Nearly tight Trotterization of correlated electrons

Yuan Su[1], Hsin-Yuan (Robert) Huang[1,2], and Earl T. Campbell[3]

[1]Institute for Quantum Information and Matter, Caltech, Pasadena, CA 91125, USA
[2]Department of Physics, Caltech, Pasadena, CA 91125, USA
[3]AWS Center for Quantum Computing, Pasadena, CA 91125, USA

November 20, 2020

Abstract

We consider simulating quantum systems on digital quantum computers. We show that the performance of quantum simulation can be improved by simultaneously exploiting the commutativity of Hamiltonian, the sparsity of interactions, and the prior knowledge of initial state. We achieve this using Trotterization for a class of correlated electrons that encompasses various physical systems, including the plane-wave-basis electronic structure and the Fermi-Hubbard model. We estimate the simulation error by taking the transition amplitude of nested commutators of Hamiltonian terms within the η-electron manifold. We develop multiple techniques for bounding the transition amplitude and the expectation of general fermionic operators, which may be of independent interest. We show that it suffices to use $\mathcal{O}\left(\frac{n^{5/3}}{\eta^{2/3}} + n^{4/3}\eta^{2/3}\right)$ gates to simulate electronic structure in the plane-wave basis with n spin orbitals and η electrons up to a negligible factor, improving the best previous result in second quantization while outperforming the first-quantized simulation when $\eta = \Omega\left(\sqrt{n}\right)$. We also obtain an improvement for simulating the Fermi-Hubbard model. We construct concrete examples for which our bounds are almost saturated, giving a nearly tight Trotterization of correlated electrons.

Contents

1 Introduction

Simulating quantum systems to model their dynamics and energy spectra is one of the most promising applications of digital quantum computers. Indeed, the difficulty of performing such simulations on classical computers led Feynman [22] and others to propose the idea of quantum computation. In 1996, Lloyd proposed the first explicit quantum algorithm for simulating local Hamiltonians [29]. Since then, various quantum simulation algorithms have been developed [7–10, 14, 17, 23, 31–34, 38], with potential applications in simulating condensed matter physics, quantum chemistry, and quantum field theories.

Lloyd's original work considered the simulation of k-local Hamiltonians. This was subsequently extended to the study of d-sparse Hamiltonians [1, 7], which provides a framework that highly abstracts the design of quantum simulation algorithms from the actual physical settings. However, despite their theoretical values, algorithms for sparse Hamiltonian simulation do not always provide the fastest approach for simulating concrete physical systems. Hamiltonians arising in practice often have additional features beyond sparseness, such as locality [23, 45], commutativity [18, 19, 42], and symmetry [46], that can be used to improve the performance of simulation. Besides, prior knowledge of the initial state [5, 20, 21, 41] and the norm distribution of Hamiltonian terms [14, 16, 24, 36] have also been proven useful for quantum simulation.

We show that a combination of these features, in particular the sparsity, commutativity, and initial-state information, can be used to give an even faster simulation. We achieve this improvement for a class of correlated electronic Hamiltonians, which includes many physically relevant systems such as the plane-wave-basis electronic-structure Hamiltonian and the Fermi-Hubbard model. Our approach uses Trotterization—by far the most widely applied method in experimental realizations of quantum simulation.

Our analysis proceeds by computing the transition amplitude of simulation error within the η-electron manifold. To this end, we develop multiple techniques for bounding the transition amplitude/expectation of a general fermionic operator, which may be of independent interest. For an n-spin-orbital electronic-structure problem in the plane-wave basis, our result improves the best previous result in second quantization [5, 19, 34] up to a negligible factor while outperforming the first-quantized result [3] for $\eta = \Omega\left(\sqrt{n}\right)$. We also obtain an improvement for simulating the Fermi-Hubbard model. We construct concrete examples for which our bounds are almost saturated, giving a nearly tight Trotterization of correlated electrons.

