upsilon',\gamma')}^{\rightarrow} e^{-\tau a_{(\upsilon',\gamma')} H_{\pi_{\upsilon'}(\gamma')}} H \prod_{(\upsilon',\gamma')}^{\leftarrow} e^{\tau a_{(\upsilon',\gamma')} H_{\pi_{\upsilon'}(\gamma')}}.$$

Furthermore, Theorem 4 implies that the operator-valued function $\mathcal{T}(\tau)$ satisfies the order condition $\mathcal{T}(\tau) = O(\tau^p)$.

We now apply Theorem 5 to expand every conjugation of matrix exponentials in $\mathcal{T}(\tau)$. In doing so, we only keep track of terms of order $O(\tau^p)$, as those terms corresponding to $1, \tau, \ldots, \tau^{p-1}$ will vanish due to the order condition. We obtain

$$\|\mathcal{T}(\tau)\| \le \sum_{(v,\gamma)} \alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}, (v',\gamma') \prec (v,\gamma)\}}, H_{\pi_v(\gamma)}) \times \frac{\tau^p}{p!}\exp\left(2\tau \sum_{(v',\gamma')\prec(v,\gamma)} \|H_{\pi_{v'}(\gamma')}\|\right) + \alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}\}}, H)\frac{\tau^p}{p!} e^{2\tau\sum_{(v',\gamma')}\|H_{\pi_{v'}(\gamma')}\|},$$

where $\vec{\{\}}$ denotes an ordered list where elements have increasing indices from left to right. This is further bounded by

$$\begin{aligned}\|\mathcal{T}(\tau)\| &\le 2\sum_{(v,\gamma)} \alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}\}}, H_{\pi_v(\gamma)}) \times \frac{\tau^p}{p!}\exp\left(2\tau \sum_{(v',\gamma')} \|H_{\pi_{v'}(\gamma')}\|\right) \\ &= 2\Upsilon \sum_{\gamma=1}^{\Gamma} \alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}\}}, H_\gamma) \times \frac{\tau^p}{p!}\exp\left(2\tau\Upsilon \sum_{\gamma'=1}^{\Gamma} \|H_{\gamma'}\|\right).\end{aligned}$$

After a final integration over τ, we have

$$\begin{aligned}\|\mathcal{S}(t) - e^{tH}\| &\le \int_0^t \mathrm{d}\tau \|e^{(t-\tau)H}\mathcal{S}(\tau)\mathcal{T}(\tau)\| \\ &\le 2\Upsilon \sum_{\gamma=1}^{\Gamma} \alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}\}}, H_\gamma) \times \frac{t^{p+1}}{(p+1)!}\exp\left(4t\Upsilon \sum_{\gamma'=1}^{\Gamma} \|H_{\gamma'}\|\right).\end{aligned}$$

This bound holds for arbitrary operators H_γ. If the operator summands are anti-Hermitian, the bound can be further tightened to

$$\|\mathcal{S}(t) - e^{tH}\| \le 2\Upsilon \sum_{\gamma=1}^{\Gamma} \alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}\}}, H_\gamma)\frac{t^{p+1}}{(p+1)!}. \quad \text{(E1)}$$

Note that our analysis depends on $\pi_{v'}$, the ordering of operator summands in stage v' of the product formula. In the following, we prove an asymptotic bound that removes this ordering constraint. The resulting bound is independent of the definition of product formula and may thus be easier to compute in practice. Our analysis here is not tight in terms of the constant prefactor, but it is sufficient to establish the commutator scaling in Theorem 6.

Recall from Theorem 5 that

$$\alpha_{\mathrm{comm}}(\overrightarrow{\{H_{\pi_{v'}(\gamma')}\}}, H_\gamma) = \sum_{q_{(1,1)}+\cdots+q_{(v',\gamma')}=p} \binom{p}{q_{(1,1)}\cdots q_{(v',\gamma')}} \times \|\mathrm{ad}^{q_{(1,1)}}_{H_{\pi_1(1)}}\cdots\mathrm{ad}^{q_{(\Upsilon,\Gamma)}}_{H_{\pi_\Upsilon(\Gamma)}}(H_\gamma)\|, \quad \text{(E2)}$$

which is upper bounded by $p!$ times $\sum_{q_{(1,1)}+\cdots+q_{(v',\gamma')}=p}\|\mathrm{ad}^{q_{(1,1)}}_{H_{\pi_1(1)}}\cdots\mathrm{ad}^{q_{(\Upsilon,\Gamma)}}_{H_{\pi_\Upsilon(\Gamma)}}(H_\gamma)\|$. Fixing the value of γ, we claim that

$$\sum_{q_{(1,1)}+\cdots+q_{(v',\gamma')}=p} \|\mathrm{ad}^{q_{(1,1)}}_{H_{\pi_1(1)}}\cdots\mathrm{ad}^{q_{(\Upsilon,\Gamma)}}_{H_{\pi_\Upsilon(\Gamma)}}(H_\gamma)\| \le \Upsilon^p \sum_{\gamma_{p+1}=1}^{\Gamma}\cdots\sum_{\gamma_2=1}^{\Gamma}\|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_\gamma]\cdots]\|.$$

This result can be seen as follows. Every nested commutator on the left-hand side has p nesting layers and must thus be of the form on the right. Conversely, we fix one term $\|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_\gamma]\cdots]\|$ from the right and bound the number of times this term might appear on the left. Each operator $H_{\gamma_2},\ldots,H_{\gamma_{p+1}}$ can appear in Υ possible stages, and hence there are Υ^p possibilities in total. When the stages are fixed, this will uniquely determine one term, $\|\mathrm{ad}^{q_{(1,1)}}_{H_{\pi_1(1)}}\cdots\mathrm{ad}^{q_{(\Upsilon,\Gamma)}}_{H_{\pi_\Upsilon(\Gamma)}}(H_\gamma)\|$, on the left. We have thus established Theorem 6.

APPENDIX F: SIMULATING SECOND-QUANTIZED ELECTRONIC STRUCTURE

In this section, we use product formulas to simulate the second-quantized plane-wave electronic structure

$$H = \underbrace{\frac{1}{2n}\sum_{j,k,\nu}\kappa_\nu^2\cos[\kappa_\nu\cdot r_{k-j}]A_j^\dagger A_k}_{T} \underbrace{-\frac{4\pi}{\omega}\sum_{j,\iota,\nu\neq 0}\frac{\zeta_\iota\cos[\kappa_\nu\cdot(\tilde{r}_\iota - r_j)]}{\kappa_\nu^2}N_j}_{U} \underbrace{+\frac{2\pi}{\omega}\sum_{\substack{j\neq k\\ \nu\neq 0}}\frac{\cos[\kappa_\nu\cdot r_{j-k}]}{\kappa_\nu^2}N_jN_k}_{V}, \quad \text{(F1)}$$

where j, k range over all n orbitals, ω is the volume of the computational cell, and we consider the constant density case where $n/\omega = \mathcal{O}(1)$. Here, $\kappa_\nu = 2\pi\nu/\omega^{1/3}$ are n vectors of the plane-wave frequencies, where ν are three-dimensional vectors of integers with elements in

$[-n^{1/3}, n^{1/3}]$, r_j are the positions of electrons, ζ_ι are nuclei charges such that $\sum_\iota |\zeta_\iota| = \mathcal{O}(n)$, and $\tilde{r}_\iota$ are the nuclear coordinates. Note that $A_j^\dagger$ and A_k are the creation and annihilation operators, and $N_j = A_j^\dagger A_j$ are the number operators.

Following the analysis in Sec. IV A, we need to bound the spectral norm of the nested commutators $[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]$, where $H_\gamma \in \{T, U, V\}$. This can be done by induction. In the base case, we need to estimate the norm of the kinetic operator T and the potential operators U and V. For readability, we use the abbreviated representation

$$T = \sum_{j,k} t_{j,k} A_j^\dagger A_k, \qquad U = \sum_j u_j N_j,$$
$$V = \sum_{j,k} v_{j,k} N_j N_k. \tag{F2}$$

Since $\|A_j^\dagger\| = \|A_j\| = \|N_j\| = 1$, we can apply the triangle inequality and upper bound $\|T\|$, $\|U\|$, and $\|V\|$ by the vector 1-norm $\|\vec{t}\|_1$, $\|\vec{u}\|_1$, and $\|\vec{v}\|_1$. We analyze this in Proposition F.2.

Lemma F.1: (See Eqs. (F6) and (F13) in Ref. [11]). Let an electronic-structure Hamiltonian be given as in Eq. (46). The following asymptotic analyses hold:

(1)

$$\sum_{\nu \neq 0} \frac{1}{\kappa_\nu^2} = \mathcal{O}(n). \tag{F3}$$

(2) For any fixed j,

$$\sum_\nu \kappa_\nu^2 \cos[\kappa_\nu \cdot r_j] = \mathcal{O}(1). \tag{F4}$$

(3)

$$\sum_\iota |\zeta_\iota| = \mathcal{O}(n). \tag{F5}$$

Proposition F.2: Let an electronic-structure Hamiltonian be given as in Eq. (46). We have the following bounds on the vector 1-norm and ∞-norm of the coefficients of the kinetic operator and the potential operators:

$$\|\vec{t}\|_\infty = \mathcal{O}\left(\frac{1}{n}\right), \qquad \|\vec{t}\|_1 = \mathcal{O}(n),$$
$$\|\vec{u}\|_\infty = \mathcal{O}(n), \qquad \|\vec{u}\|_1 = \mathcal{O}(n^2),$$
$$\|\vec{v}\|_\infty = \mathcal{O}(1), \qquad \|\vec{v}\|_1 = \mathcal{O}(n^2). \tag{F6}$$

Proof.—The claims about the asymptotic scaling of $\|\vec{t}\|_\infty$, $\|\vec{u}\|_\infty$, and $\|\vec{v}\|_\infty$ follow from Lemma F.1. We then obtain the scaling of the vector 1-norm from the triangle inequality. ∎

For the inductive step, we consider a general second-quantized operator of the form

$$W = \sum_{\vec{j},\vec{k},\vec{l}} w_{\vec{j},\vec{k},\vec{l}} \underbrace{\cdots (A_{j_x}^\dagger A_{k_x}) \cdots (N_{l_y}) \cdots}_{\text{at most } q \text{ operators}}, \tag{F7}$$

where $\vec{j}$, $\vec{k}$, and $\vec{l}$ denote vectors of orbitals, with total length of at most q. We keep track of the number of $A_{j_x}^\dagger A_{k_x}$ and N_{l_y} in each summand; the largest such number q is called the "layer" of W. We compute the commutator between the kinetic or potential operator and a general second-quantized operator in Proposition F.4.

Lemma F.3: (Commutation rules of second-quantized operators). The following commutation rules hold for second-quantized operators:

$$[A_j^\dagger A_k, A_l^\dagger A_m] = \delta_{kl} A_j^\dagger A_m - \delta_{jm} A_l^\dagger A_k,$$
$$[A_j^\dagger A_k, N_l] = \delta_{kl} A_j^\dagger A_l - \delta_{jl} A_l^\dagger A_k,$$
$$[A_j^\dagger A_k, N_l N_m] = (\delta_{kl} A_j^\dagger A_l - \delta_{jl} A_l^\dagger A_k) N_m + N_l (\delta_{km} A_j^\dagger A_m - \delta_{jm} A_m^\dagger A_k),$$

where δ_{kl} is the Kronecker-delta function.

Proof.—The first rule is proved by Eq. (1.8.14) in Ref. [84]. The other rules follow from the definition of the number operator $N_l = A_l^\dagger A_l$ and the commutation relation $[AB, C] = A[B, C] + [A, C]B$ for any operators A, B, and C. ∎

Proposition F.4: Let an electronic-structure Hamiltonian be given as in Eq. (46). The following statements hold for a general second-quantized operator W with q layers:

(1) $\tilde{W} = [T, W]$ is an operator with q layers and $\|\vec{\tilde{w}}\|_1 \leq 2qn \|\vec{t}\|_\infty \|\vec{w}\|_1$;
(2) $\tilde{W} = [U, W]$ is an operator with q layers and $\|\vec{\tilde{w}}\|_1 \leq 2q \|\vec{u}\|_\infty \|\vec{w}\|_1$; and
(3) $\tilde{W} = [V, W]$ is an operator with $q + 1$ layers and $\|\vec{\tilde{w}}\|_1 \leq 4qn \|\vec{v}\|_\infty \|\vec{w}\|_1$.

Proof.—For Statement 1, we have

$$\tilde{W} = [T, W] = \left[\sum_{\alpha,\beta} t_{\alpha,\beta} A_\alpha^\dagger A_\beta, \sum_{\vec{j},\vec{k},\vec{l}} w_{\vec{j},\vec{k},\vec{l}} \cdots (A_{j_x}^\dagger A_{k_x}) \cdots (N_{l_y}) \cdots\right]$$
$$= \sum_{\alpha,\beta} \sum_{\vec{j},\vec{k},\vec{l}} t_{\alpha,\beta} w_{\vec{j},\vec{k},\vec{l}} [A_\alpha^\dagger A_\beta, \cdots (A_{j_x}^\dagger A_{k_x}) \cdots (N_{l_y}) \cdots].$$

Performing the commutation sequentially, it suffices to consider

$$\cdots [A_\alpha^\dagger A_\beta, A_{j_x}^\dagger A_{k_x}] \cdots (N_{l_y}) \cdots$$
$$\cdots (A_{j_x}^\dagger A_{k_x}) \cdots [A_\alpha^\dagger A_\beta, N_{l_y}] \cdots \tag{F8}$$

For fixed $\alpha, \beta, \vec{j}, \vec{k}, \vec{l}$, there are at most q such commutators.

