Destructive Error Interference in Product-Formula Lattice Simulation

Minh C. Tran,[1,2,3] Su-Kuan Chu,[1,2] Yuan Su,[1,4,5] Andrew M. Childs,[1,4,5] and Alexey V. Gorshkov[1,2,3]
[1]Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, Maryland 20742, USA
[2]Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA
[3]Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA
[4]Department of Computer Science, University of Maryland, College Park, Maryland 20742, USA
[5]Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742, USA

(Received 22 January 2020; accepted 6 May 2020; published 4 June 2020)

Quantum computers can efficiently simulate the dynamics of quantum systems. In this Letter, we study the cost of digitally simulating the dynamics of several physically relevant systems using the first-order product-formula algorithm. We show that the errors from different Trotterization steps in the algorithm can interfere destructively, yielding a much smaller error than previously estimated. In particular, we prove that the total error in simulating a nearest-neighbor interacting system of n sites for time t using the first-order product formula with r time slices is $O(nt/r + nt^3/r^2)$ when nt^2/r is less than a small constant. Given an error tolerance ε, the error bound yields an estimate of $\max\{O(n^2t/\varepsilon), O(n^2t^{3/2}/\varepsilon^{1/2})\}$ for the total gate count of the simulation. The estimate is tighter than previous bounds and matches the empirical performance observed in Childs *et al.* [Proc. Natl. Acad. Sci. U.S.A. **115**, 9456 (2018)]. We also provide numerical evidence for potential improvements and conjecture an even tighter estimate for the gate count.

DOI: 10.1103/PhysRevLett.124.220502

Simulating the dynamics of quantum systems is one of the primary applications of quantum computers. While analog quantum simulations rely on engineering physical systems to mimic other systems, digital quantum simulations use algorithms to decompose the evolution unitary into a sequence of elementary quantum gates. The first quantum simulation algorithm proposed by Lloyd [1] uses the Lie-Trotter product formula, also known as the first-order product formula (PF1) [2,3]. Since then, more advanced quantum simulation algorithms have been developed, including algorithms based on the higher-order product formulas [4–7], linear combinations of unitaries [8,9], quantum signal processing [10], and Lieb-Robinson bounds [11,12], which all asymptotically reduce the cost of digital quantum simulation in terms of the number of gates used in the limit of large time or large system size.

Despite these developments, PF1 remains one of the most popular algorithms for near-term implementations of digital quantum simulation due to its simplicity. In practice, the small prefactor in the scaling of the gate count of PF1 compared to more advanced quantum simulation algorithms makes it attractive for simulations where the evolution time and the system size are not too large [7].

Despite its simplicity and wide applicability, a tight error bound for PF1 in simulating many physically relevant systems remains elusive. Recent works [2–4] estimated that $O(n^2t^2)$ elementary gates suffice to simulate the dynamics of a nearest-neighbor interacting system consisting of n sites for time t using PF1 [13]. However, the numerical evidence in Ref. [7] suggests that PF1 performs much better than this in practice. In particular, the gate count for simulating the dynamics of a nearest-neighbor Heisenberg spin chain of length n for time $t = n$ scales only as $O(n^{2.964})$. In addition, Heyl *et al.* [14] also found that the error of simulating the time evolution of a local observable using PF1 can be much smaller than theoretically estimated.

In this Letter, we provide an approach to tighten the error bound of PF1 for simulating several physically relevant systems, including those with nearest-neighbor interactions. The key finding of the Letter is that the errors from different steps of the algorithm can combine destructively, resulting in a smaller total error than previous analysis estimates. In particular, the tighter error bound suggests that simulating the dynamics of a nearest-neighbor interacting system of n sites for time t up to an error tolerance ε requires only $\max\{O(n^2t/\varepsilon), O(n^2t^{3/2}/\varepsilon^{1/2})\}$ quantum gates, which is asymptotically smaller than the state-of-the-art bound $O(n^2t^2/\varepsilon)$ in Refs. [2–4]. At $t = n$ and at a fixed ε, our estimate $O(n^3)$ also closely matches the empirical gate count $O(n^{2.964})$ computed in Ref. [7].