1.1 Combining sparsity, commutativity, and initial-state information

Sparsity can be used to improve quantum simulation in multiple ways. A common notion of d-sparsity concerns the target Hamiltonian itself, where each row and column of the Hamiltonian contains d nonzero elements accessed by querying quantum oracles. As aforementioned, this provides an abstract framework for designing efficient simulation algorithms and is versatile for establishing lower bounds [7], although it sometimes ignores other important properties of the target system, such as locality, commutativity, and symmetry. Another notion of sparsity, closely related to our paper, considers the interactions between the underlying qubits or modes [11, 39, 48]. The sparsity of interactions does not in general imply the underlying Hamiltonian is sparse, but it provides a tighter bound on the number of terms in the Hamiltonian and may thus be favorable to quantum simulation.

Trotterization and its alternative variants [17, 19, 24, 33, 38] provide a simple approach to quantum simulation and are by far the only known approach that can exploit the commutativity of Hamiltonian. Indeed, in the extreme case where all the terms in the Hamiltonian commute, we can

simultaneously diagonalize them and apply the first-order Lie-Trotter formula $\mathscr{S}_1(t)$ without error. Previous studies have also established commutator error bounds for certain low-order formulas [43] and specific systems [18, 42]. An analysis of a general formula $\mathscr{S}_p(t)$ is, however, considerably more difficult and has remained elusive until the recent proof of the commutator scaling of Trotter error [19].

A different direction to speeding up quantum simulation is to exploit the initial-state information. The error of quantum simulation is commonly quantified in previous work by the spectral-norm distance, which considers all possible states in the underlying Hilbert space. But if the state is known to be within some subspace throughout the simulation, then in principle this knowledge could be used to improve the analysis. For instance, quantum simulation in practice often starts with an initial state in the low-energy subspace of the Hamiltonian, so a worst-case spectral-norm analysis will inevitably overestimate the error. To address this, recent studies have considered a low-energy projection on the simulation error and provided improved approaches, using either Trotterization [5, 20, 21, 41] or more advanced quantum algorithms [30], that can be advantageous when the energy of initial state is sufficiently small.

Ideally, the sparsity of interactions, the commutativity of Hamiltonian, and the prior knowledge about the initial state can be combined to yield an even faster quantum simulation. This combination, however, appears to be technically challenging to achieve. Indeed, the state-of-the-art analysis of Trotterization represents simulation error in terms of nested commutators of Hamiltonian terms with exponential conjugations [19, Theorem 10]. This error representation is versatile for computing the commutator scaling of Trotter error, but it yields little information about the energy of initial state. To the best of our knowledge, the only previous attempt to address this problem was made by Somma for simulating bosonic Hamiltonians [41], whose solution unfortunately suffers from a divergence issue. Instead, we combine the sparsity, commutativity and the initial-state information to give an improved simulation of a class of correlated electrons.

1.2 Simulating correlated electrons

Simulating correlated electrons has emerged as one of the most important applications of digital quantum simulation [6]. The first efficient quantum algorithm for simulating electronic Hamiltonians is developed by Aspuru-Guzik et al. [2] using phase estimation; later developments have dramatically reduced the cost through various techniques [15, 35].

Here, we consider simulating the following class of correlated electrons by Trotterization:

$$H = T + V := \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} \nu_{l,m} N_l N_m, \tag{1}$$

where $A_j^\dagger$ and A_k are the fermionic *creation* and *annihilation operators*, N_l are the *occupation-number operators*, τ and ν are coefficient matrices, and the summation is over n spin orbitals. The specific definitions of these fermionic operators are given in Section 2.2. We say the interactions are *d-sparse* if there are at most d nonzero elements within each row/column of τ and ν. This model represents various systems arising in physics and chemistry, including the electronic-structure Hamiltonians in the plane-wave basis [5] and the Fermi-Hubbard model [20, 26].