For the first type of commutator, we have from Lemma F.3 that

$$[A_\alpha^\dagger A_\beta, A_{j_x}^\dagger A_{k_x}] = \delta_{\beta,j_x} A_\alpha^\dagger A_{k_x} - \delta_{\alpha,k_x} A_{j_x}^\dagger A_\beta. \tag{F9}$$

Without loss of generality, consider the first term; its contribution to $\|\vec{\tilde{w}}\|_1$ is at most

$$\sum_{\alpha,\beta} \sum_{\vec{j},\vec{k},\vec{l}} \delta_{\beta,j_x} |t_{\alpha,\beta} w_{\vec{j},\vec{k},\vec{l}}| = \sum_{\alpha,\vec{j},\vec{k},\vec{l}} |t_{\alpha,j_x} w_{\vec{j},\vec{k},\vec{l}}| \leq n \|\vec{t}\|_\infty \|\vec{w}\|_1. \tag{F10}$$

Similarly, we use Lemma F.3 to analyze the second type of commutator,

$$[A_\alpha^\dagger A_\beta, N_{l_y}] = \delta_{l_y,\beta} A_\alpha^\dagger A_\beta - \delta_{l_y,\alpha} A_\alpha^\dagger A_\beta, \tag{F11}$$

and find its contribution to $\|\vec{\tilde{w}}\|_1$ as

$$\sum_{\alpha,\beta} \sum_{\vec{j},\vec{k},\vec{l}} \delta_{l_y,\beta} |t_{\alpha,\beta} w_{\vec{j},\vec{k},\vec{l}}| = \sum_{\alpha,\vec{j},\vec{k},\vec{l}} |t_{\alpha,l_y} w_{\vec{j},\vec{k},\vec{l}}| \leq n \|\vec{t}\|_\infty \|\vec{w}\|_1. \tag{F12}$$

For Statement 2, we have

$$\begin{aligned} \tilde{W} &= [U, W] \\ &= \left[\sum_\alpha u_\alpha N_\alpha, \sum_{\vec{j},\vec{k},\vec{l}} w_{\vec{j},\vec{k},\vec{l}} \cdots (A_{j_x}^\dagger A_{j_x}) \cdots (N_{l_y}) \cdots \right] \\ &= \sum_\alpha \sum_{\vec{j},\vec{k},\vec{l}} u_\alpha w_{\vec{j},\vec{k},\vec{l}} [N_\alpha, \cdots (A_{j_x}^\dagger A_{k_x}) \cdots (N_{l_y}) \cdots]. \end{aligned}$$

Performing the commutation sequentially, it suffices to consider

$$\cdots [N_\alpha, A_{j_x}^\dagger A_{k_x}] \cdots (N_{l_y}) \cdots \tag{F13}$$

For fixed $\alpha, \vec{j}, \vec{k}, \vec{l}$, there are at most q such commutators. We use Lemma F.3 again to get

$$[N_\alpha, A_{j_x}^\dagger A_{k_x}] = \delta_{\alpha,j_x} A_{j_x}^\dagger A_{k_x} - \delta_{\alpha,k_x} A_{j_x}^\dagger A_{k_x}$$

and find its contribution to $\|\vec{\tilde{w}}\|_1$ as

$$\sum_\alpha \sum_{\vec{j},\vec{k},\vec{l}} \delta_{\alpha,j_x} |u_\alpha w_{\vec{j},\vec{k},\vec{l}}| = \sum_{\vec{j},\vec{k},\vec{l}} |u_{j_x} w_{\vec{j},\vec{k},\vec{l}}| \leq \|\vec{u}\|_\infty \|\vec{w}\|_1.$$

For Statement 3, we have

$$\begin{aligned} \tilde{W} &= [V, W] \\ &= \left[\sum_{\alpha,\beta} v_{\alpha,\beta} N_\alpha N_\beta, \sum_{\vec{j},\vec{k},\vec{l}} w_{\vec{j},\vec{k},\vec{l}} \cdots (A_{j_x}^\dagger A_{k_x}) \cdots (N_{l_y}) \cdots \right] \\ &= \sum_{\alpha,\beta} \sum_{\vec{j},\vec{k},\vec{l}} v_{\alpha,\beta} w_{\vec{j},\vec{k},\vec{l}} [N_\alpha N_\beta, \cdots (A_{j_x}^\dagger A_{k_x}) \cdots (N_{l_y}) \cdots]. \end{aligned}$$

Performing the commutation sequentially, it suffices to consider

$$\cdots [N_\alpha N_\beta, A_{j_x}^\dagger A_{k_x}] \cdots (N_{l_y}) \cdots \tag{F14}$$

For fixed $\alpha, \beta, \vec{j}, \vec{k}, \vec{l}$, there are at most q such commutators. Using Lemma F.3, we have

$$\begin{aligned} [N_\alpha N_\beta, A_{j_x}^\dagger A_{k_x}] = {} & (\delta_{\alpha,j_x} A_{j_x}^\dagger A_{k_x} - \delta_{\alpha,k_x} A_{j_x}^\dagger A_{k_x}) N_\beta \\ & + N_\alpha (\delta_{\beta,j_x} A_{j_x}^\dagger A_{k_x} - \delta_{\beta,k_x} A_{j_x}^\dagger A_{k_x}). \end{aligned}$$

Without loss of generality, consider the first term; its contribution to $\|\vec{\tilde{w}}\|_1$ is at most

$$\sum_{\alpha,\beta} \sum_{\vec{j},\vec{k},\vec{l}} \delta_{\alpha,j_x} |v_{\alpha,\beta} w_{\vec{j},\vec{k},\vec{l}}| = \sum_{\beta,\vec{j},\vec{k},\vec{l}} |v_{j_x,\beta} w_{\vec{j},\vec{k},\vec{l}}| \leq n \|\vec{v}\|_\infty \|\vec{w}\|_1. \tag{F15}$$

■

Theorem F.5: [Product-formula simulation of second-quantized plane-wave electronic structure] Let $H = T + U + V$ be a second-quantized plane-wave electronic-structure Hamiltonian with n orbitals as in Eq. (46). Let $\mathcal{S}(t)$ be a pth-order product formula as in Eq. (48). Then, the Trotter error has the scaling

$$\|\mathcal{S}(t) - e^{-itH}\| = \mathcal{O}((nt)^{p+1}). \tag{F16}$$

To simulate with accuracy ϵ, it thus suffices to choose a Trotter number of

$$r = \mathcal{O}\left(\frac{(nt)^{1+1/p}}{\epsilon^{1/p}} \right). \tag{F17}$$

Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step as in Refs. [25,61], we have the gate complexity

$$n^{2+o(1)} t^{1+o(1)}. \tag{F18}$$

Proof.—We compute the scaling of the spectral norm of

$$W = [H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots], \tag{F19}$$

by induction, where $H_\gamma \in \{T, U, V\}$. In the base case where $p = 1$, we have from Proposition F.2 and Proposition F.4 that the coefficients of W have 1-norm in $\mathcal{O}(n^2)$, which implies $\|W\| = \mathcal{O}(n^2)$. For the inductive step, suppose that $W = [H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]$ is a second-quantized operator whose coefficients have vector 1-norm in $\mathcal{O}(n^p)$. Then, Proposition F.4 implies that $[T, W]$, $[U, W]$, and $[V, W]$ are second-quantized operators and their coefficients have 1-norm in $\mathcal{O}(n^{p+1})$. This proves that

$$\tilde{\alpha}_{\mathrm{comm}} = \sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\| = \mathcal{O}(n^{p+1}). \tag{F20}$$

The theorem then follows from Theorem 6 and Corollary 7. ∎

APPENDIX G: SIMULATING k-LOCAL HAMILTONIANS

In this section, we consider simulating k-local Hamiltonians using product formulas.

Recall from Sec. IV A that a k-local Hamiltonian on n qubits can be expressed as

$$H = \sum_{j_1,\ldots,j_k} H_{j_1,\ldots,j_k}, \tag{G1}$$

where each $H_{j_1,\ldots,j_k}$ acts nontrivially only on qubits labeled by $j_1, \ldots, j_k$. Our goal is to analyze the nested commutators

$$\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\|, \tag{G2}$$

where $\Gamma = n^k$ and H_{γ_j} are local operators $H_{j_1,\ldots,j_k}$. We then bound the Trotter error and the complexity of the product-formula algorithm using Theorem 6 and Corollary 7.

We claim that the operator

$$W_{\gamma_1,\ldots,\gamma_{p+1}} \equiv [H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots] \tag{G3}$$

is supported on at most $k + p(k-1)$ qubits and $\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \|W_{\gamma_1,\ldots,\gamma_{p+1}}\| = \mathcal{O}(|||H|||_1^p \|H\|_1)$, where we have used the 1-norm $\|H\|_1 = \sum_{j_1,\ldots,j_k} \|H_{j_1,\ldots,j_k}\|$ and the induced 1-norm $|||H|||_1 = \max_l \max_{j_l} \sum_{j_1,\ldots,j_{l-1},j_{l+1},\ldots,j_k} \|H_{j_1,\ldots,j_k}\|$. We prove this claim by induction on p. For $p = 1$, the commutator $W_{\gamma_1,\gamma_2} = [H_{\gamma_2}, H_{\gamma_1}]$ takes the form $[H_{j_1,\ldots,j_k}, H_{i_1,\ldots,i_k}]$, which is nonzero only when there exist $l, m = 1, \ldots, k$ such that $j_l = i_m$. It then follows that W_{γ_1,γ_2} is supported on at most $2k - 1$ qubits and that

$$\sum_{\substack{j_1,\ldots,j_k,\\ i_1,\ldots,i_k}} \|[H_{j_1,\ldots,j_k}, H_{i_1,\ldots,i_k}]\| \le 2k^2 \max_l \max_{j_l} \sum_{\substack{j_1,\ldots,j_{l-1},\\ j_{l+1},\ldots,j_k}} \|H_{j_1,\ldots,j_k}\| \sum_{i_1,\ldots,i_k} \|H_{i_1,\ldots,i_k}\| = \mathcal{O}(|||H|||_1 \|H\|_1), \tag{G4}$$

which proves the claim for $p = 1$.

Suppose that the claim holds up to $p - 1$. Following a similar argument, we have

$$\begin{aligned}
\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \|W_{\gamma_1,\ldots,\gamma_{p+1}}\| &= \sum_{j_1,\ldots,j_k} \sum_{\gamma_1,\ldots,\gamma_p=1}^{\Gamma} \|[H_{j_1,\ldots,j_k}, W_{\gamma_1,\ldots,\gamma_p}]\| \\
&\le 2k(k + (p-1)(k-1)) \cdot \max_l \max_{j_l} \sum_{\substack{j_1,\ldots,j_{l-1},\\ j_{l+1},\ldots,j_k}} \|H_{j_1,\ldots,j_k}\| \sum_{\gamma_1,\ldots,\gamma_p=1}^{\Gamma} \|W_{\gamma_1,\ldots,\gamma_p}\| \\
&= 2k(k + (p-1)(k-1)) |||H|||_1 \cdot \mathcal{O}(|||H|||_1^{p-1} \|H\|_1) \\
&= \mathcal{O}(|||H|||_1^p \|H\|_1).
\end{aligned}$$

Since the support of $H_{j_1,\ldots,j_k}$ and $W_{\gamma_1,\ldots,\gamma_p}$ overlaps, the operator $W_{\gamma_1,\ldots,\gamma_{p+1}}$ acts nontrivially on at most $k + p(k-1)$ qubits. This completes the induction.

Theorem G.1: (Product-formula simulation of k-local Hamiltonians). Let H be a k-local Hamiltonian on n qubits as in Eq. (51). Let $\mathcal{S}(t)$ be a pth-order product formula. Then, the Trotter error has the scaling

$$\|\mathcal{S}(t) - e^{-itH}\| = \mathcal{O}(|||H|||_1^p \|H\|_1). \tag{G5}$$

To simulate with accuracy ϵ, it thus suffices to choose a Trotter number of

$$r = \mathcal{O}\left(\frac{|||H|||_1 \|H\|_1^{1/p} t^{1+1/p}}{\epsilon^{1/p}}\right). \tag{G6}$$

Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step using $\mathcal{O}(n^k)$ gates, we have the gate complexity

$$n^k |||H|||_1 \|H\|_1^{o(1)} t^{1+o(1)}. \tag{G7}$$

APPENDIX H: SIMULATING POWER-LAW INTERACTIONS

In this section, we analyze the performance of product formulas for simulating power-law interactions (Sec. IV A). Let $\Lambda \subseteq \mathbb{R}^d$ be an n-qubit d-dimensional square lattice. We say that H is a power-law Hamiltonian on Λ with an exponent α if it can be written as

$$H = \sum_{\vec{i},\vec{j}\in\Lambda} H_{\vec{i},\vec{j}}, \tag{H1}$$

where $H_{\vec{i},\vec{j}}$ is an operator that acts nontrivially only on two qubits, $\vec{i}, \vec{j} \in \Lambda$, and

$$\|H_{\vec{i},\vec{j}}\| \leq \begin{cases} 1 & \text{if } \vec{i} = \vec{j} \\ \frac{1}{\|\vec{i}-\vec{j}\|_2^\alpha} & \text{if } \vec{i} \neq \vec{j}, \end{cases} \tag{H2}$$

where $\|\vec{i} - \vec{j}\|_2$ is the Euclidean distance between $\vec{i}$ and $\vec{j}$ on the lattice.

Our analysis uses the following lemma.