Setup.—We assume that the system evolves under a Hamiltonian $H = \sum_X h_X$, which is a sum of time-independent terms h_X, each acting nontrivially on a subset X of constant size. Our approach applies if there exists a partition $H = H_1 + H_2$ such that the terms h_X in H_1 mutually commute and the terms h_X in H_2 also mutually commute. Examples of Hamiltonians that satisfy this assumption include all one-dimensional, finite-range [15] interacting systems, such as the Heisenberg model and the transverse field Ising model in one dimension with either

open or periodic boundary conditions and with or without disorder. Additionally, this assumption also covers some physically relevant systems in higher dimensions, such as the transverse field Ising model with either finite- or long-range interactions.

To simulate the time evolution of the system for time t using elementary quantum gates, we use the first-order product formula [1]

$$U_t \approx [U_{t/r}^{(1)} U_{t/r}^{(2)}]^r, \tag{1}$$

where $U_t := \exp(-iHt)$, $U_{t/r}^{(p)} := \exp(-iH_p t/r)$ for $p = 1$, 2, and r is the number of time segments to be chosen later so that the norm of the total error $\Delta := U_t - [U_{t/r}^{(1)} U_{t/r}^{(2)}]^r$ is, at most, a constant ε. By our assumption that the terms within H_p ($p = 1$, 2) mutually commute, we can further decompose the evolution $U_{t/r}^{(p)}$ into a product of elementary quantum gates with no additional error.

For simplicity, we demonstrate our approach to estimating the gate count of PF1 on a one-dimensional lattice of n sites, evolving under a time-independent, nearest-neighbor Hamiltonian $H = \sum_{i=1}^{n-1} h_i$, where h_i is supported only on sites i, $i+1$, $\|h_i\| \le J$ for all i, J is a constant, and $\|\cdot\|$ denotes the operator norm. Without loss of generality, we also assume $J = 1$, which sets the timescale for the dynamics of the system. We then apply PF1 to the partition $H = H_1 + H_2$, where $H_1 = \sum_{\text{odd}\, j} h_j$ and $H_2 = \sum_{\text{even}\, j} h_j$. Note that the terms within H_1 (H_2) mutually commute and therefore satisfy the aforementioned assumption.

Leading contributions.—To estimate the gate count, we first need a bound on the total error Δ. The previous best bound from Ref. [4] gives $\|\Delta\| \le O(nt^2/r)$, so that $r = \Theta(nt^2/\varepsilon)$ suffices to ensure the total error at most ε, giving gate count $nr = O(t^2 n^2/\varepsilon)$. Before we prove our tighter bound, we will first argue simply based on the lowest-order error that $\|\Delta\| \approx O(nt/r)$, which would result in a gate count $O(tn^2/\varepsilon)$, matching the empirical estimate of about $O(n^3)$ for $t = n$ in Ref. [7].

Let $\delta = U_{t/r} - U_{t/r}^{(1)} U_{t/r}^{(2)}$ be the error of the approximation in each time segment. In the limit $r \gg t$, the leading contribution to δ is given by the commutator between H_1 and H_2 [1],

$$\|\delta\| \approx \frac{1}{2}\frac{t^2}{r^2}\|[H_1, H_2]\| = O\left(\frac{nt^2}{r^2}\right). \tag{2}$$

Replacing $U_{t/r}^{(1)} U_{t/r}^{(2)}$ by $U_{t/r} + \delta$ on the right-hand side of Eq. (1) and expanding to first order in δ, we have an approximation for the total error

$$\Delta \approx \sum_{j=0}^{r-1} U_{t/r}^{j} \delta U_{t/r}^{r-1-j} = \left(\sum_{j=0}^{r-1} U_{t/r}^{j} \delta U_{t/r}^{-j}\right) U_{t/r}^{r-1}, \tag{3}$$

where $U_{t/r}^{j} := (U_{t/r})^j$. If we bound $\|\Delta\|$ using the triangle inequality, i.e.,

$$\|\Delta\| \approx \left\| \sum_{j=0}^{r-1} U_{t/r}^{j} \delta U_{t/r}^{r-1-j} \right\| \le r\|\delta\| \approx O\left(\frac{nt^2}{r}\right), \tag{4}$$

we get the same error bound (and hence the same gate count) as Ref. [4].