To apply Trotterization, we need to express the Hamiltonian as a sum of elementary terms, each of which can be directly exponentiated on a quantum computer; see Section 2.1 for a review of this algorithm. For the electronic Hamiltonian (1), it suffices to consider the two-term decomposition $H = T + V$, as the exponentials of T and V can be directly implemented using various quantum circuits. For instance, all the terms in V commute with each other so $e^{-itV} = \prod_{l,m} e^{-it\nu_{l,m} N_l N_m}$, where

each $e^{-it\nu_{l,m}N_lN_m}$ corresponds to a two-qubit gate under the Jordan-Wigner transformation. On the other hand, exponential e^{-itT} can be implemented by diagonalization, i.e., $e^{-itT} = Ue^{-i\sum\lambda_\ell N_\ell}U^\dagger$, where U can be efficiently implemented using Givens rotations [27, 40]. In cases where τ and ν are translationally invariant $\tau_{j,k} = \tau_{j+q,k+q}$, $\nu_{l,m} = \nu_{l+q,m+q}$, we can implement e^{-itT} using the fast fermionic Fourier transform [5] and a related circuit implementation exists for e^{-itV} [34].

We now apply a pth-order Trotterization $\mathscr{S}_p(t)$ to approximate the evolution of the electronic Hamiltonian (1) for time t. We prove the following bound on the error of this approximation.

Theorem 1 (Fermionic-seminorm of Trotter error)**.** *Let $H = \sum_{j,k}\tau_{j,k}A_j^\dagger A_k + \sum_{l,m}\nu_{l,m}N_lN_m$ be a correlated-electronic Hamiltonian (1) with n orbitals, which we simulate using a pth-order product formula $\mathscr{S}_p(t)$. Then,*

$$\left\|\mathscr{S}_p(t) - e^{-itH}\right\|_\eta = \mathcal{O}\left(\left(\|\tau\| + \|\nu\|_{\max}\eta\right)^{p-1}\|\tau\|\,\|\nu\|_{\max}\,\eta^2 t^{p+1}\right). \tag{2}$$

Furthermore, if the interactions are d sparse,

$$\left\|\mathscr{S}_p(t) - e^{-itH}\right\|_\eta = \mathcal{O}\left(\left(\|\tau\|_{\max} + \|\nu\|_{\max}\right)^{p-1}\|\tau\|_{\max}\,\|\nu\|_{\max}\,d^{p+1}\eta t^{p+1}\right). \tag{3}$$

Here, $\|\cdot\|$ is the spectral norm*, $\|\cdot\|_{\max}$ is the* max-norm *denoting the largest matrix element in absolute value, and*

$$\|X\|_\eta := \max_{|\psi_\eta\rangle,|\phi_\eta\rangle}|\langle\phi_\eta|X|\psi_\eta\rangle| \tag{4}$$

is the fermionic seminorm *for number-preserving operator X, where $|\psi_\eta\rangle$, $|\phi_\eta\rangle$ are quantum states with η electrons.*

This theorem follows from an inductive estimate of the fermionic seminorm of nested commutators of Hamiltonian terms, and will be formally proved in Section 3 and Section 4. Note that in order to use prior knowledge of the initial state, we have considered the fermionic seminorm $\|\cdot\|_\eta$ of Trotter error with respect to the η-electron manifold. This seminorm is closely related to the other metrics used by previous work to quantify the impact of initial-state information to quantum simulation [5, 20, 21, 41]; see Section 2.3 for a detailed discussion. The resulting bound depends on the number of electrons η, as well as the spectral norm $\|\tau\|$, the max-norm $\|\tau\|_{\max}, \|\nu\|_{\max}$, and the sparsity d of interactions, but there is no dependence on the total number of spin orbitals n. This improves over previous work [19, Theorem D.5] where an explicit n-scaling seems unavoidable. Meanwhile, other prior estimates of the fermionic seminorm [5, Appendix G] [20, Theorem 13] did not exploit the commutativity of Hamiltonian and would introduce an addition factor of η^p in the Trotter error bound. Our result thus improves the performance of simulation by combining the initial-state information, the interaction sparsity, and the commutativity of Hamiltonian.

A common issue with the Trotterization algorithm is that existing analyses can be very loose for simulating specific physical systems. However, we address this with the following theorem, which shows that the asymptotic scaling of our bound is nearly tight.