Lemma H.1: Given an n-qubit d-dimensional square lattice $\Lambda \subseteq \mathbb{R}^d$, it holds that

$$\sum_{\vec{j}\in\Lambda\setminus\{\vec{0}\}} \frac{1}{\|\vec{j}\|_2^\alpha} = \begin{cases} \mathcal{O}(n^{1-\alpha/d}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(\log n) & \text{for } \alpha = d \\ \mathcal{O}(1) & \text{for } \alpha > d. \end{cases} \tag{H3}$$

Furthermore, for $\alpha > d$ and $x > 0$, we have

$$\sum_{\vec{j}\in\Lambda, \|\vec{j}\|_2 \geq x} \frac{1}{\|\vec{j}\|_2^\alpha} = \mathcal{O}\left(\frac{1}{x^{\alpha-d}}\right). \tag{H4}$$

Proof.—Reference [52] provides a detailed proof of the lemma, which follows from rewriting the left-hand side of Eq. (H3) as a Riemann sum of the d-dimensional integral $\int_{\|\vec{j}\|_2\geq 1} \mathrm{d}^d\vec{j}/\|\vec{j}\|_2$. Evaluating the integral gives the right-hand side of Eq. (H3). Similarly, Eq. (H4) follows from evaluating the integral $\int_{\|\vec{j}\|_2\geq x} \mathrm{d}^d\vec{j}/\|\vec{j}\|_2 \propto (1/x^{\alpha-d})$. ∎

Theorem H.2: (Product-formula simulation of power-law interactions). Let $\Lambda \subseteq \mathbb{R}^d$ be an n-qubit d-dimensional square lattice, and let H be a power-law Hamiltonian (55) with exponent α. Let $\mathcal{S}(t)$ be a pth-order product formula. Then, the Trotter error has the scaling

$$\|\mathcal{S}(t) - e^{-itH}\| = \begin{cases} \mathcal{O}(n^{1+(p+1)(1-\alpha/d)} t^{p+1}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(n(\log n)^{p+1} t^{p+1}) & \text{for } \alpha = d \\ \mathcal{O}(nt^{p+1}) & \text{for } \alpha > d. \end{cases}$$

To simulate with accuracy ϵ, it thus suffices to choose a Trotter number of

$$r = \begin{cases} \mathcal{O}(n^{1-\frac{\alpha}{d}+\frac{1}{p}(2-\frac{\alpha}{d})} t^{1+\frac{1}{p}}/\epsilon^{\frac{1}{p}}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(n^{\frac{1}{p}}(\log n)^{1+\frac{1}{p}} t^{1+\frac{1}{p}}/\epsilon^{\frac{1}{p}}) & \text{for } \alpha = d \\ \mathcal{O}(n^{\frac{1}{p}} t^{1+\frac{1}{p}}/\epsilon^{\frac{1}{p}}) & \text{for } \alpha > d. \end{cases}$$

Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step using $\mathcal{O}(n^2)$ gates, we have the gate complexity

$$g_\alpha = \begin{cases} n^{3-\frac{\alpha}{d}+o(1)} t^{1+o(1)} & \text{for } 0 \leq \alpha < d \\ n^{2+o(1)} t^{1+o(1)} & \text{for } \alpha \geq d. \end{cases} \tag{H5}$$

Proof.—Given a power-law Hamiltonian H with exponent α, we use Lemma H.1 to compute the scaling of its induced 1-norm,

$$|||H|||_1 \leq 1 + \max_{\vec{i}} \sum_{\vec{j}\neq\vec{i}} \frac{1}{\|\vec{i}-\vec{j}\|_2^\alpha} \tag{H6}$$

$$= \begin{cases} \mathcal{O}(n^{1-\alpha/d}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(\log n) & \text{for } \alpha = d \\ \mathcal{O}(1) & \text{for } \alpha > d, \end{cases} \tag{H7}$$

and 1-norm

$$\|H\|_1 \leq \sum_{\vec{i}} \left(1 + \sum_{\vec{j}\neq\vec{i}} \frac{1}{\|\vec{i}-\vec{j}\|_2^\alpha}\right) \tag{H8}$$

$$= \begin{cases} \mathcal{O}(n^{2-\alpha/d}) & \text{for } 0 \leq \alpha < d \\ \mathcal{O}(n \log n) & \text{for } \alpha = d \\ \mathcal{O}(n) & \text{for } \alpha > d. \end{cases} \tag{H9}$$

The claim then follows from Theorem G.1. ∎

As mentioned in Sec. IV A, the performance of product formulas can be further improved for rapidly decaying power-law interactions (Theorem H.3) and quasilocal interactions (Theorem H.4).

Theorem H.3: (Product-formula simulation of rapidly decaying power-law interactions). Let $\Lambda \subseteq \mathbb{R}^d$ be an n-qubit d-dimensional square lattice, and let H be a power-law Hamiltonian (55) with exponent $\alpha > d$. Let $\mathcal{S}(t)$ be a pth-order product formula for $\tilde{H}$, the truncated Hamiltonian where summands acting on sites with distance larger than ℓ are removed. Then, the Trotter error has the scaling

$$\|\mathcal{S}(t) - e^{-it\tilde{H}}\| = \mathcal{O}(nt^{p+1}). \tag{H10}$$

To simulate with accuracy ϵ, it thus suffices to choose the cutoff $\ell = \Theta((nt/\epsilon)^{1/(\alpha-d)})$ and a Trotter number of

$$r = \mathcal{O}(n^{\frac{1}{p}} t^{1+\frac{1}{p}} / \epsilon^{\frac{1}{p}}). \tag{H11}$$

Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step using $\mathcal{O}(n\ell^d)$ gates, we have the gate complexity

$$g_\alpha = (nt)^{1+\frac{d}{\alpha-d}+o(1)}. \tag{H12}$$

Proof.—We use Lemma H.1 to bound the distance between the original and the truncated Hamiltonian,

$$\|H - \tilde{H}\| \le \sum_{\vec{i}} \sum_{\|\vec{j}-\vec{i}\|_2 > \ell} \frac{1}{\|\vec{i}-\vec{j}\|_2^\alpha} = \mathcal{O}\left(\frac{n}{\ell^{\alpha-d}}\right). \tag{H13}$$

We choose a cutoff value $\ell = \Theta((nt/\epsilon)^{1/(\alpha-d)})$, and Corollary A.5 implies that the truncation error is at most

$$\|e^{-itH} - e^{-it\tilde{H}}\| \le \|H - \tilde{H}\| t = \mathcal{O}\left(\frac{nt}{\ell^{\alpha-d}}\right) = \mathcal{O}(\epsilon). \tag{H14}$$

The theorem is then proved in a similar way as Theorem H.2. ∎

Theorem H.4: (Product-formula simulation of quasilocal interactions). Let $\Lambda \subseteq \mathbb{R}^d$ be an n-qubit d-dimensional square lattice, and let H be a quasilocal Hamiltonian (61) with constant $\beta > 0$. Let $\mathcal{S}(t)$ be a pth-order product formula for $\tilde{H}$, the truncated Hamiltonian where summands acting on sites with distance larger than ℓ are removed. Then, the Trotter error has the scaling

$$\|\mathcal{S}(t) - e^{-it\tilde{H}}\| = \mathcal{O}(nt^{p+1}). \tag{H15}$$

To simulate with accuracy ϵ, it thus suffices to choose the cutoff $\ell = \Theta(\log(nt/\epsilon))$ and a Trotter number of

$$r = \mathcal{O}\left(n^{\frac{1}{p}} t^{1+\frac{1}{p}} / \epsilon^{\frac{1}{p}}\right). \tag{H16}$$

Choosing p sufficiently large, letting ϵ be constant, and implementing each Trotter step using $\mathcal{O}(n\ell^d)$ gates, we have the gate complexity

$$g_\beta = (nt)^{1+o(1)}. \tag{H17}$$

Proof.—We choose $\ell = \Theta(\log(nt/\epsilon))$ so that the truncation error is at most

$$\begin{aligned}\|e^{-itH} - e^{-it\tilde{H}}\| &\le \|H - \tilde{H}\| t \\ &\le \sum_{\vec{i}} \sum_{\|\vec{j}-\vec{i}\|_2 > \ell} e^{-\beta\|\vec{i}-\vec{j}\|_2} \cdot t = \mathcal{O}(\epsilon).\end{aligned} \tag{H18}$$

The remaining analysis proceeds in a similar way as in Theorem H.3. ∎

APPENDIX I: SIMULATING CLUSTERED HAMILTONIANS

We continue the analysis in Sec. IV A of the hybrid algorithm for simulating clustered Hamiltonians [53]. An essential step of this algorithm is to decompose the Hamiltonian into parties using product formulas. We show that our Trotter error bound implies a more efficient decomposition and thereby gives a faster simulation of clustered Hamiltonians.

Let H be an n-qubit Hamiltonian. Assume that each term in H acts on at most two qubits with spectral norm at most one, and each qubit is interacted with at most a constant number d' of qubits. We further group the qubits into multiple parties and write

$$H = A + B = \sum_l H_l^{(1)} + \sum_l H_l^{(2)}, \tag{I1}$$

for all l: $\|H_l^{(1)}\|, \|H_l^{(2)}\| \le 1$, where terms in A act on qubits within a single party and terms in B act between two different parties.

The hybrid algorithm of Ref. [53] applies the first-order Lie-Trotter formula to decompose the Hamiltonian $H = A + \sum_l H_l^{(2)}$. Their analysis shows that a Trotter number of

$$r = \mathcal{O}\left(\frac{h_B^2 t^2}{\epsilon}\right) \tag{I2}$$

suffices to achieve error at most ϵ, where $h_B = \sum_l \|H_l^{(2)}\|$ is the interaction strength. Here, we show that it suffices to take

$$r = \mathcal{O}\left(\frac{d'^{\frac{1+p}{2}} h_B^{\frac{1}{p}} t^{1+\frac{1}{p}}}{\epsilon^{\frac{1}{p}}}\right) = \mathcal{O}\left(\frac{h_B^{1/p} t^{1+1/p}}{\epsilon^{1/p}}\right) \tag{I3}$$

using a pth-order product formula. This method improves the analysis of Ref. [53] for $p = 1$ and extends their result to higher-order cases.

In light of Theorem 6, we need to compute

$$\sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \cdots]\|, \tag{I4}$$

where each H_γ is either $H_l^{(2)}$ or A. Since $[A, A] = 0$ and $[H_\gamma, A] = -[A, H_\gamma]$, we may, without loss of generality, assume that $H_{\gamma_1} = H_{l_1}^{(2)}$, i.e.,

$$\sum_{\gamma_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_{\gamma_1}]\cdots]\| = \sum_{l_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_{l_1}^{(2)}]\cdots]\|. \quad \text{(I5)}$$

We now replace each A by $\sum_l H_l^{(1)}$ and apply the triangle inequality to get

$$\sum_{l_1,\gamma_2,\ldots,\gamma_{p+1}} \|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_{l_1}^{(2)}]\cdots]\| \le \sum_{l_1,l_2,\ldots,l_{p+1}} \|[K_{l_{p+1}},\cdots[K_{l_2},H_{l_1}^{(2)}]\cdots]\|, \quad \text{(I6)}$$

where each K_l is either $H_l^{(1)}$ or $H_l^{(2)}$. Since each qubit supports at most d' terms and each term acts on at most two qubits,

$$\sum_{l_1,l_2,\ldots,l_{p+1}} \|[K_{l_{p+1}},\cdots[K_{l_2},H_{l_1}^{(2)}]\cdots]\| = \mathcal{O}\left(d'^p\cdots d'^2 d'\sum_{l_1}\|H_{l_1}^{(2)}\|\right) = \mathcal{O}(d'^{\frac{(1+p)p}{2}}h_B).$$

We have thus established the following theorem.

Theorem I.1: (Product-formula decomposition of evolutions of clustered Hamiltonians). Let $H = A + \sum_l H_l^{(2)}$ be a clustered Hamiltonian as in Eq. (64), where each qubit is interacted with at most a constant number d' of qubits and the interaction strength is $h_B = \sum_l \|H_l^{(2)}\|$. Let $\mathcal{S}(t)$ be a pth-order product formula as in Eq. (67). Then, the Trotter error has the scaling

$$\|\mathcal{S}(t) - e^{-itH}\| = \mathcal{O}(d'^{\frac{(1+p)p}{2}}h_B t^{p+1}). \quad \text{(I7)}$$

To decompose with accuracy ϵ, it thus suffices to choose a Trotter number of

$$r = \mathcal{O}\left(\frac{h_B^{1/p}t^{1+1/p}}{\epsilon^{1/p}}\right). \quad \text{(I8)}$$

Choosing p sufficiently large, we have

$$r = \mathcal{O}\left(\frac{h_B^{o(1)}t^{1+o(1)}}{\epsilon^{o(1)}}\right). \quad \text{(I9)}$$

The hybrid simulator of Ref. [53] has runtime $2^{O(r\cdot \mathrm{cc}(g))}$, where $\mathrm{cc}(g)$ is the contraction complexity of the interaction graph g between the parties. Theorem I.1 thus gives a hybrid simulator with complexity $2^{\mathcal{O}(h_B^{o(1)}t^{1+o(1)}\mathrm{cc}(g)/\epsilon^{o(1)})}$, dramatically improving the previous result of $2^{\mathcal{O}(h_B^2 t^2 \mathrm{cc}(g)/\epsilon)}$ [53].

APPENDIX J: SIMULATING LOCAL OBSERVABLES

In this section, we analyze the performance of product formulas for simulating local observables. Following Sec. IV B, we consider a power-law Hamiltonian $H = \sum_{\vec{i},\vec{j}\in\Lambda} H_{\vec{i}\vec{j}}$ on an n-qubit d-dimensional lattice $\Lambda \subseteq \mathbb{R}^d$ with exponent $\alpha > 2d$. Our goal is to simulate the time evolution $\mathcal{A}(t) = e^{itH}Ae^{-itH}$ of a local observable A with support $\mathcal{S}(A)$ enclosed in a d-dimensional ball of constant radius x_0.

As mentioned in Sec. IV B, our approach is to construct a Hamiltonian H_{lc} whose support has radius independent of the system size. To this end, we consider a general observable B and assume that $\mathcal{S}(B)$—the support of B—is a d-dimensional ball of radius y_0 centered on the origin. We define

$$H_1 = \sum_{\vec{i},\vec{j}\in\mathcal{B}_\ell} H_{\vec{i},\vec{j}}, \quad \text{(J1)}$$

$$H_\gamma = \sum_{\vec{i},\vec{j}\in\Delta\mathcal{B}_{\gamma\ell}} H_{\vec{i},\vec{j}} + \sum_{\substack{\vec{i}\in\Delta\mathcal{B}_{(\gamma-1)\ell} \\ \vec{j}\in\Delta\mathcal{B}_{\gamma\ell}}} H_{\vec{i},\vec{j}} \quad \text{for } \gamma = 2,\ldots,\Gamma-1, \quad \text{(J2)}$$

$$H_\Gamma = \sum_{\vec{i},\vec{j}\notin\mathcal{B}_{(\Gamma-2)\ell}} H_{\vec{i},\vec{j}}, \quad \text{(J3)}$$

where $\mathcal{B}_y = \{\vec{i}\in\Lambda : \inf_{\vec{j}\in\mathcal{S}(B)}\|\vec{i}-\vec{j}\|_2 \le y\}$ is a ball of radius $y + y_0$ centered on $\mathcal{S}(B)$, $\Delta\mathcal{B}_{\gamma\ell} = \mathcal{B}_{\gamma\ell}\backslash\mathcal{B}_{(\gamma-1)\ell}$ is the shell containing sites between distance $(\gamma-1)\ell$ and $\gamma\ell$ from $\mathcal{S}(B)$, and ℓ, Γ are positive integers to be chosen later. We then define the truncated Hamiltonian $H_{\mathrm{trunc}} = \sum_{\gamma=1}^{\Gamma} H_\gamma$. We analyze the truncation error in the lemma below.