To understand the key idea for improving the bound, imagine δ as a vector in the space of operators and the unitary evolution $U_{t/r}^{j} \delta U_{t/r}^{-j}$ as a rotation of the vector by a small angle proportional to jt/r. Since the terms in Eq. (3) correspond to rotations of δ by evenly spaced angles [16], the sum involves significant cancellation, making it much smaller than the upper bound derived using the triangle inequality [Eq. (4)].

To realize this intuition, we make a change of variables to $x = tj/r$ and approximate the sum in Δ by an integral,

$$\|\Delta\| \approx \left\| \sum_{j=0}^{r-1} U_{t/r}^{j} \delta U_{t/r}^{-j} \right\| \approx \frac{r}{t}\left\| \int_0^t dx U_x \delta U_{-x} \right\|. \tag{5}$$

With the assumption that $H = H_1 + H_2$ is a sum of two terms, we rewrite δ (to leading order in t/r) as

$$\delta \approx \frac{1}{2}[H_1, H_2]\frac{t^2}{r^2} = \frac{1}{2}[H, H_2]\frac{t^2}{r^2} \tag{6}$$

and use the identity

$$U_t A U_{-t} - A = -i \int_0^t dx U_x [H, A] U_{-x}, \tag{7}$$

with $A = (t^2/2r^2) H_2$, to evaluate the integral in Eq. (5) and arrive at an estimate for the norm of Δ,

$$\begin{aligned} \|\Delta\| &\approx \frac{r}{t}\left\| \int_0^t dx U_x \left[H, \frac{t^2}{2r^2} H_2\right] U_{-x} \right\| \\ &\le \frac{t}{2r} 2\|H_2\| = O\left(\frac{nt}{r}\right), \end{aligned} \tag{8}$$

which is a factor of t tighter than Eq. (4). We attribute this improvement to the destructive interference between the error terms in Eq. (3). To ensure that the total error $\|\Delta\|$ is at most ε, we choose $r = \Theta(nt/\varepsilon)$, leading to the total gate count $O(nr) = O(n^2 t/\varepsilon)$, which has optimal scaling in t [17]. At $t = n$ and fixed ε, the gate count becomes $O(n^3)$, which closely matches the empirical performance $O(n^{2.964})$ observed in Ref. [7].

Additionally, if the time step $t/r = \tau$ is a constant, the total error of the simulation $\|\Delta\| = O(n\tau)$ appears to be independent of the total number of time segments. This feature agrees well with Ref. [14], where the authors argue

that, for a fixed, small value of τ, the error in simulating the evolution of a local observable using PF1 would not increase with the total simulation time t. However, our bound is more general; it applies to the error in simulating the evolution unitary of the system and, hence, any observable.

Higher-order contributions.—We made three approximations in deriving Eq. (8). First, in Eq. (6), we considered δ to only the leading order in t/r and discarded terms of higher order in t/r. We then expanded Δ in Eq. (3) to only the first order in δ while ignoring the higher-order terms in δ^k. Additionally, we evaluated the sum in Eq. (5) by approximating it with an integral. We now make the estimation rigorous by considering the errors incurred upon making the three approximations.

First, we show that higher-order terms in t/r in the expansion of δ are indeed dominated by the second order. For that, we write δ as a series in t/r,

$$\delta \coloneqq U_{t/r} - U_{t/r}^{(1)} U_{t/r}^{(2)} = \sum_{k=2}^{\infty} \frac{(-it)^k}{k! r^k} \delta_k, \tag{9}$$

where δ_k are operators independent of t, r. If we only need a bound on the norm of δ, it is sufficient to bound the norms of δ_k. However, in addition to the norm, we are also interested in the structure of δ_k, described in Lemma 1, which is crucial for evaluating the total error [see Eq. (6)].