Theorem 2 (Tightness)**.** *For $s, w > 0$, there exists a correlated-electronic Hamiltonian $H = T + V = \sum_{j,k}\tau_{j,k}A_j^\dagger A_k + \sum_{l,m}\nu_{l,m}N_lN_m$ such that*

$$\|\tau\| = s, \quad \|\nu\|_{\max} = w, \quad \left\|\underbrace{[T,\ldots[T}_{p},V]]\right\|_\eta = \Omega\left(s^p w\eta\right), \quad \left\|\underbrace{[V,\ldots[V}_{p},T]]\right\|_\eta = \Omega\left((w\eta)^p\, s/n\right). \tag{5}$$

In addition, for $u, w > 0$ and positive integer $1 \leq d \leq \eta$, there exists a d-sparse correlated-electronic Hamiltonian such that

$$\|\tau\|_{\max} = u, \quad \|\nu\|_{\max} = w, \quad \left\| \underbrace{[T, \ldots [T}_{p}, V]] \right\|_\eta = \Omega\left((ud)^p w\eta\right), \quad \left\| \underbrace{[V, \ldots [V}_{p}, T]] \right\|_\eta = \Omega\left((wd)^p u\right). \tag{6}$$

We prove the above theorem by choosing $T = \sum_{j,k=0}^{n-1} A_j^\dagger A_k$ and $V = \sum_{l,m=0}^{n/2-1} N_l N_m$ and computing their rescaled nested commutators, both in the original basis and the Fourier basis; see Section 5 for the proof. Note that both commutators $[T, \ldots [T, V]]$ and $[V, \ldots, [V, T]]$ contribute to the Trotter error, as well as other types of nested commutators which do not dominate the error scaling (Proposition 1). Modulo an application of the triangle inequality, Theorem 2 then shows that our bound (2) overestimates the Trotter error by a factor of $n\eta$ in the worst case, whereas (3) overestimates a factor of at most η. For p sufficiently large, this only contributes $n^{o(1)}$ and $\eta^{o(1)}$ to the gate complexity, respectively. In this sense, we have given a nearly tight Trotterization of the correlated-electronic Hamiltonians (1).

1.3 Main techniques

The proof of Theorem 1 relies on multiple approaches we develop for bounding the fermionic seminorm, which may be of independent interest. Recall from (4) that the fermionic seminorm $\|X\|_\eta$ for a number-preserving operator X is the maximum transition amplitude of X within the η-electron manifold.

Our first approach is based on the observation that the fermionic seminorm of a number-preserving X can be alternatively represented using the expectation of $X^\dagger X$, i.e.,

$$\|X\|_\eta = \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} |\langle \phi_\eta | X | \psi_\eta \rangle| = \max_{|\psi_\eta\rangle} \sqrt{\langle \psi_\eta | X^\dagger X | \psi_\eta \rangle}. \tag{7}$$

We then upper bound $X^\dagger X$ in terms of the *particle-number operator* $N = \sum_j N_j$, so that the expectation scales with the number of electrons $\eta = \langle \psi_\eta | N | \psi_\eta \rangle$ instead of the total number of spin orbitals. Assuming X is a sum of product of fermionic operators, we contract the summation indices in $X^\dagger X$ by either diagonalization or using an operator Cauchy-Schwarz inequality, followed by an application of a Hölder-type inequality. We perform this procedure recursively to simplify $X^\dagger X$. We detail this recursive approach in Section 3 and apply it to prove (2).

Our second approach starts by bounding the fermionic seminorm of X in terms of its maximum expectation value:

$$\|X\|_\eta = \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} |\langle \phi_\eta | X | \psi_\eta \rangle| \leq 2 \max_{|\psi_\eta\rangle} |\langle \psi_\eta | X | \psi_\eta \rangle|. \tag{8}$$

We then expand X and $|\psi_\eta\rangle$ and give a combinatorial argument to count the number of "paths" which have nonzero contribution to the expectation. We discuss this path-counting approach in more detail in Section 4 and use it to prove (3). The underlying idea of path counting is conceptually simple and may have potential applications in other contexts beyond the analysis of Trotter error.