Lemma J.1: Let $\Lambda \subseteq \mathbb{R}^d$ be a d-dimensional square lattice of n qubits. Let H be a power-law Hamiltonian with exponent $\alpha > d$, and let B be an observable with support enclosed in a d-dimensional ball of radius y_0. Let $H_{\mathrm{trunc}} = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be the truncated Hamiltonian as defined above. Assuming $\Gamma = \mathcal{O}(1)$, we have

$$\|H - H_{\mathrm{trunc}}\| = \mathcal{O}\left(\frac{(y_0+\Gamma\ell)^{d-1}}{\ell^{\alpha-d-1}}\right). \quad \text{(J4)}$$

Proof.—We expand $H - H_{\mathrm{trunc}}$ as

$$H - H_{\mathrm{trunc}} = \sum_{\gamma=0}^{\Gamma-2}\sum_{\vec{i}\in\Delta\mathcal{B}_{\gamma\ell}}\sum_{\nu\ge\gamma+2}\sum_{\vec{j}\in\Delta\mathcal{B}_{\nu\ell}} H_{\vec{i},\vec{j}}. \quad \text{(J5)}$$

Applying the triangle inequality, we have

$$\|H - H_{\text{trunc}}\| \le \sum_{\gamma=0}^{\Gamma-2} \sum_{\vec{i}\in\Delta\mathcal{B}_{\gamma\ell}} \sum_{\nu\ge\gamma+2} \sum_{\vec{j}\in\Delta\mathcal{B}_{\nu\ell}} \|H_{\vec{i},\vec{j}}\| \tag{J6}$$

$$\le \sum_{\gamma=0}^{\Gamma-2} \sum_{\vec{i}\in\mathcal{B}_{\gamma\ell}} \sum_{\vec{j}\notin\mathcal{B}_{(\gamma+1)\ell}} \|H_{\vec{i},\vec{j}}\| \tag{J7}$$

$$\le \sum_{\gamma=0}^{\Gamma-2} \frac{(y_0+\Gamma\ell)^{d-1}}{\ell^{\alpha-d-1}} = \mathcal{O}\left(\frac{(y_0+\Gamma\ell)^{d-1}}{\ell^{\alpha-d-1}}\right), \tag{J8}$$

where the third inequality follows from Lemma 9 of Ref. [52] (see also the derivation of Eq. (A1) in Ref. [52], with A and C being $\mathcal{B}_{\gamma\ell}$ and the complement of $\mathcal{B}_{(\gamma+1)\ell}$, respectively). The factor $(y_0+\Gamma\ell)^{d-1}$ estimates the boundary area of $\mathcal{B}_{\gamma\ell}$. This establishes the claimed scaling of the truncation error. ∎

Next, we simulate the evolution $e^{-itH_{\text{trunc}}}$ using the pth-order product formula

$$\mathcal{S}_{\text{trunc}}(t) = \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\Gamma} e^{-ita_{(\upsilon,\gamma)}H_{\pi_\upsilon(\gamma)}}, \tag{J9}$$

where we put additional constraints on the permutation function π_ν:

$$\pi_\upsilon(1,2,3,4,5,6,\ldots) = \begin{cases} (2,4,6,\ldots,1,3,5,\ldots) & \text{if } \upsilon \text{ is odd} \\ (1,3,5,\ldots,2,4,6,\ldots) & \text{if } \upsilon \text{ is even.} \end{cases} \tag{J10}$$

By Theorem 6, the Trotter error of approximating $e^{-itH_{\text{trunc}}}$ by $\mathcal{S}_{\text{trunc}}(t)$ depends on

$$\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_{\gamma_1}]\cdots]\|, \tag{J11}$$

which we analyze in the following lemma.

Lemma J.2: Let $\Lambda \subseteq \mathbb{R}^d$ be a d-dimensional square lattice of n qubits. Let H be a power-law Hamiltonian with exponent $\alpha > d$, and let B be an observable with support enclosed in a d-dimensional ball of radius y_0. Let $H_{\text{trunc}} = \sum_{\gamma=1}^{\Gamma} H_\gamma$ be the truncated Hamiltonian as defined above. Assuming $\Gamma = \mathcal{O}(1)$, we have

$$\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \|[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_{\gamma_1}]\cdots]\| = \mathcal{O}((y_0+\Gamma\ell)^{d-1}\ell). \tag{J12}$$

Proof.—For convenience, we define

$$\mathcal{S}_1 = \{(\vec{i},\vec{j}) : \vec{i},\vec{j} \in \mathcal{B}_\ell\}, \tag{J13}$$

$$\mathcal{S}_\gamma = \{(\vec{i},\vec{j}) : \vec{i} \in \Delta\mathcal{B}_{(\gamma-1)\ell} \cup \Delta\mathcal{B}_{\gamma\ell}, \vec{j} \in \Delta\mathcal{B}_{\gamma\ell}\} \quad \text{for } \gamma = 2,\ldots,\Gamma-1, \tag{J14}$$

$$\mathcal{S}_\Gamma = \{(\vec{i},\vec{j}) : \vec{i},\vec{j} \notin \mathcal{B}_{(\Gamma-2)\ell}\}, \tag{J15}$$

so that $H_\gamma = \sum_{(\vec{i},\vec{j})\in\mathcal{S}_\gamma} H_{\vec{i},\vec{j}}$ for $\gamma = 1,\ldots,\Gamma$. Our goal is to analyze

$$\sum_{\gamma_1,\ldots,\gamma_{p+1}=1}^{\Gamma} \left\| \left[\sum_{(\vec{i}_{p+1},\vec{j}_{p+1})\in\mathcal{S}_{\gamma_{p+1}}} H_{\vec{i}_{p+1},\vec{j}_{p+1}}, \cdots \left[\sum_{(\vec{i}_2,\vec{j}_2)\in\mathcal{S}_{\gamma_2}} H_{\vec{i}_2,\vec{j}_2}, \sum_{(\vec{i}_1,\vec{j}_1)\in\mathcal{S}_{\gamma_1}} H_{\vec{i}_1,\vec{j}_1} \right] \cdots \right] \right\|.$$

Note that at least one of γ_1, γ_2 must be different from $1, \Gamma$; otherwise, $[H_{\gamma_1}, H_{\gamma_2}] = 0$. Therefore, we may assume $1 < \gamma_1 < \Gamma$ without loss of generality and bound the norm of commutators as

$$\sum_{\gamma_1=2}^{\Gamma-1} \sum_{\gamma_2,\ldots,\gamma_{p+1}=1}^{\Gamma} \sum_{\substack{(\vec{i}_1,\vec{j}_1)\in\mathcal{S}_{\gamma_1},(\vec{i}_2,\vec{j}_2)\in\mathcal{S}_{\gamma_2},\\ \ldots,(\vec{i}_{p+1},\vec{j}_{p+1})\in\mathcal{S}_{\gamma_{p+1}}}} \|[H_{\vec{i}_{p+1},\vec{j}_{p+1}},\cdots[H_{\vec{i}_2,\vec{j}_2},H_{\vec{i}_1,\vec{j}_1}]\cdots]\|. \tag{J16}$$

By a similar argument as in Theorem G.1 and Theorem H.2, we find the upper bound

$$\begin{aligned} &\mathcal{O}\left(\sum_{\gamma_1=2}^{\Gamma-1} \sum_{(\vec{i}_1,\vec{j}_1)\in\mathcal{S}_{\gamma_1}} |||H|||_1^p \|H_{\vec{i}_1,\vec{j}_1}\|\right) \\ &= \mathcal{O}\left(\sum_{\gamma_1=2}^{\Gamma-1} \sum_{\vec{j}_1\in\Delta\mathcal{B}_{\gamma_1\ell}} |||H|||_1^{p+1}\right) \\ &= \mathcal{O}((y_0+\Gamma\ell)^{d-1}\ell), \end{aligned} \tag{J17}$$

where we use the fact that $|||H|||_1 = \mathcal{O}(1)$ [Eq. (H6)] and upper bound the volume of $\Delta\mathcal{B}_{\gamma_1\ell}$ by $\mathcal{O}((y_0+\Gamma\ell)^{d-1}\ell)$—the product of its boundary area $\mathcal{O}((y_0+\Gamma\ell)^{d-1})$ with its thickness ℓ. ∎

Using the fact that product formulas can preserve the locality of the simulated system, we commute the matrix exponentials in $\mathcal{S}_{\text{trunc}}(t)$ through B to cancel with their counterpart in $\mathcal{S}^\dagger_{\text{trunc}}(t)$. By choosing $\Gamma = \Upsilon + 1$, we have

$$\mathcal{S}^\dagger_{\text{trunc}}(t) B \mathcal{S}_{\text{trunc}}(t) = \mathcal{S}^\dagger_{\text{reduce}}(t) B \mathcal{S}_{\text{reduce}}(t),$$

where

$$\mathcal{S}_{\text{reduce}}(t) = \prod_{\upsilon=1}^{\Upsilon} \prod_{\gamma=1}^{\upsilon} e^{-ita_{(\upsilon,\pi_\upsilon^{-1}(\gamma))}H_\gamma} \tag{J18}$$

is the reduced product formula of $\mathscr{A}(t)$. This method gives a decomposition of the evolution of local observable B with error

$$\|e^{itH}Be^{-itH} - \mathcal{S}^{\dagger}_{\text{reduce}}(t)B\mathcal{S}_{\text{reduce}}(t)\| = \mathcal{O}\left(\|B\|t(y_0+\Gamma\ell)^{d-1}\left(\frac{1}{\ell^{\alpha-d-1}}+\ell t^p\right)\right).$$

The remaining analysis proceeds in the same way as in Sec. IV B. We have then proved Proposition 8.

Assuming $x_0 = \mathcal{O}(1)$ and $\Gamma = \Upsilon = \mathcal{O}(1)$, we have

$$\|e^{itH}Ae^{-itH} - e^{itH_{\text{lc}}}Ae^{-itH_{\text{lc}}}\| = \mathcal{O}\left(t(r\ell)^{d-1}\left(\frac{1}{\ell^{\alpha-d-1}}+\frac{\ell t^p}{r^p}\right)\right), \tag{J19}$$

where H_{lc} is supported on a ball of radius $x = x_0 + r\Gamma\ell = \mathcal{O}(r\ell)$. To minimize the error, we choose $\ell = \Theta((r/t)^{(p/\alpha-d)}) \geq 1$, which is possible if $r \geq t$ and $\alpha > 2d$. With this choice of ℓ, the error becomes

$$\|e^{itH}Ae^{-itH} - e^{itH_{\text{lc}}}Ae^{-itH_{\text{lc}}}\| = \mathcal{O}\left(\frac{t^{\frac{p(\alpha-2d)+\alpha-d}{\alpha-d}}}{r^{\frac{p(\alpha-2d)-(\alpha-d)(d-1)}{\alpha-d}}}\right). \tag{J20}$$

To ensure that the error is at most ϵ, we choose

$$r = \Theta\left(\frac{t^{\frac{p(\alpha-2d)+\alpha-d}{p(\alpha-2d)-(\alpha-d)(d-1)}}}{\epsilon^{\frac{\alpha-d}{p(\alpha-2d)-(\alpha-d)(d-1)}}}\right). \tag{J21}$$

Note that r can be made to be greater than 1 for large times if the exponent of t in the above equation is positive; i.e., we require (assuming $\alpha > 2d$)

$$\alpha > \frac{2d-\frac{d(d-1)}{p}}{1-\frac{d-1}{p}} \Leftrightarrow p > \frac{(\alpha-d)(d-1)}{\alpha-2d}. \tag{J22}$$

In addition, the choice of r above is also consistent with the condition $r \geq t$ because

$$\frac{p(\alpha-2d)+\alpha-d}{p(\alpha-2d)-(\alpha-d)(d-1)} > 1. \tag{J23}$$

Recall that $e^{-itH_{\text{lc}}}$ is an evolution supported on a ball of radius $x = \mathcal{O}(r\ell)$. Invoking Theorem H.3, we obtain the gate count

$$g_\alpha = \mathcal{O}\left((x^d t)^{1+\frac{1}{p}+\frac{d}{\alpha-d}}\right) = \mathcal{O}\left(t^{\frac{(\alpha(p+1)-d)(\alpha(dp+p+1)-(d+2)dp-d)}{p(\alpha-d)(\alpha+d^2-d(\alpha+2p+1)+\alpha p)}}\right) \tag{J24}$$

for simulating local observable A with constant accuracy, which simplifies to

$$g_\alpha = t^{\frac{\alpha(\alpha(d+1)-(d+2)d)}{(\alpha-d)(\alpha-2d)}+o(1)} = t^{(1+d\frac{\alpha-d}{\alpha-2d})(1+\frac{d}{\alpha-d})+o(1)} \tag{J25}$$

in the large p limit. The remaining analysis proceeds as in Sec. IV B.

APPENDIX K: QUANTUM MONTE CARLO SIMULATION

In this section, we apply our Trotter error bound to improve the performance of quantum Monte Carlo simulation. This analysis is sketched in Sec. IV C and detailed here.

We will use the following lemma.

Lemma K.1: (Relative perturbation of eigenvalues in Theorem 2.1 of Ref. [85] and Theorem 5.4 of Ref. [86]). Let matrix C be positive semidefinite, and let D be nonsingular. Assume that the eigenvalues $\lambda_i(C)$ and $\lambda_i(D^\dagger CD)$ are ordered nonincreasingly. Then,

$$\lambda_i(D^\dagger CD) \leq \lambda_i(C)\|D^\dagger D\|. \tag{K1}$$

1. Transverse field Ising Hamiltonian

We first consider simulating an n-qubit transverse field Ising Hamiltonian $H = -A - B$ [38], where

$$A = \sum_{1\leq u<v\leq n} j_{u,v}Z_uZ_v, \qquad B = \sum_{1\leq u\leq n} h_uX_u. \tag{K2}$$

Note that X_u and Z_u are Pauli operators acting on the uth qubit, and $j_{u,v} \geq 0$ and $h_u \geq 0$ are non-negative coefficients. Our goal is to approximate the partition function $\mathcal{Z} = \text{Tr}(e^{-H})$ up to multiplicative error $0 < \epsilon < 1$.