Lemma 1.—For all $k \geq 2$, there exist S_k, V_k such that $\delta_k = [H, S_k] + V_k$ and

$$\|V_k\| = O(e^{k-2} n^{k-2}), \tag{10}$$

$$\|S_k\| = O(k^2 n^{k-1}), \tag{11}$$

$$\|[H, S_k]\| = O(k^3 n^{k-1}), \tag{12}$$

where the big-O constants do not depend on k.

Lemma 1 holds for $k = 2$, with $S_2 = H_2$ and $V_2 = 0$ [see Eq. (6)]. For $k > 2$, we construct S_k, V_k inductively using the definition of δ_k in Eq. (9). The factor n^{k-2} in the norm of V_k comes from the $(k-2)$th nested commutators in the expansion of δ_k. We provide a detailed proof of the lemma in the Supplemental Material [18].

A corollary of Lemma 1 is $\|\delta_k\| = O(e^k n^{k-1})$, and therefore, we can immediately bound the norm of δ,

$$\begin{aligned} \|\delta\| &\leq \sum_{k=2}^{\infty} \frac{t^k}{k! r^k} \|\delta_k\| = O\left(\frac{nt^2}{r^2} \sum_{k=0}^{\infty} \frac{(ent)^k}{k! r^k}\right) \\ &= O\left(\frac{nt^2}{r^2} \exp \frac{ent}{r}\right) = O\left(\frac{nt^2}{r^2}\right), \end{aligned} \tag{13}$$

where we assume $r > ent$. We later fulfill this condition by choosing an appropriate value for r.

Another corollary of Lemma 1 is that $\delta = [H, S] + V$, where $S = \sum_{k=2}^{\infty} [(-it)^k / k! r^k] S_k$ and $V = \sum_{k=3}^{\infty} [(-it)^k / k! r^k] V_k$. It is straightforward to verify the bounds on the norms of S and V,

$$\|S\| = O\left(\frac{nt^2}{r^2}\right), \qquad \|V\| = O\left(\frac{nt^3}{r^3}\right), \tag{14}$$

where we again assume $r > ent$.

Next, we rectify the approximation in Eq. (5) by rigorously bounding the norm of the sum.

Lemma 2.—For any positive integer $a \geq 1$,

$$\left\| \sum_{j=0}^{a-1} U_{t/r}^j \delta U_{t/r}^{-j} \right\| = O\left(\frac{nt}{r}\right) + O\left(a \frac{nt^3}{r^3}\right). \tag{15}$$

When $a = r$, the left-hand side of Eq. (15) is exactly the sum in Eq. (5). We bound the sum by approximating it with an integral, which yields $O(nt/r)$ after evaluation. Carefully bounding the error of the approximation results in the second term $O(ant^3/r^3)$. We present the detailed proof of the lemma in the Supplemental Material [18].

Given Lemma 1 and Lemma 2, we now bound the total error $\|\Delta\|$. We expand Δ as a series in δ and write $\Delta = \sum_{k=1}^{r} \Delta_k$, where Δ_k involves only the kth order in δ. For example, $\Delta_1 = \sum_{j=0}^{r-1} U_{t/r}^j \delta U_{t/r}^{-j}$, the norm of which we can already bound using Lemma 2. We can use the same technique to estimate $\|\Delta_k\|$ for all $k \geq 1$ [18],

$$\|\Delta_k\| \leq r^{k-1} \|\delta\|^{k-1} O\left(\frac{nt}{r} + \frac{nt^3}{r^2}\right). \tag{16}$$

Finally, we bound $\|\Delta\|$ using the triangle inequality

$$\|\Delta\| \leq \sum_{k=1}^{r} \|\Delta_k\| = O\left(\frac{nt}{r} + \frac{nt^3}{r^2}\right), \tag{17}$$

where we assume $r\|\delta\| < 1/2$ so that $\sum_{k=1}^{r} (r\|\delta\|)^{k-1} = O(1)$. With our choice of r, this assumption later reduces to $\varepsilon t \leq 1$, where ε is the error tolerance of the simulation.