1.4 Applications

The nearly tight Trotterization of electronic Hamiltonian (1) gives improved simulations of many systems arising in condensed matter physics and quantum chemistry, including the plane-wave-basis electronic-structure Hamiltonian and the Fermi-Hubbard model.

The electronic-structure problem considers electrons interacting with each other and some fixed nuclei. An efficient simulation of such systems could help understand chemical reactions, and provide insight into material properties. Here, we consider representing the electronic-structure Hamiltonian in the plane-wave basis [5]:

$$\begin{aligned} H = & \frac{1}{2n} \sum_{j,k,\mu} \kappa_\mu^2 \cos[\kappa_\mu \cdot r_{k-j}] A_j^\dagger A_k \\ & - \frac{4\pi}{\omega} \sum_{j,\iota,\mu \neq 0} \frac{\zeta_\iota \cos[\kappa_\mu \cdot (\widetilde{r}_\iota - r_j)]}{\kappa_\mu^2} N_j + \frac{2\pi}{\omega} \sum_{\substack{j \neq k \\ \mu \neq 0}} \frac{\cos[\kappa_\mu \cdot r_{j-k}]}{\kappa_\mu^2} N_j N_k, \end{aligned} \tag{9}$$

where ω is the volume of the computational cell, $\kappa_\mu = 2\pi\mu/\omega^{1/3}$ are n vectors of plane-wave frequencies, μ 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; and $\widetilde{r}_\iota$ are the nuclear coordinates. For this Hamiltonian, we have

$$\|\tau\| = \mathcal{O}\left(\frac{n^{2/3}}{\omega^{2/3}}\right), \qquad \|\nu\|_{\max} = \mathcal{O}\left(\frac{n^{1/3}}{\omega^{1/3}}\right). \tag{10}$$

Assuming a constant system density $\eta = \mathcal{O}(\omega)$, Theorem 1 then implies that

$$\left\| \mathscr{S}_p(t) - e^{-itH} \right\|_\eta = \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3} \right)^p n^{1/3}\eta^{2/3} t^{p+1} \right). \tag{11}$$

This approximation is accurate for a short-time evolution. To simulate for a longer time, we divide the evolution into r steps and apply $\mathscr{S}_p(t/r)$ within each step, obtaining

$$\left\| \mathscr{S}_p^r(t/r) - e^{-itH} \right\|_\eta = \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3} \right)^p n^{1/3}\eta^{2/3} \frac{t^{p+1}}{r^p} \right). \tag{12}$$

Therefore,

$$r = \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3} \right) \left(n^{2/3}\eta^{1/3} \right)^{1/p} \right) \tag{13}$$

steps suffices to simulate for constant time and accuracy with a pth-order Trotterization. Implementing each step using the approach of [34, Sect. 5] and choosing p sufficiently large, we obtain the gate complexity

$$\left(\frac{n^{5/3}}{\eta^{2/3}} + n^{4/3}\eta^{2/3} \right) n^{o(1)}. \tag{14}$$

Up to the negligible factor $n^{o(1)}$, this improves the best previous result in second quantization while outperforming the first-quantized simulation when $\eta = \Omega(\sqrt{n})$. See Table 1 for a gate-count comparison. We discuss this in detail in Section 6.1.

We also consider applications to the Fermi-Hubbard model, which is believed to capture the physics of some high temperature superconductors. This model is classically challenging to simulate [28, 49], but is a potential candidate for near-term quantum simulation [12, 13]. We have

$$H = -s \sum_{\langle j,k \rangle, \sigma} \left(A_{j,\sigma}^\dagger A_{k,\sigma} + A_{k,\sigma}^\dagger A_{j,\sigma} \right) + v \sum_j N_{j,0} N_{j,1}, \tag{15}$$