A key step in the algorithm of Ref. [38] is to decompose the evolution operator using the second-order Suzuki formula. Note that all the summands in A (or B) commute with each other, so no error is introduced when the evolution under A (or B) is further decomposed into elementary exponentials. It thus suffices to analyze the Trotter error of approximating $e^{t(A+B)}$ by $(e^{(t/2r)A}e^{(t/r)B}e^{(t/2r)A})^r$ for time $t > 0$ and Trotter number r. To this end, we define

$$U := e^{\frac{t}{r}(A+B)}, \tag{K3}$$

$$V := e^{\frac{t}{2r}A}e^{\frac{t}{r}B}e^{\frac{t}{2r}A}, \tag{K4}$$

$$W := \exp_{\mathcal{T}}\left[\int_0^{\frac{t}{r}} \mathrm{d}\tau\, e^{-\tau\text{ad}_{A+B}}\left(e^{\frac{\tau}{2}\text{ad}_A}B - B + e^{\frac{\tau}{2}\text{ad}_A}e^{\tau\text{ad}_B}\frac{A}{2} - \frac{A}{2}\right)\right], \tag{K5}$$

so Theorem 3 implies $V = UW$.

Our goal is to bound the difference between the eigenvalues $\lambda_i(V^r)$ and $\lambda_i(U^r)$, which can be done recursively as follows. We first replace the rightmost V by UW and the leftmost V by $W^\dagger U$. Invoking Lemma K.1, we have

$$\lambda_i(V^r) = \lambda_i(W^\dagger U V^{r-2} U W) \le \lambda_i(U V^{r-2} U)\|W\|^2.$$

By Theorem 1.3.22 of Ref. [54],

$$\lambda_i(UV^{r-2}U) = \lambda_i(V^{\frac{r}{2}-1} U U V^{\frac{r}{2}-1}). \tag{K6}$$

We now apply a similar procedure to obtain

$$\begin{aligned}
&\lambda_i(V^{\frac{r}{2}-1} U U V^{\frac{r}{2}-1}) \\
&= \lambda_i(W^\dagger U V^{\frac{r}{2}-2} U U V^{\frac{r}{2}-2} U W) \\
&\le \lambda_i(U V^{\frac{r}{2}-2} U U V^{\frac{r}{2}-2} U)\|W\|^2 \\
&= \lambda_i(V^{\frac{r}{4}-1} U U V^{\frac{r}{2}-2} U U V^{\frac{r}{4}-1})\|W\|^2 \\
&\le \lambda_i(U V^{\frac{r}{4}-2} U U V^{\frac{r}{2}-2} U U V^{\frac{r}{4}-2} U)\|W\|^4 \\
&= \lambda_i(V^{\frac{r}{4}-1} U U V^{\frac{r}{4}-2} U U V^{\frac{r}{4}-2} U U V^{\frac{r}{4}-1})\|W\|^4 \\
&\le \lambda_i(U V^{\frac{r}{4}-2} U U V^{\frac{r}{4}-2} U U V^{\frac{r}{4}-2} U U V^{\frac{r}{4}-2} U)\|W\|^6.
\end{aligned}$$

To ensure that this recursion is valid, we choose r to be a power of 2. Since any positive integer is between 2^m and 2^{m+1} for some $m \ge 0$, this choice only enlarges r by a factor of at most 2. Overall, we have

$$\lambda_i(V^r) \le \lambda_i(U^r)\|W\|^r. \tag{K7}$$

To analyze the operator W, we further compute

$$\begin{aligned}
&e^{\frac{\tau}{2}\mathrm{ad}_A} B - B \\
&\quad = \tau \mathrm{ad}_{\frac{A}{2}} B + \int_0^\tau \mathrm{d}\tau_2 \int_0^{\tau_2} \mathrm{d}\tau_3 e^{\frac{\tau_3}{2}\mathrm{ad}_A} \mathrm{ad}^2_{\frac{A}{2}} B,
\end{aligned} \tag{K8}$$

and

$$\begin{aligned}
&e^{\frac{\tau}{2}\mathrm{ad}_A} e^{\tau \mathrm{ad}_B} \frac{A}{2} - \frac{A}{2} \\
&\quad = e^{\frac{\tau}{2}\mathrm{ad}_A} \left(\int_0^\tau \mathrm{d}\tau_2 e^{\tau_2 \mathrm{ad}_B} \mathrm{ad}_B \frac{A}{2} \right) \\
&\quad = e^{\frac{\tau}{2}\mathrm{ad}_A} \left(\tau \mathrm{ad}_B \frac{A}{2} + \int_0^\tau \mathrm{d}\tau_2 \int_0^{\tau_2} \mathrm{d}\tau_3 e^{\tau_3 \mathrm{ad}_B} \mathrm{ad}^2_B \frac{A}{2} \right) \\
&\quad = \tau \mathrm{ad}_B \frac{A}{2} + \tau \int_0^\tau \mathrm{d}\tau_2 e^{\frac{\tau_2}{2}\mathrm{ad}_A} \mathrm{ad}^2_B \frac{A}{2} \\
&\qquad + \int_0^\tau \mathrm{d}\tau_2 \int_0^{\tau_2} \mathrm{d}\tau_3 e^{\tau_3 \mathrm{ad}_B} \mathrm{ad}^2_B \frac{A}{2}.
\end{aligned} \tag{K9}$$

By Lemma A.4, we have

$$\begin{aligned}
\|W\| \le \exp\Bigg(& e^{2\frac{t}{r}\|H\| + \frac{t}{r}\|A\|} \frac{t^3}{24 r^3} \|\mathrm{ad}^2_A B\| \\
& + e^{2\frac{t}{r}\|H\| + \frac{t}{r}\|A\|} \frac{t^3}{6 r^3} \|\mathrm{ad}^2_B A\| \\
& + e^{2\frac{t}{r}\|H\| + 2\frac{t}{r}\|B\|} \frac{t^3}{12 r^3} \|\mathrm{ad}^2_B A\| \Bigg).
\end{aligned} \tag{K10}$$

This bound is tighter than the previous result of Lemma of Ref. [38] in that it exploits the commutativity of operator summands. For the transverse field Ising model, this leads to an asymptotic improvement on the performance of Monte Carlo simulation. The remaining analysis proceeds as in Sec. IV C.

2. Ferromagnetic spin systems

In this section, we consider simulating partition functions of ferromagnetic quantum spin systems,

$$\begin{aligned}
H = & \sum_{1 \le u < v \le n} p_{uv}(-X_u X_v - Y_u Y_v) \\
& + \sum_{1 \le u < v \le n} q_{uv}(-X_u X_v + Y_u Y_v) + \sum_{u=1}^n d_u (I + Z_u),
\end{aligned} \tag{K11}$$

where $p_{uv}, q_{uv} \in [0,1]$. Our goal is to approximate the partition function $\mathcal{Z}(\beta, H) = \mathrm{Tr}[e^{-\beta H}]$ for $\beta > 0$. Following Ref. [39], we restrict ourselves to the n-qubit (nonunitary) gate set

$$\begin{aligned}
&\left\{ f_u(e^{\pm t}), g_{uv}(t), h_{uv}(t) | u, v = 1, \ldots, n, \right. \\
&\qquad \left. u \ne v, 0 < t < \frac{1}{2} \right\},
\end{aligned} \tag{K12}$$

where

$$\begin{aligned}
f(e^{\pm t}) &= \begin{bmatrix} e^{\pm t} & 0 \\ 0 & 1 \end{bmatrix}, \qquad g(t) = \begin{bmatrix} 1+t^2 & 0 & 0 & t \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ t & 0 & 0 & 1 \end{bmatrix}, \\
h(t) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1+t^2 & t & 0 \\ 0 & t & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned} \tag{K13}$$

and the subscripts u, v indicate the qubits on which the gates act nontrivially. These gates approximate the exponentials of terms of the original Hamiltonian. Specifically, we represent the gates as

$$f_u(e^{\pm t}) = e^{\pm\frac{t}{2}F_u}, \qquad g_{uv}(t) = e^{-\frac{t}{2}\tilde{\mathcal{G}}_{uv}(t)}, \qquad h_{uv}(t) = e^{-\frac{t}{2}\tilde{\mathcal{H}}_{uv}(t)}, \tag{K14}$$

where $0 < t < 1/2$ and

$$F_i = (I + Z_u), \quad \tilde{\mathcal{G}}_{uv}(t) = (-X_u X_v + Y_u Y_v) - \frac{2}{t}\mathcal{G}_{uv}(t), \quad \tilde{\mathcal{H}}_{uv}(t) = (-X_u X_v - Y_u Y_v) - \frac{2}{t}\mathcal{H}_{uv}(t). \tag{K15}$$

By Proposition 1 of Ref. [39], we have $\|\mathcal{G}_{uv}(t)\| \le t^2$ and $\|\mathcal{H}_{uv}(t)\| \le t^2$.

We divide the evolution into r steps. We choose $r > 2\beta$ so that we can implement the product formula using gates from Eq. (98) with parameters

$$-\frac{1}{2} < -\frac{\beta}{r}d_u < \frac{1}{2}, \qquad 0 < \frac{\beta}{r}q_{uv} < \frac{1}{2}, \qquad 0 < \frac{\beta}{r}p_{uv} < \frac{1}{2}. \tag{K16}$$

Consider the gate sequence

$$\begin{aligned}
&\prod_{1\le u\le n} f_u(e^{-\frac{\beta}{r}d_u}) \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \cdot \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \prod_{1\le u\le n} f_u(e^{-\frac{\beta}{r}d_u}) \\
&= \prod_{1\le u\le n} e^{-\frac{\beta}{2r}d_u F_u} \prod_{1\le u<v\le n} e^{-\frac{\beta}{2r}q_{uv}\tilde{\mathcal{G}}_{uv}(\frac{\beta}{r}q_{uv})} \prod_{1\le u<v\le n} e^{-\frac{\beta}{2r}p_{uv}\tilde{\mathcal{H}}_{uv}(\frac{\beta}{r}p_{uv})} \cdot \prod_{1\le u<v\le n} e^{-\frac{\beta}{2r}p_{uv}\tilde{\mathcal{H}}_{uv}(\frac{\beta}{r}p_{uv})} \prod_{1\le u<v\le n} e^{-\frac{\beta}{2r}q_{uv}\tilde{\mathcal{G}}_{uv}(\frac{\beta}{r}q_{uv})} \prod_{1\le u\le n} e^{-\frac{\beta}{2r}d_u F_u} \\
&= \exp\left(-\frac{\beta}{r}\left(\sum_{1\le u<v\le n} p_{uv}\tilde{\mathcal{H}}_{uv}\left(\frac{\beta}{r}p_{uv}\right) + \sum_{1\le u<v\le n} q_{uv}\tilde{\mathcal{G}}_{uv}\left(\frac{\beta}{r}q_{uv}\right) + \sum_{u=1}^{n} d_u F_u\right)\right) \cdot W,
\end{aligned} \tag{K17}$$

which implements the second-order Suzuki formula, where we have applied Theorem 3 in the last line. Since

$$\|F_u\| \le 2, \qquad \left\|\tilde{\mathcal{G}}_{uv}\left(\frac{\beta}{r}q_{uv}\right)\right\| \le 2 + 2\frac{\beta}{r}q_{uv} \le 3, \qquad \left\|\tilde{\mathcal{H}}_{uv}\left(\frac{\beta}{r}p_{uv}\right)\right\| \le 2 + 2\frac{\beta}{r}p_{uv} \le 3, \tag{K18}$$

the perturbed Hamiltonian satisfies

$$\sum_{1\le u<v\le n} p_{uv}\left\|\tilde{\mathcal{H}}_{uv}\left(\frac{\beta}{r}p_{uv}\right)\right\| + \sum_{1\le u<v\le n} q_{uv}\left\|\tilde{\mathcal{G}}_{uv}\left(\frac{\beta}{r}q_{uv}\right)\right\| + \sum_{u=1}^{n}|d_u|\|F_u\| \le \binom{n}{2}3 + \binom{n}{2}3 + 2n \le 3n^2. \tag{K19}$$

We also need to bound nested commutators of Hamiltonian terms with two layers of nesting. This analysis is similar to that for the transverse field Ising model; the resulting scaling is $\mathcal{O}(n^4)$. By Theorem 6, there exists a constant $c > 0$ such that

$$\|W\| \le \exp\left(\frac{cn^4\beta^3}{r^3}e^{\frac{12n^2\beta}{r}}\right). \tag{K20}$$

To proceed, we apply Lemma A.2 to switch to the interaction picture, giving

$$\begin{aligned}
&\exp\left(-\frac{\beta}{r}\left(\sum_{1\le u<v\le n} p_{uv}\tilde{\mathcal{H}}_{uv}\left(\frac{\beta}{r}p_{uv}\right) + \sum_{1\le u<v\le n} q_{uv}\tilde{\mathcal{G}}_{uv}\left(\frac{\beta}{r}q_{uv}\right) + \sum_{u=1}^{n} d_u F_u\right)\right) \\
&= e^{-\frac{\beta}{r}H}V,
\end{aligned} \tag{K21}$$

where

$$V = \exp_{\mathcal{T}}\left(-\int_0^{\frac{\beta}{r}} \mathrm{d}\tau e^{\tau H}\left(\sum_{1\le u<v\le n} p_{uv}\tilde{\mathscr{H}}_{uv}\left(\frac{\beta}{r}p_{uv}\right) + \sum_{1\le u<v\le n} q_{uv}\tilde{\mathcal{G}}_{uv}\left(\frac{\beta}{r}q_{uv}\right) + \sum_{u=1}^{n} d_u F_u - H\right)e^{-\tau H}\right).$$

From Eq. (K15),

$$\begin{aligned}
&\left\|\sum_{1\le u<v\le n} p_{uv}\tilde{\mathscr{H}}_{uv}\left(\frac{\beta}{r}p_{uv}\right) + \sum_{1\le u<v\le n} q_{uv}\tilde{\mathcal{G}}_{uv}\left(\frac{\beta}{r}q_{uv}\right) + \sum_{u=1}^{n} d_u F_u - H\right\| = \left\|\sum_{1\le u<v\le n} p_{uv}\frac{2r}{\beta p_{uv}}\mathscr{H}_{uv}\left(\frac{\beta}{r}p_{uv}\right) + \sum_{1\le u<v\le n} q_{uv}\frac{2r}{\beta q_{uv}}\mathcal{G}_{uv}\left(\frac{\beta}{r}q_{uv}\right)\right\| \\
&\le \binom{n}{2}2\frac{\beta}{r}p_{uv} + \binom{n}{2}2\frac{\beta}{r}q_{uv} = 2n^2\frac{\beta}{r},
\end{aligned}$$

whereas the original Hamiltonian has spectral norm

$$\|H\| \le \sum_{1\le u<v\le n} p_{uv}\|-X_uX_v - Y_uY_v\| + \sum_{1\le u<v\le n} q_{uv}\|-X_uX_v + Y_uY_v\| + \sum_{u=1}^{n} |d_u|\|I+Z_u\| \tag{K22}$$

$$\le \binom{n}{2}\cdot 2 + \binom{n}{2}\cdot 2 + n\cdot 2 = 2n^2. \tag{K23}$$

Thus, Lemma A.4 implies

$$\|V\| \le \exp\left(\frac{2n^2\beta^2}{r^2} e^{\frac{4n^2\beta}{r}}\right). \tag{K24}$$

Altogether, we obtain

$$\prod_{1\le u\le n} f_u\left(e^{-\frac{\beta}{r}d_u}\right) \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \cdot \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \prod_{1\le u<v\le n} h_{uv}\left(\frac{\beta}{r}p_{uv}\right) \cdot \prod_{1\le u<v\le n} g_{uv}\left(\frac{\beta}{r}q_{uv}\right) \prod_{1\le u\le n} f_u\left(e^{-\frac{\beta}{r}d_u}\right) = e^{-\frac{\beta}{r}H}U, \tag{K25}$$

where the operator $U = VW$ has spectral norm bounded by

$$\|U\| = \|VW\| \le \exp\left(\frac{2n^2\beta^2}{r^2} e^{\frac{4n^2\beta}{r}} + \frac{cn^4\beta^3}{r^3} e^{\frac{12n^2\beta}{r}}\right) \tag{K26}$$

for some constant $c > 0$. The remaining analysis continues as in Sec. IV C. Similar to the case of the transverse field Ising model, our Trotter error bound gives improved quantum Monte Carlo simulation of the ferromagnetic quantum spin systems.