Empirical error scaling.—We now benchmark the bound in Eq. (17) against the empirical error in simulating the dynamics of a nearest-neighbor Heisenberg chain

$$H = \sum_{i=1}^{n-1} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1}, \tag{18}$$

where $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ denotes the Pauli matrices on qubit i. Using fixed values for n and r, we compute the total error of PF1 at different times t and plot the result in Fig. 1. We also plot in Fig. 2 the empirical errors of simulating the same system using the second-order (PF2) and the fourth-order (PF4) product formulas [6].

From Fig. 1, the total error of PF1 appears to agree well with our bound in Eq. (17). The change in the error scaling

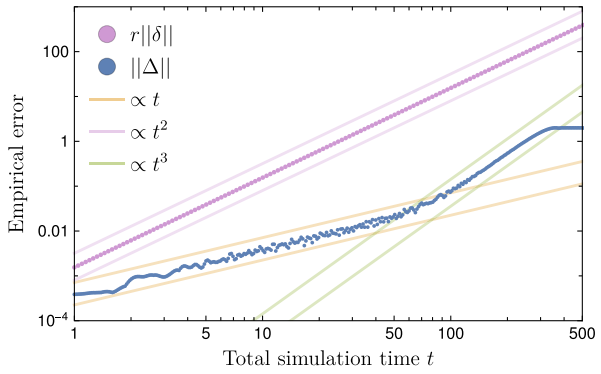

FIG. 1. The total error $\|\Delta\|$ (blue dots) of PF1 in simulating the Heisenberg chain in Eq. (18) is numerically evaluated at $n = 8$, $r = 10\,000$, and variable time t between 0 and 1000. The purple dots represent the error estimate $r\|\delta\|$ one would get using the triangle inequalities [Eq. (4)]. We also plot functions proportional to t (orange lines), t^2 (purple lines), and t^3 (green lines) for reference.

from $O(t)$ at small time to $O(t^3)$ at large time can be explained by the destructive error interference between the time slices as follows. While the leading error terms in each time slice scale as $O(t^2)$, they interfere destructively between time slices, resulting in a total contribution that increases with time at a slower rate $O(t)$ [recall Eq. (8)]. Meanwhile, some higher-order error terms do not interfere destructively. They scale as $O(t^3)$ and eventually take over as the primary contribution to the total error. This intuition also explains the similarity between the error scalings of PF1 (at late time) and PF2 (Fig. 2). On the other hand, if there were no destructive error interference between the time slices, the contribution from the leading error terms to the total error of PF1 would have scaled as $O(t^2)$ (Fig. 1, purple dots) and saturated at 2 before the higher-order terms could take over.

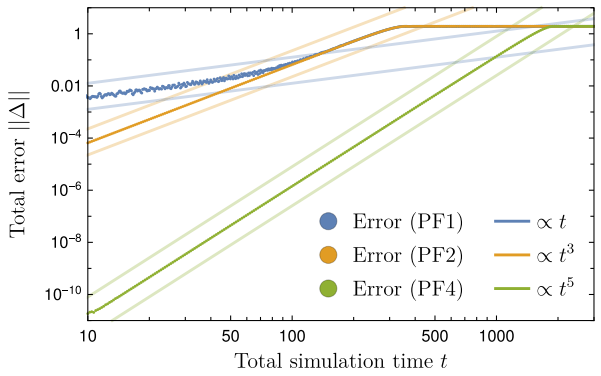

FIG. 2. The total error of simulating the Heisenberg chain with $n = 8$ spins in Eq. (18) using PF1 (blue dots), PF2 (orange dots), and PF4 (green dots) is numerically computed at $r = 10000$, and variable time t between 10 and 3000. We also plot functions proportional to t (blue lines), t^3 (orange lines), and t^5 (green lines) for reference.

We also note that the error of PF2 (PF4) scales as t^3 (t^5) initially before saturating at a later time, in agreement with the existing bounds using the triangle inequality for the higher-order product formulas [4,7]. This suggests the absence of significant destructive error interference for PF2 and PF4 in our numerical simulation.