Algorithm/Bound	n, η	$\eta = n$
Interaction-picture (Ref. [3], first quantization)	$\widetilde{\mathcal{O}}\left(n^{\frac{1}{3}}\eta^{\frac{8}{3}}\right)$	$\widetilde{\mathcal{O}}\left(n^3\right)$
Qubitization (Ref. [3], first quantization)	$\widetilde{\mathcal{O}}\left(n^{\frac{2}{3}}\eta^{\frac{4}{3}} + n^{\frac{1}{3}}\eta^{\frac{8}{3}}\right)$	$\widetilde{\mathcal{O}}\left(n^3\right)$
Interaction-picture (Ref. [34], second quantization)	$\widetilde{\mathcal{O}}\left(\frac{n^{8/3}}{\eta^{2/3}}\right)$	$\widetilde{\mathcal{O}}\left(n^2\right)$
Trotterization (Ref. [5], second quantization)	$\left(n^{5/3}\eta^{1/3} + n^{4/3}\eta^{5/3}\right) n^{o(1)}$	$n^{3+o(1)}$
Trotterization (Ref. [19], second quantization)	$\left(\frac{n^{7/3}}{\eta^{1/3}}\right) n^{o(1)}$	$n^{2+o(1)}$
Trotterization (Theorem 1, second quantization)	$\left(\frac{n^{5/3}}{\eta^{2/3}} + n^{4/3}\eta^{2/3}\right) n^{o(1)}$	$n^{2+o(1)}$

Table 1: Comparison of our result and previous results for simulating the plane-wave-basis electronic structure with n spin orbitals and η electrons. We use $\widetilde{\mathcal{O}}(\cdot)$ to suppress polylogarithmic factors in the gate complexity scaling.

where $\langle j, k\rangle$ denotes a summation over nearest-neighbor lattice sites and $\sigma \in \{0, 1\}$. The Fermi-Hubbard model represents a lattice system with nearest-neighbor interactions and, according to [18], can be simulated with $\mathcal{O}\left(n^{1+1/p}\right)$ gates using a pth-order Trotterization for constant time and accuracy. On the other hand, recent work [20] shows that Trotterization algorithm has gate complexity $\mathcal{O}\left(n\eta^{1+1/p}\right)$ when restricted to the η-electron manifold. By simultaneously using the sparsity of interactions, the commutativity of Hamiltonian and the initial-state information, we show in Section 6.2 that $\mathcal{O}\left(n\eta^{1/p}\right)$ gates suffices, improving both Trotterization results for the Fermi-Hubbard model.[1]

We conclude the paper in Section 7 with a discussion of the results and some open questions.

2 Preliminaries

In this section, we summarize preliminaries of this paper, including a discussion of the Trotterization algorithm and their error analysis in Section 2.1, a brief summary of the second-quantization representation in Section 2.2, and an introduction to the fermionic seminorm and their properties in Section 2.3.

2.1 Trotterization and Trotter error

The Trotterization algorithm approximates the evolution of a sum of Hamiltonian terms using exponentials of the individual terms. For the correlated-electronic Hamiltonian (1), it suffices to consider a two-term Hamiltonian $H = T + V$, as the exponentials of T and V can be directly implemented on a quantum computer. Then, the ideal evolution under H for time t is given by $e^{-itH} = e^{-it(T+V)}$, which can be approximated by a pth-order product formula $\mathscr{S}_p(t)$, such as the first-order Lie-Trotter formula

$$\mathscr{S}_1(t) := e^{-itT}e^{-itV} \tag{16}$$

and $(2k)$th-order Suzuki formulas [44]

$$\begin{aligned} \mathscr{S}_2(t) &:= e^{-i\frac{t}{2}V}e^{-itT}e^{-i\frac{t}{2}V}, \\ \mathscr{S}_{2k}(t) &:= \mathscr{S}_{2k-2}(u_k t)^2\, \mathscr{S}_{2k-2}((1-4u_k)t)\, \mathscr{S}_{2k-2}(u_k t)^2, \end{aligned} \tag{17}$$

[1] Note however that this does not significantly improve the approach based on Lieb-Robinson bounds [23], since that approach has gate complexity $\widetilde{\mathcal{O}}(nt)$ when using a high-precision quantum simulation algorithm as subroutine.

where $u_k := 1/(4-4^{1/(2k-1)})$. This approximation is accurate when t is small. To simulate for a longer time, we divide the evolution into r *Trotter steps* and apply $\mathscr{S}_p(t/r)$ with *Trotter error* at most ϵ/r. We choose r sufficiently large so that the simulation error, as quantified by the spectral norm $\left\|\mathscr{S}_p^r(t/r) - e^{-itH}\right\|$, is at most ϵ.