APPENDIX L: SECOND-ORDER ERROR BOUND WITH SMALL PREFACTOR

In this section, we prove the tight error bound for the second-order Suzuki formula in Proposition 10.

For the two-term case, our goal is to decompose the evolution $e^{-itH} = e^{-it(A+B)}$ using the product formula $\mathcal{S}_2(t) = e^{-i(t/2)A}e^{-itB}e^{-i(t/2)A}$. Using the variation-of-parameters formula (Lemma A.1), we have

$$\mathcal{S}_2(t) = e^{-itH} + \int_0^t \mathrm{d}\tau_1 e^{-i(t-\tau_1)H} e^{-i\frac{\tau_1}{2}A} \cdot \mathcal{T}_2(\tau_1) e^{-\tau_1 B} e^{-i\frac{\tau_1}{2}A}, \tag{L1}$$

where

$$\mathcal{T}_2(\tau_1) = e^{-i\tau_1 B}\left(-i\frac{A}{2}\right)e^{i\tau_1 B} + i\frac{A}{2} + e^{i\frac{\tau_1}{2}A}(iB)e^{-i\frac{\tau_1}{2}A} - iB. \tag{L2}$$

By Theorem 4 or a direct calculation, we find the order condition $\mathcal{T}_2(\tau_1) = O(\tau_1^2)$, which implies

$$\mathcal{T}_2(\tau_1) = \int_0^{\tau_1}\mathrm{d}\tau_2\int_0^{\tau_2}\mathrm{d}\tau_3\left[e^{-i\tau_3\mathrm{ad}_B}\mathrm{ad}_{-iB}^2\left(-i\frac{A}{2}\right) + e^{i\frac{\tau_3}{2}\mathrm{ad}_A}\mathrm{ad}_{i\frac{A}{2}}^2(iB)\right]. \tag{L3}$$

Altogether, we have the representation

$$\mathcal{S}_2(t) = e^{-itH} + \int_0^t \mathrm{d}\tau_1\int_0^{\tau_1}\mathrm{d}\tau_2\int_0^{\tau_2}\mathrm{d}\tau_3 e^{-i(t-\tau_1)H} \cdot e^{-i\frac{\tau_1}{2}A}\left[e^{-i\tau_3\mathrm{ad}_B}\mathrm{ad}_{-iB}^2\left(-i\frac{A}{2}\right) + e^{i\frac{\tau_3}{2}\mathrm{ad}_A}\mathrm{ad}_{i\frac{A}{2}}^2(iB)\right]e^{-\tau_1 B}e^{-i\frac{\tau_1}{2}A}, \tag{L4}$$

and the error bound for $t \ge 0$,

$$\|\mathcal{S}_2(t) - e^{-itH}\| \le \frac{t^3}{12}\|[B,[B,A]]\| + \frac{t^3}{24}\|[A,[A,B]]\|. \tag{L5}$$

For a general Hamiltonian $H = \sum_{\gamma=1}^{\Gamma} H_\gamma$, we apply the triangle inequality to get

$$\left\|\prod_{\gamma=\Gamma}^{1} e^{-i\frac{t}{2}H_\gamma}\prod_{\gamma=1}^{\Gamma} e^{-i\frac{t}{2}H_\gamma} - e^{-it\sum_{\gamma=1}^{\Gamma}H_\gamma}\right\| \le \sum_{\gamma_1=1}^{\Gamma}\left\|\prod_{\gamma_2=\gamma_1}^{1} e^{-i\frac{t}{2}H_{\gamma_2}}e^{-it\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}}\prod_{\gamma_2=1}^{\gamma_1} e^{-i\frac{t}{2}H_{\gamma_2}} - \prod_{\gamma_2=\gamma_1-1}^{1} e^{-i\frac{t}{2}H_{\gamma_2}}e^{-it\sum_{\gamma_2=\gamma_1}^{\Gamma}H_{\gamma_2}}\prod_{\gamma_2=1}^{\gamma_1-1} e^{-i\frac{t}{2}H_{\gamma_2}}\right\|$$
$$= \sum_{\gamma_1=1}^{\Gamma}\left\|e^{-i\frac{t}{2}H_{\gamma_1}}e^{-it\sum_{\gamma_2=\gamma_1+}^{\Gamma}H_{\gamma_2}}e^{-i\frac{t}{2}H_{\gamma_1}} - e^{-it\sum_{\gamma_2=\gamma_1}^{\Gamma}H_{\gamma_2}}\right\|$$
$$\le \frac{t^3}{12}\sum_{\gamma_1=1}^{\Gamma}\left\|\left[\sum_{\gamma_3=\gamma_1+1}^{\Gamma}H_{\gamma_3},\left[\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2},H_{\gamma_1}\right]\right]\right\| + \frac{t^3}{24}\sum_{\gamma_1=1}^{\Gamma}\left\|\left[H_{\gamma_1},\left[H_{\gamma_1},\sum_{\gamma_2=\gamma_1+1}^{\Gamma}H_{\gamma_2}\right]\right]\right\|. \tag{L6}$$

This completes the proof of Proposition 10.

APPENDIX M: HIGHER-ORDER ERROR BOUNDS WITH SMALL PREFACTORS

We have shown in Sec. V A that our analysis reproduces known tight error bounds for first- and second-order formulas. In this section, we give heuristic strategies to derive higher-order Trotter error bounds with small prefactors. We illustrate this for the fourth-order formula, which is advantageous for simulating small-size systems [31] but does not have a tight error analysis. We further benchmark our bounds in Sec. V B by numerically simulating systems with nearest-neighbor and power-law interactions. Throughout this section, we assume H is Hermitian, $t \in \mathbb{R}$, and consider the real-time evolution e^{-itH}.

We first consider a Hamiltonian $H = A + B$ consisting of two summands. This Hamiltonian models systems with nearest-neighbor interactions where summands are grouped in an even-odd pattern [Eq. (124)]. The ideal evolution under H for time t is e^{-itH}, which we decompose using the fourth-order product formula $\mathcal{S}_4(t)$. Recall from Eq. (11) that $\mathcal{S}_4(t)$ is defined by

$$\begin{aligned}\mathcal{S}_2(t) &:= e^{-i\frac{t}{2}A} e^{-itB} e^{-i\frac{t}{2}A},\\ \mathcal{S}_4(t) &:= [\mathcal{S}_2(u_2 t)]^2 \mathcal{S}_2((1-4u_2)t)[\mathcal{S}_2(u_2 t)]^2,\end{aligned} \tag{M1}$$

with $u_2 := 1/(4-4^{1/3})$. Expanding this definition, we obtain

$$\begin{aligned}\mathcal{S}_4(t) = {} & e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A}\\ & \cdot e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A},\end{aligned}$$

where

$$\begin{aligned}a_1 &:= a_6 := \frac{u_2}{2},\\ b_1 &:= a_2 := b_2 := b_4 := a_5 := b_5 := u_2,\\ a_3 &:= a_4 := \frac{1-3u_2}{2},\\ b_3 &:= 1-4u_2.\end{aligned} \tag{M2}$$

Without loss of generality, we analyze the additive Trotter error of $\mathcal{S}_4(t)$. In Sec. III B, we gave an analysis that works for a general product formula, and we improve that here to obtain an error bound for $\mathcal{S}_4(t)$ with a small prefactor.

To this end, we compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_4(t) - (-iH)\mathcal{S}_4(t) = {} & [e^{-ita_6 A}, -ib_5 B] e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + [e^{-ita_6 A} e^{-itb_5 B}, -ia_5 A] e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A} + \cdots\\ & + [e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A}, -ib_1 B] e^{-itb_1 B} e^{-ita_1 A}\\ & + [e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B}, -ia_1 A] e^{-ita_1 A}.\end{aligned} \tag{M3}$$

Performing the commutation sequentially, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_4(t) - (-iH)\mathcal{S}_4(t) = {} & [e^{-ita_6 A}, -ib_5 B] e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A}[e^{-itb_5 B}, -ia_5 A] e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A} + \cdots\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B}[e^{-ita_2 A}, -ib_1 B] e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B}[e^{-ita_3 A}, -ib_1 B] e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B}[e^{-ita_4 A}, -ib_1 B] e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B}[e^{-ita_5 A}, -ib_1 B] e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + [e^{-ita_6 A}, -ib_1 B] e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A}[e^{-itb_1 B}, -ia_1 A] e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A}[e^{-itb_2 B}, -ia_1 A] e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A}[e^{-itb_3 B}, -ia_1 A] e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A} e^{-itb_5 B} e^{-ita_5 A}[e^{-itb_4 B}, -ia_1 A] e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}\\ & + e^{-ita_6 A}[e^{-itb_5 B}, -ia_1 A] e^{-ita_5 A} e^{-itb_4 B} e^{-ita_4 A} e^{-itb_3 B} e^{-ita_3 A} e^{-itb_2 B} e^{-ita_2 A} e^{-itb_1 B} e^{-ita_1 A}.\end{aligned} \tag{M4}$$

We further define

$$\begin{aligned} c_1 &:= a_1, \qquad d_1 := b_1, \\ c_2 &:= a_1 + a_2, \qquad d_2 := b_1 + b_2, \\ c_3 &:= a_1 + a_2 + a_3, \qquad d_3 := b_1 + b_2 + b_3, \\ c_4 &:= a_1 + a_2 + a_3 + a_4, \qquad d_4 := b_1 + b_2 + b_3 + b_4, \\ c_5 &:= a_1 + a_2 + a_3 + a_4 + a_5, \qquad d_5 := b_1 + b_2 + b_3 + b_4 + b_5, \end{aligned} \tag{M5}$$

so that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_4(t) - (-iH)\mathcal{S}_4(t) = {} & [e^{-ita_6A}, -id_5B]e^{-itb_5B}e^{-ita_5A}e^{-itb_4B}e^{-ita_4A}e^{-itb_3B}e^{-ita_3A}e^{-itb_2B}e^{-ita_2A}e^{-itb_1B}e^{-ita_1A} \\ & + e^{-ita_6A}[e^{-itb_5B}, -ic_5A]e^{-ita_5A}e^{-itb_4B}e^{-ita_4A}e^{-itb_3B}e^{-ita_3A}e^{-itb_2B}e^{-ita_2A}e^{-itb_1B}e^{-ita_1A} + \cdots \\ & + e^{-ita_6A}e^{-itb_5B}e^{-ita_5A}e^{-itb_4B}e^{-ita_4A}e^{-itb_3B}e^{-ita_3A}e^{-itb_2B}[e^{-ita_2A}, -id_1B]e^{-itb_1B}e^{-ita_1A} \\ & + e^{-ita_6A}e^{-itb_5B}e^{-ita_5A}e^{-itb_4B}e^{-ita_4A}e^{-itb_3B}e^{-ita_3A}e^{-itb_2B}e^{-ita_2A}[e^{-itb_1B}, -ic_1A]e^{-ita_1A}. \end{aligned} \tag{M6}$$

In Sec. III B and Appendix C, we factor out the operator-valued function $\mathcal{S}_4(t)$ from the left-hand side of the above equation as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_4(t) - (-iH)\mathcal{S}_4(t) = \mathcal{S}_4(t)\mathcal{T}(t). \tag{M7}$$

This approach suffices to establish the asymptotic bound in Theorem 6 and Corollary 7. However, the resulting function $\mathcal{T}(t)$ contains unitary conjugations with a large number of conjugating layers, which defeats the goal of establishing tight error bounds. We improve this by simultaneously factoring out $\mathcal{S}_{4,\mathrm{left}}(t)$ from the left-hand side of the equation and $\mathcal{S}_{4,\mathrm{right}}(t)$ from the right-hand side, obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_4(t) - (-iH)\mathcal{S}_4(t) = \mathcal{S}_{4,\mathrm{left}}(t)\mathcal{T}_4(t)\mathcal{S}_{4,\mathrm{right}}(t), \tag{M8}$$

where

$$\begin{aligned} \mathcal{S}_{\mathrm{left}}(t) &:= e^{-ita_6A}e^{-itb_5B}e^{-ita_5A}e^{-itb_4B}e^{-ita_4A}, \\ \mathcal{S}_{\mathrm{right}}(t) &:= e^{-itb_3B}e^{-ita_3A}e^{-itb_2B}e^{-ita_2A}e^{-itb_1B}e^{-ita_1A}. \end{aligned} \tag{M9}$$

It then remains to analyze $\mathcal{T}_4(t)$.