Gate count.—Given the error bound in Eq. (17), we now count the number of gates for PF1. Equation (17) suggests we should choose

$$r \propto \max\left\{\frac{nt}{\varepsilon}, \sqrt{\frac{nt^3}{\varepsilon}}, 1\right\}, \tag{19}$$

so that the total error $\|\Delta\|$ is at most ε. First, we assume $nt \geq \varepsilon$ and consider two cases, corresponding to $\varepsilon t \leq 1$ (small time) and $\varepsilon t > 1$ (large time). The former condition implies that the first term in Eq. (19) dominates and therefore we should choose $r = \Theta(nt/\varepsilon)$. This choice of r together with $\varepsilon t \leq 1$ also fulfills the condition $r\|\delta\| < 1/2$ required earlier, as long as we choose a large enough prefactor in $\Theta(nt/\varepsilon)$. Thus, when $\varepsilon t \leq 1$, the gate count of PF1 is

$$O(rn) = O\left(\frac{n^2 t}{\varepsilon}\right). \tag{20}$$

On the other hand, when $\varepsilon t > 1$, we divide the simulation into m stages. In each stage, we simulate the evolution for time t/m with an error at most ε/m by further dividing the stage into r time segments. In order to apply the above analysis in each stage, we require m to be large enough so that $\varepsilon t/m^2 \leq 1$. Since the resulting gate count $O(mn^2 t/\varepsilon)$ increases with m, it is optimal to choose m as small as possible, i.e., $m = \lceil\sqrt{\varepsilon t}\rceil$. Therefore, the total gate count in this case is

$$O\left(\sqrt{\varepsilon t}\frac{n^2 t}{\varepsilon}\right) = O\left(\frac{n^2 t^{3/2}}{\varepsilon^{1/2}}\right). \tag{21}$$

Finally, when $nt < \varepsilon$, we simply choose $r = \Theta(1)$, giving gate count $O(1)$. Combining the above arguments, we have an upper bound on the total gate count of

$$\max\left\{O\left(\frac{n^2 t}{\varepsilon}\right), O\left(\frac{n^2 t^{3/2}}{\varepsilon^{1/2}}\right), O(1)\right\}, \tag{22}$$

which is valid for all times t and is tighter than the previous best estimate in Ref. [4].

Discussion and outlook.—As mentioned earlier, we assume that the terms of the Hamiltonian can be separated into two parts such that the terms within each part mutually commute. Therefore, our results apply to translationally invariant spin chains in one dimension with finite-range interactions and with either open or periodic boundary conditions, as well as disordered spin chains, such as those

featuring many-body localization [19]. Additionally, our analysis also holds for some systems in higher dimensions, such as the transverse field Ising model with either finite- or long-range interactions, where the two mutually commuting parts of the Hamiltonian are the spin-spin interactions and the field terms. However, for long-range interactions, the number of interaction terms can scale as $O(n^2)$ [instead of $O(n)$ for the finite-range interactions], so the scalings of the error bound and of the gate count as functions of n must be adjusted accordingly. Furthermore, our technique can also be used to bound the error in simulating materials where the electronic structure Hamiltonian in the plane wave dual basis [20] is a sum of mutually commuting kinetic energy terms and Coulomb interactions.

However, it is unclear whether our approach generalizes to Hamiltonians that can only be separated into three or more mutually commuting parts, such as those that typically occur in higher dimensions and systems with general long-range interactions, where the simple relation between δ and H in Eq. (6) no longer holds in general. Despite our numerics for the Heisenberg interactions, it remains open whether significant destructive error interference can be achieved for higher-order product formulas in simulating other classes of Hamiltonians. In addition, although our main focus in this Letter is on real-time simulation, it would be interesting to consider the implications of our bound for the error of the product formula in simulating imaginary time evolution, which is relevant for path integral quantum Monte Carlo algorithms [21].

We also note that, while our analysis requires $r\|\delta\| < 1/2$, our numerical calculation (see Fig. 1) shows that our error bound agrees well with the empirical scaling even at large values of t, where $r\|\delta\| \gg 1/2$. Therefore, we conjecture that the error bound in Eq. (17) is valid regardless of whether εt is less than 1. If the conjecture holds, Eq. (19) implies that we should choose $r \propto nt/\varepsilon$ and $r \propto \sqrt{nt^3/\varepsilon}$ for $\varepsilon t \leq n$ and $\varepsilon t > n$, respectively (in the limit of large n and t). The former choice yields the same gate count $O(n^2 t/\varepsilon)$ as in Eq. (20), but the latter choice leads to a gate count of $O(nr) = O(\sqrt{n^3 t^3/\varepsilon})$, which is tighter than the estimate in Eq. (21). Thus, the conjecture would imply that PF1 performs as well as PF2—whose gate count is also $O(\sqrt{n^3 t^3/\varepsilon})$ [4]—in the large-time limit. We consider proving the conjecture a very interesting future direction.