To this end, we use the fact that

$$\begin{aligned} [e^{tX}, Y] &= e^{tX}\int_0^t \mathrm{d}\tau\, e^{-\tau X}[X, Y]e^{\tau X} \\ &= \int_0^t \mathrm{d}\tau\, e^{\tau X}[X, Y]e^{-\tau X}e^{tX}, \end{aligned} \tag{M10}$$

for any $t \in \mathbb{R}$ and operators X, Y. We then have from Lemma A.1 that

$$\mathcal{S}_4(t) = e^{-itH} + \int_0^t \mathrm{d}\tau_1\, e^{-i(t-\tau_1)H}\mathcal{S}_{4,\mathrm{left}}(\tau_1)\mathcal{T}_4(\tau_1)\mathcal{S}_{4,\mathrm{right}}(\tau_1),$$

where

$$\begin{aligned} \mathcal{T}_4(\tau_1) = {} & \int_0^{\tau_1} \mathrm{d}\tau_2 e^{i\tau_1 a_4A}e^{i\tau_1 b_4B}e^{i\tau_1 a_5A}e^{i\tau_1 b_5B}e^{i\tau_2 a_6A}[-ia_6A, -id_5B]e^{-i\tau_2 a_6A}e^{-i\tau_1 b_5B}e^{-i\tau_1 a_5A}e^{-i\tau_1 b_4B}e^{-i\tau_1 a_4A} \\ & + \int_0^{\tau_1} \mathrm{d}\tau_2 e^{i\tau_1 a_4A}e^{i\tau_1 b_4B}e^{i\tau_1 a_5A}e^{i\tau_2 b_5B}[-ib_5B, -ic_5A]e^{-i\tau_2 b_5B}e^{-i\tau_1 a_5A}e^{-i\tau_1 b_4B}e^{-i\tau_1 a_4A} \\ & + \int_0^{\tau_1} \mathrm{d}\tau_2 e^{i\tau_1 a_4A}e^{i\tau_1 b_4B}e^{i\tau_2 a_5A}[-ia_5A, -id_4B]e^{-i\tau_2 a_5A}e^{-i\tau_1 b_4B}e^{-i\tau_1 a_4A} \\ & + \int_0^{\tau_1} \mathrm{d}\tau_2 e^{i\tau_1 a_4A}e^{i\tau_2 b_4B}[-ib_4B, -ic_4A]e^{-i\tau_2 b_4B}e^{-i\tau_1 a_4A} + \int_0^{\tau_1} \mathrm{d}\tau_2 e^{i\tau_2 a_4A}[-ia_4A, -id_3B]e^{-i\tau_2 a_4A} \\ & + \int_0^{\tau_1} \mathrm{d}\tau_2 e^{-i\tau_2 b_3B}[-ib_3B, -ic_3A]e^{i\tau_2 b_3B} + \int_0^{\tau_1} \mathrm{d}\tau_2 e^{-i\tau_1 b_3B}e^{-i\tau_2 a_3A}[-ia_3A, -id_2B]e^{i\tau_2 a_3A}e^{i\tau_1 b_3B} \end{aligned}$$

$$+\int_0^{\tau_1} \mathrm{d}\tau_2 e^{-i\tau_1 b_3 B} e^{-i\tau_1 a_3 A} e^{-i\tau_2 b_2 B}[-ib_2 B, -ic_2 A] e^{i\tau_2 b_2 B} e^{i\tau_1 a_3 A} e^{i\tau_1 b_3 B}$$
$$+\int_0^{\tau_1} \mathrm{d}\tau_2 e^{-i\tau_1 b_3 B} e^{-i\tau_1 a_3 A} e^{-i\tau_1 b_2 B} e^{-i\tau_2 a_2 A}[-ia_2 A, -id_1 B] e^{i\tau_2 a_2 A} e^{i\tau_1 b_2 B} e^{i\tau_1 a_3 A} e^{i\tau_1 b_3 B}$$
$$+\int_0^{\tau_1} \mathrm{d}\tau_2 e^{-i\tau_1 b_3 B} e^{-i\tau_1 a_3 A} e^{-i\tau_1 b_2 B} e^{-i\tau_1 a_2 A} e^{-i\tau_2 b_1 B}[-ib_1 B, -ic_1 A] e^{i\tau_2 b_1 B} e^{i\tau_1 a_2 A} e^{i\tau_1 b_2 B} e^{i\tau_1 a_3 A} e^{i\tau_1 b_3 B}. \quad \text{(M11)}$$

The operator-valued function $\mathcal{T}_4(\tau_1)$ has the order condition $\mathcal{T}_4(\tau_1) = O(\tau_1^4)$, which follows from Proposition D.3 and the fact that $\mathcal{S}_4(t) = e^{-itH} + O(t^5)$. For terms in $\mathcal{T}_4(\tau_1)$, we compute the Taylor expansion of each layer of unitary conjugation as in Sec. III D. In light of Lemma D.2, we expand the time variables τ_1 and τ_2 to third order, as the double integral $\int_0^t \mathrm{d}\tau \int_0^{\tau_1} \mathrm{d}\tau_2$ already exists. We then apply the triangle inequality to bound the spectral norm of a linear combination of nested commutators of A and B with four nesting layers. Since $[A, A] = [B, B] = 0$ and $[A, B] = [B, A]$, the bound only contains $2^5/4 = 8$ nonzero terms.

Altogether, we obtain

$$\begin{aligned}\|\mathcal{S}_4(t) - e^{-itH}\| \le t^5 (&0.0047\|[A,[A,[A,[B,A]]]]\| + 0.0057\|[A,[A,[B,[B,A]]]]\| \\ &+ 0.0046\|[A,[B,[A,[B,A]]]]\| + 0.0074\|[A,[B,[B,[B,A]]]]\| \\ &+ 0.0097\|[B,[A,[A,[B,A]]]]\| + 0.0097\|[B,[A,[B,[B,A]]]]\| \\ &+ 0.0173\|[B,[B,[A,[B,A]]]]\| + 0.0284\|[B,[B,[B,[B,A]]]]\|),\end{aligned} \quad \text{(M12)}$$

assuming $t \ge 0$.

TABLE II. Coefficients of the fourth-order Trotter error bound [Eq. (M14)] for Hamiltonians with three summands.

Commutator	Coefficient	Commutator	Coefficient	Commutator	Coefficient
$\|[H_1,[H_1,[H_1,[H_2,H_1]]]]\|$	0.0047	$\|[H_1,[H_1,[H_1,[H_3,H_1]]]]\|$	0.0047	$\|[H_1,[H_1,[H_1,[H_3,H_2]]]]\|$	0.0043
$\|[H_1,[H_1,[H_2,[H_2,H_1]]]]\|$	0.0057	$\|[H_1,[H_1,[H_2,[H_3,H_1]]]]\|$	0.0057	$\|[H_1,[H_1,[H_2,[H_3,H_2]]]]\|$	0.0057
$\|[H_1,[H_1,[H_3,[H_2,H_1]]]]\|$	0.0057	$\|[H_1,[H_1,[H_3,[H_3,H_1]]]]\|$	0.0057	$\|[H_1,[H_1,[H_3,[H_3,H_2]]]]\|$	0.0057
$\|[H_1,[H_2,[H_1,[H_2,H_1]]]]\|$	0.0046	$\|[H_1,[H_2,[H_1,[H_3,H_1]]]]\|$	0.0046	$\|[H_1,[H_2,[H_1,[H_3,H_2]]]]\|$	0.0035
$\|[H_1,[H_2,[H_2,[H_2,H_1]]]]\|$	0.0074	$\|[H_1,[H_2,[H_2,[H_3,H_1]]]]\|$	0.0070	$\|[H_1,[H_2,[H_2,[H_3,H_2]]]]\|$	0.0062
$\|[H_1,[H_2,[H_3,[H_2,H_1]]]]\|$	0.0082	$\|[H_1,[H_2,[H_3,[H_3,H_1]]]]\|$	0.0082	$\|[H_1,[H_2,[H_3,[H_3,H_2]]]]\|$	0.0082
$\|[H_1,[H_3,[H_1,[H_2,H_1]]]]\|$	0.0046	$\|[H_1,[H_3,[H_1,[H_3,H_1]]]]\|$	0.0046	$\|[H_1,[H_3,[H_1,[H_3,H_2]]]]\|$	0.0035
$\|[H_1,[H_3,[H_2,[H_2,H_1]]]]\|$	0.0070	$\|[H_1,[H_3,[H_2,[H_3,H_1]]]]\|$	0.0058	$\|[H_1,[H_3,[H_2,[H_3,H_2]]]]\|$	0.0046
$\|[H_1,[H_3,[H_3,[H_2,H_1]]]]\|$	0.0082	$\|[H_1,[H_3,[H_3,[H_3,H_1]]]]\|$	0.0074	$\|[H_1,[H_3,[H_3,[H_3,H_2]]]]\|$	0.0074
$\|[H_2,[H_1,[H_1,[H_2,H_1]]]]\|$	0.0150	$\|[H_2,[H_1,[H_1,[H_3,H_1]]]]\|$	0.0150	$\|[H_2,[H_1,[H_1,[H_3,H_2]]]]\|$	0.0141
$\|[H_2,[H_1,[H_2,[H_2,H_1]]]]\|$	0.0161	$\|[H_2,[H_1,[H_2,[H_3,H_1]]]]\|$	0.0161	$\|[H_2,[H_1,[H_2,[H_3,H_2]]]]\|$	0.0161
$\|[H_2,[H_1,[H_3,[H_2,H_1]]]]\|$	0.0161	$\|[H_2,[H_1,[H_3,[H_3,H_1]]]]\|$	0.0161	$\|[H_2,[H_1,[H_3,[H_3,H_2]]]]\|$	0.0161
$\|[H_2,[H_2,[H_1,[H_2,H_1]]]]\|$	0.0239	$\|[H_2,[H_2,[H_1,[H_3,H_1]]]]\|$	0.0239	$\|[H_2,[H_2,[H_1,[H_3,H_2]]]]\|$	0.0212
$\|[H_2,[H_2,[H_2,[H_2,H_1]]]]\|$	0.0315	$\|[H_2,[H_2,[H_2,[H_3,H_1]]]]\|$	0.0306	$\|[H_2,[H_2,[H_2,[H_3,H_2]]]]\|$	0.0290
$\|[H_2,[H_2,[H_3,[H_2,H_1]]]]\|$	0.0303	$\|[H_2,[H_2,[H_3,[H_3,H_1]]]]\|$	0.0303	$\|[H_2,[H_2,[H_3,[H_3,H_2]]]]\|$	0.0303
$\|[H_2,[H_3,[H_1,[H_2,H_1]]]]\|$	0.0179	$\|[H_2,[H_3,[H_1,[H_3,H_1]]]]\|$	0.0179	$\|[H_2,[H_3,[H_1,[H_3,H_2]]]]\|$	0.0153
$\|[H_2,[H_3,[H_2,[H_2,H_1]]]]\|$	0.0232	$\|[H_2,[H_3,[H_2,[H_3,H_1]]]]\|$	0.0206	$\|[H_2,[H_3,[H_2,[H_3,H_2]]]]\|$	0.0179
$\|[H_2,[H_3,[H_3,[H_2,H_1]]]]\|$	0.0259	$\|[H_2,[H_3,[H_3,[H_3,H_1]]]]\|$	0.0241	$\|[H_2,[H_3,[H_3,[H_3,H_2]]]]\|$	0.0241
$\|[H_3,[H_1,[H_1,[H_2,H_1]]]]\|$	0.0204	$\|[H_3,[H_1,[H_1,[H_3,H_1]]]]\|$	0.0204	$\|[H_3,[H_1,[H_1,[H_3,H_2]]]]\|$	0.0186
$\|[H_3,[H_1,[H_2,[H_2,H_1]]]]\|$	0.0225	$\|[H_3,[H_1,[H_2,[H_3,H_1]]]]\|$	0.0225	$\|[H_3,[H_1,[H_2,[H_3,H_2]]]]\|$	0.0217
$\|[H_3,[H_1,[H_3,[H_2,H_1]]]]\|$	0.0225	$\|[H_3,[H_1,[H_3,[H_3,H_1]]]]\|$	0.0225	$\|[H_3,[H_1,[H_3,[H_3,H_2]]]]\|$	0.0225
$\|[H_3,[H_2,[H_1,[H_2,H_1]]]]\|$	0.0423	$\|[H_3,[H_2,[H_1,[H_3,H_1]]]]\|$	0.0423	$\|[H_3,[H_2,[H_1,[H_3,H_2]]]]\|$	0.0377
$\|[H_3,[H_2,[H_2,[H_2,H_1]]]]\|$	0.0585	$\|[H_3,[H_2,[H_2,[H_3,H_1]]]]\|$	0.0571	$\|[H_3,[H_2,[H_2,[H_3,H_2]]]]\|$	0.0537
$\|[H_3,[H_2,[H_3,[H_2,H_1]]]]\|$	0.0502	$\|[H_3,[H_2,[H_3,[H_3,H_1]]]]\|$	0.0502	$\|[H_3,[H_2,[H_3,[H_3,H_2]]]]\|$	0.0502
$\|[H_3,[H_3,[H_1,[H_2,H_1]]]]\|$	0.0423	$\|[H_3,[H_3,[H_1,[H_3,H_1]]]]\|$	0.0423	$\|[H_3,[H_3,[H_1,[H_3,H_2]]]]\|$	0.0377
$\|[H_3,[H_3,[H_2,[H_2,H_1]]]]\|$	0.0681	$\|[H_3,[H_3,[H_2,[H_3,H_1]]]]\|$	0.0641	$\|[H_3,[H_3,[H_2,[H_3,H_2]]]]\|$	0.0601
$\|[H_3,[H_3,[H_3,[H_2,H_1]]]]\|$	0.0648	$\|[H_3,[H_3,[H_3,[H_3,H_1]]]]\|$	0.0621	$\|[H_3,[H_3,[H_3,[H_3,H_2]]]]\|$	0.0628

Proposition M.1: (Trotter error bound for the fourth-order Suzuki formula with two summands). Let $H = A + B$ be a Hamiltonian consisting of two summands and $t \geq 0$. Let $\mathcal{S}_4(t)$ be the fourth-order Suzuki formula (11). Then,

$$\begin{aligned}\|\mathcal{S}_4(t) - e^{-itH}\| \leq t^5 (&0.0047\|[A,[A,[A,[B,A]]]]\| + 0.0057\|[A,[A,[B,[B,A]]]]\| \\ &+ 0.0046\|[A,[B,[A,[B,A]]]]\| + 0.0074\|[A,[B,[B,[B,A]]]]\| \\ &+ 0.0097\|[B,[A,[A,[B,A]]]]\| + 0.0097\|[B,[A,[B,[B,A]]]]\| \\ &+ 0.0173\|[B,[B,[A,[B,A]]]]\| + 0.0284\|[B,[B,[B,[B,A]]]]\|). \end{aligned} \tag{M13}$$

A generalization of this approach analyzes Hamiltonians with three summands, which is relevant for certain nearest-neighbor and power-law systems where terms are ordered in an X-Y-Z pattern [Eq. (125)].

Proposition M.2: (Trotter error bound for the fourth-order Suzuki formula with three summands). Let $H = H_1 + H_2 + H_3$ be a Hamiltonian consisting of three summands and $t \geq 0$. Let $\mathcal{S}_4(t)$ be the fourth-order Suzuki formula (11). Then,

$$U\|\mathcal{S}_4(t) - e^{-itH}\| \leq t^5 \sum_{i,j,k,l,m=1}^{3} c_{i,j,k,l,m}\|[H_i,[H_j,[H_k,[H_l,H_m]]]]\|, \tag{M14}$$

where the coefficients $c_{i,j,k,l,m}$ are given by Table II.

Unlike the first- and second-order cases, we do not have a rigorous proof of the tightness of these bounds. However, our numerical result suggests that these bounds are close to tight for one-dimensional Heisenberg models with nearest-neighbor [Eq. (122)] and power-law [Eq. (126)] interactions. We hope future work will shed light on the tightness of our analysis through either theoretical justification or numerical calculation.

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

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

[3] D. Aharonov and A. Ta-Shma, *Adiabatic Quantum State Generation and Statistical Zero Knowledge*, in *Proceedings of the 35th ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2003), pp. 20–29, https://doi.org/10.1145/780542.780546.

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

[5] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, *Exponential Improvement in Precision for Simulating Sparse Hamiltonians*, in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2014), pp. 283–292, https://doi.org/10.1145/2591796.2591854.

[6] D. W. Berry, A. M. Childs, and R. Kothari, *Hamiltonian Simulation with Nearly Optimal Dependence on All Parameters*, in *Proceedings of the 56th IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, 2015), pp. 792–809, https://doi.org/10.1109/FOCS.2015.54.

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

[8] G. H. Low, *Hamiltonian Simulation with Nearly Optimal Dependence on Spectral Norm*, in *Proceedings of the 51th ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2019), pp. 491–502, https://doi.org/10.1145/3313276.3316386.

[9] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, *Gate Count Estimates for Performing Quantum Chemistry on Small Quantum Computers*, Phys. Rev. A **90**, 022305 (2014).

[10] 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).

[11] 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).

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

[13] S. P. Jordan, K. S. M. Lee, and J. Preskill, *Quantum Algorithms for Quantum Field Theories*, Science **336**, 1130 (2012).

[14] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri *et al.*, *Towards Quantum Chemistry on a Quantum Computer*, Nat. Chem. **2**, 106 (2010).

[15] 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).

[16] A. W. Harrow, A. Hassidim, and S. Lloyd, *Quantum Algorithm for Linear Systems of Equations*, Phys. Rev. Lett. **103**, 150502 (2009).
[17] F. G. S. L. Brandao and K. M. Svore, *Quantum Speed-ups for Solving Semidefinite Programs*, in *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, 2017), pp. 415–426, https://doi.org/10.1109/FOCS.2017.45.
[18] E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Algorithm for the Hamiltonian NAND Tree*, Theory Comput. **4**, 169 (2008).
[19] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, *Exponential Algorithmic Speedup by a Quantum Walk*, in *Proceedings of the 35th ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2003), pp. 59–68, https://doi.org/10.1145/780542.780552.
[20] D. W. Berry, *High-Order Quantum Algorithm for Solving Linear Differential Equations*, J. Phys. A **47**, 105301 (2014).
[21] M. Suzuki, *General Theory of Fractal Path Integrals with Applications to Many-Body Theories and Statistical Physics*, J. Math. Phys. (N.Y.) **32**, 400 (1991).
[22] S. Blanes and F. Casas, *A Concise Introduction to Geometric Numerical Integration* (Chapman and Hall/CRC, London, 2016)https://doi.org/10.1201/b21563.
[23] 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).
[24] G. H. Low and I. L. Chuang, *Hamiltonian Simulation by Qubitization*, Quantum **3**, 163 (2019).
[25] G. H. Low and N. Wiebe, *Hamiltonian Simulation in the Interaction Picture*, arXiv:1805.00675.
[26] S. Hadfield and A. Papageorgiou, *Divide and Conquer Approach to Quantum Hamiltonian Simulation*, New J. Phys. **20**, 043003 (2018).
[27] A. M. Childs, A. Ostrander, and Y. Su, *Faster Quantum Simulation by Randomization*, Quantum **3**, 182 (2019).
[28] G. H. Low, V. Kliuchnikov, and N. Wiebe, *Well-Conditioned Multiproduct Hamiltonian Simulation*, arXiv:1907.11679.
[29] Y. Ouyang, D. R. White, and E. Campbell, *Compilation by Stochastic Hamiltonian Sparsification*, Quantum **4**, 235 (2020).
[30] E. Campbell, *Random Compiler for Fast Hamiltonian Simulation*, Phys. Rev. Lett. **123**, 070503 (2019).
[31] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, *Toward the First Quantum Simulation with Quantum Speedup*," Proc. Natl. Acad. Sci. U.S.A. **115**, 9456 (2018).
[32] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, *Elucidating Reaction Mechanisms on Quantum Computers*, Proc. Natl. Acad. Sci. U.S.A. **114**, 7555 (2017).
[33] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe, *Chemical Basis of Trotter-Suzuki Errors in Quantum Chemistry Simulation*, Phys. Rev. A **91**, 022311 (2015).
[34] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, *Solving Strongly Correlated Electron Models on a Quantum Computer*, Phys. Rev. A **92**, 062318 (2015).
[35] A. M. Childs and Y. Su, *Nearly Optimal Lattice Simulation by Product Formulas*, Phys. Rev. Lett. **123**, 050503 (2019).
[36] M. Troyer, N. Pearson, and D. Poulin, *Trotter Error Scaling with System Size in Quantum Simulations*, in *APS Meeting Abstracts* (2019).
[37] M. Kliesch, C. Gogolin, and J. Eisert, *Lieb-Robinson Bounds and the Simulation of Time-Evolution of Local Observables in Lattice Systems*, in *Many-Electron Approaches in Physics, Chemistry and Mathematics* (Springer, New York, 2014), pp. 301–318, https://doi.org/10.1007/978-3-319-06379-9_17.
[38] S. Bravyi, *Monte Carlo Simulation of Stoquastic Hamiltonians*, Quantum Inf. Comput. **15**, 1122 (2015).
[39] S. Bravyi and D. Gosset, *Polynomial-Time Classical Simulation of Quantum Ferromagnets*, Phys. Rev. Lett. **119**, 100503 (2017).
[40] M. Suzuki, *Decomposition Formulas of Exponential Operators and Lie Exponentials with Some Applications to Quantum Mechanics and Statistical Physics*, J. Math. Phys. (N.Y.) **26**, 601 (1985).
[41] J. Huyghebaert and H. De Raedt, *Product Formula Methods for Time-Dependent Schrödinger Problems*, J. Phys. A **23**, 5777 (1990).
[42] S. Descombes and M. Thalhammer, *An Exact Local Error Representation of Exponential Operator Splitting Methods for Evolutionary Problems and Applications to Linear Schrödinger Equations in the Semi-classical Regime*, BIT **50**, 729 (2010).
[43] 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).
[44] R. D. Somma, *A Trotter-Suzuki Approximation for Lie Groups with Applications to Hamiltonian Simulation*, J. Math. Phys. (N.Y.) **57**, 062202 (2016).
[45] M. Thalhammer, *High-Order Exponential Operator Splitting Methods for Time-Dependent Schrödinger Equations*, SIAM J. Numer. Anal. **46**, 2022 (2008).
[46] T. Jahnke and C. Lubich, *Error Bounds for Exponential Operator Splittings*, BIT **40**, 735 (2000).
[47] M. Thalhammer, *Convergence Analysis of High-Order Time-Splitting Pseudospectral Methods for Nonlinear Schrödinger Equations*, SIAM J. Numer. Anal. **50**, 3231 (2012).
[48] R. I. McLachlan and G. R. W. Quispel, *Splitting Methods*, Acta Numer. **11**, 341 (2002).
[49] R. I. McLachlan, *On the Numerical Integration of Ordinary Differential Equations by Symmetric Composition Methods*, SIAM J. Sci. Comput. **16**, 151 (1995).
[50] E. Hairer, G. Wanner, and C. Lubich, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations* (Springer Science & Business Media, New York, 2006), Vol. 31, https://doi.org/10.1007/978-3-662-05018-7.
[51] The 1-norm $\|H\|_1$ and the induced 1-norm $|||H|||_1$ are formally defined in Sec. II A. For now, it suffices to know

that $|||H|||_1 \leq \|H\|_1$ and that the gap can be significant for many k-local Hamiltonians.

[52] 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).

[53] T. Peng, A. Harrow, M. Ozols, and X. Wu, *Simulating Large Quantum Circuits on a Small Quantum Computer*, Phys. Rev. Lett. **125**, 150504 (2020).

[54] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, England, 2012), https://doi.org/10.1017/CBO9781139020411.

[55] J. D. Dollard and C. N. Friedman, *Product Integration with Application to Differential Equations* (Cambridge University Press, Cambridge, England, 1984), https://doi.org/http://dx.doi.org/10.1017/CBO9781107340701.

[56] Alternatively, we may define a time-ordered exponential by its Dyson series or by a convergent sequence of products of ordinary matrix exponentials, and verify that this alternative definition satisfies a certain differential equation. We prefer the differential-equation definition since it is more versatile for the analysis in this paper.

[57] N. Wiebe, D. Berry, P. Høyer, and B. C. Sanders, *Higher Order Decompositions of Ordered Operator Exponentials*, J. Phys. A **43**, 065203 (2010).

[58] W. Auzinger and W. Herfort, *Local Error Structures and Order Conditions in Terms of Lie Elements for Exponential Splitting Schemes*, Opusc. Math. **34**, 243 (2014).

[59] W. Auzinger, O. Koch, and M. Thalhammer, *Defect-Based Local Error Estimators for Splitting Methods, with Application to Schrödinger Equations, Part II. Higher-Order Methods for Linear Problems*, J. Comput. Appl. Math. **255**, 384 (2014).

[60] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, *Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity*, Phys. Rev. X **8**, 041015 (2018).

[61] A. J. Ferris, *Fourier Transform for Fermionic Systems and the Spectral Tensor Network*, Phys. Rev. Lett. **113**, 010401 (2014).

[62] For Hamiltonians that are supported on finite lattices, we simply add trivial terms supported outside the lattices.

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

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

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

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

[67] Z.-X. Gong, M. Foss-Feig, S. Michalakis, and A. V. Gorshkov, *Persistence of Locality in Systems with Power-Law Interactions*, Phys. Rev. Lett. **113**, 030602 (2014).

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

[69] D.-M. Storch, M. Van Den Worm, and M. Kastner, *Interplay of Soundcone and Supersonic Propagation in Lattice Models with Power Law Interactions*, New J. Phys. **17**, 063021 (2015).

[70] C.-F. Chen and A. Lucas, *Finite Speed of quantum scrambling with long range interactions*, Phys. Rev. Lett. **123**, 250605 (2019).

[71] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low, *Quantum Algorithm for Simulating Real Time Evolution of Lattice Hamiltonians*, in *Proceedings of the 59th IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, 2018), pp. 350–360, https://doi.org/10.1109/FOCS.2018.00041.

[72] More recent bounds [70,73] provide tighter light cones than in Tran *et al.* [52] for $\alpha > 2d + 1$.

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

[74] D. J. Luitz, N. Laflorencie, and F. Alet, *Many-Body Localization Edge in the Random-Field Heisenberg Chain*, Phys. Rev. B **91**, 081103(R) (2015).

[75] Y. Atia and D. Aharonov, *Fast-Forwarding of Hamiltonians and Exponentially Precise Measurements*, Nat. Commun. **8**, 1572 (2017).

[76] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. Kin-Lic Chan, and R. Babbush, *Quantum Simulation of Electronic Structure with Linear Depth and Connectivity*, Phys. Rev. Lett. **120**, 110501 (2018).

[77] D. Poulin, A. Qarry, R. D. Somma, and F. Verstraete, *Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space*, Phys. Rev. Lett. **106**, 170501 (2011).

[78] M. Kieferová, A. Scherer, and D. W. Berry, *Simulating the Dynamics of Time-Dependent Hamiltonians with a Truncated Dyson Series*, Phys. Rev. A **99**, 042314 (2019).

[79] D. W. Berry, A. M. Childs, Y. Su, X. Wang, and N. Wiebe, *Time-Dependent Hamiltonian Simulation with L^1-Norm Scaling*, Quantum **4**, 254 (2020).

[80] M. Heyl, Philipp. Hauke, and P. Zoller, *Quantum Localization Bounds Trotter Errors in Digital Quantum Simulation*, Sci. Adv. **5**, eaau8342 (2019).

[81] L. M. Sieberer, T. Olsacher, A. Elben, M. Heyl, P. Hauke, F. Haake, and P. Zoller, *Digital Quantum Simulation, Trotter Errors, and Quantum Chaos of the Kicked Top*, npj Quantum Inf. **5**, 78 (2019).

[82] S. McArdle, X. Yuan, and S. Benjamin, *Error-Mitigated Digital Quantum Simulation*, Phys. Rev. Lett. **122**, 180501 (2019).

[83] A. W. Knapp, *Basic Real Analysis*, Digital 2nd ed. (Birkhëuser, Boston, 2005), https://doi.org/http://dx.doi.org/10.3792/euclid/9781429799997.

[84] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, New York, 2014), https://doi.org/10.1002/9781119019572.

[85] S. C. Eisenstat and I. C. F. Ipsen, *Relative Perturbation Techniques for Singular Value Problems*, SIAM J. Numer. Anal. **32**, 1972 (1995).

[86] I. C. F. Ipsen, *Relative Perturbation Results for Matrix Eigenvalues and Singular Values*, Acta Numer. **7**, 151 (1998).