M. C. T., S. K. C., and A. V. G. acknowledge funding from U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Accelerated Research in Quantum Computing program (Award No. DE-SC0020312), U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials and Chemical Sciences Research for Quantum Information Science program (Award No. DE-SC0019449), U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Quantum Testbed Pathfinder program (Award No. DE-SC0019040), NSF PFCQC program, ARO MURI, AFOSR, ARL CDQI, and NSF PFC at JQI. A. M. C. and Y. S. acknowledge funding from ARO MURI, NSF (Grant No. CCF-1813814), and the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Quantum Algorithm Teams, Quantum Testbed Pathfinder (Award No. DE-SC0019040), and Accelerated Research in Quantum Computing (Award No. DE-SC0020312) programs. M. C. T. is supported in part by the NSF Grant No. NSF PHY-1748958 and the Heising-Simons Foundation. S. K. C. also acknowledges the support from the Studying Abroad Scholarship by Ministry of Education in Taiwan (R. O. C.).

[1] S. Lloyd, Science **273**, 1073 (1996).
[2] M. Suzuki, J. Math. Phys. (N.Y.) **26**, 601 (1985).
[3] J. Huyghebaert and H. D. Raedt, J. Phys. A **23**, 5777 (1990).
[4] A. M. Childs and Y. Su, Phys. Rev. Lett. **123**, 050503 (2019).
[5] A. M. Childs, Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2004.
[6] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Commun. Math. Phys. **270**, 359 (2007).
[7] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Proc. Natl. Acad. Sci. U.S.A. **115**, 9456 (2018).
[8] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Phys. Rev. Lett. **114**, 090502 (2015).
[9] G. H. Low, V. Kliuchnikov, and N. Wiebe, arXiv:1907.11679.
[10] G. H. Low and I. L. Chuang, Phys. Rev. Lett. **118**, 010501 (2017).
[11] J. Haah, M. B. Hastings, R. Kothari, and G. Hao Low, arXiv:1801.03922.
[12] M. C. Tran, A. Y. Guo, Y. Su, J. R. Garrison, Z. Eldredge, M. Foss-Feig, A. M. Childs, and A. V. Gorshkov, Phys. Rev. X **9**, 031006 (2019).
[13] References [2–4] took into account the commutativity between some interaction terms in the Hamiltonian of a nearest-neighbor interacting system. Without this commutativity, the gate count would be $O(n^3 t^2)$ [5,6].
[14] M. Heyl, P. Hauke, and P. Zoller, Sci. Adv. **5**, eaau8342 (2019).
[15] For interactions of maximum range R, we can group $\lceil (R+1)/2 \rceil$ consecutive sites into distinct blocks such that the Hamiltonian consists of only interactions between nearest-neighbor blocks. The error analysis for using the first-order product formula to simulate such a system would follow from our analysis for simulating nearest-neighbor interactions. Note, however, that we assume that the exact simulation of the evolution of each constant-size block requires only a constant amount of elementary gates.
[16] While it is common in practice to divide the evolution into even time slices, our interpretation of Eq. (3) as a sum of the same vector δ rotated by different angles suggests that we can also achieve destructive error interference even if the slices are uneven.

[17] D. W. Berry, A. M. Childs, and R. Kothari, in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (IEEE, New York, 2015), pp. 792–809.

[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.220502 for detailed proofs of Lemmas 1 and 2 and Eq. (16).

[19] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).

[20] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and Garnet Kin-Lic Chan, Phys. Rev. X **8**, 011044 (2018).

[21] A. W. Sandvik, AIP Conf. Proc. **1297**, 135 (2010).