Trotterization provides a simple approach to quantum simulation and is by far the only known approach that can exploit the commutativity of Hamiltonian. Indeed, in the extreme case where all the Hamiltonian terms commute, Trotterization can implement the exact evolution without error. Previous studies have also established commutator analysis of Trotter error for systems with geometrical locality and Lie-algebraic structures, as well as certain low-order formulas, including the first-order Lie-Trotter formula

$$\mathscr{S}_1(t) - e^{-itH} = \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \; e^{-i(t-\tau_1)H} e^{-i\tau_1 T} e^{i\tau_2 T} \left[iT, iV\right] e^{-i\tau_2 T} e^{-i\tau_1 V} \tag{18}$$

and the second-order Suzuki formula

$$\begin{aligned}\mathscr{S}_2(t) - e^{-itH} = &\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \int_0^{\tau_2} \mathrm{d}\tau_3 \; e^{-i(t-\tau_1)H} e^{-i\frac{\tau_1}{2}V} \\ &\cdot \left(e^{-i\tau_3 T} \left[-iT, \left[-iT, -i\frac{V}{2}\right]\right] e^{i\tau_3 T} + e^{i\frac{\tau_3}{2}V} \left[i\frac{V}{2}, \left[i\frac{V}{2}, iT\right]\right] e^{-i\frac{\tau_3}{2}V} \right) e^{-i\tau_1 T} e^{-i\frac{\tau_1}{2}V}.\end{aligned} \tag{19}$$

An analysis of the general case is, however, considerably more difficult and has remained elusive until the recent proof of commutator scaling of Trotter error [19]. Here, we introduce a stronger version of [19, Theorem 11] which can be proved by combining their Theorem 8, 9, and 10 without invoking the triangle inequality.

Proposition 1 (Commutator representation of Trotter error)**.** *Let $H = T + V$ be a two-term Hamiltonian and $\mathscr{S}_p(t)$ be a pth-order product formula. Define $H_0 = V$ and $H_1 = T$. Then,*

$$\begin{aligned}\mathscr{S}_p(t) - e^{-itH} = &\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \sum_{\gamma,j} a_{\gamma,j}(\tau_1,\tau_2) e^{-i(t-\tau_1)H} \\ &\cdot \mathscr{U}_{\gamma,j}(\tau_1,\tau_2) \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right] \mathscr{W}_{\gamma,j}(\tau_1,\tau_2),\end{aligned} \tag{20}$$

where γ goes over all $(p+1)$-dimensional binary vectors and j goes through a constant range of numbers (depending on the order p). Here, $\mathscr{U}_{\gamma,j}(\tau_1,\tau_2)$ and $\mathscr{W}_{\gamma,j}(\tau_1,\tau_2)$ are products of evolutions of T and V with time variables τ_1 and τ_2 and $a_\gamma(\tau_1,\tau_2)$ are coefficients such that

$$\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \; \left|a_{\gamma,j}(\tau_1,\tau_2)\right| = \mathcal{O}\left(t^{p+1}\right). \tag{21}$$

As an immediate application, we find that the spectral norm of Trotter error scales with nested commutators of the Hamiltonian terms, i.e.,

$$\left\|\mathscr{S}_p(t) - e^{-itH}\right\| = \mathcal{O}\left(\max_\gamma \left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\| t^{p+1}\right). \tag{22}$$

Note that the use of $\max_\gamma$ in place of $\sum_\gamma$ does not change the scaling as γ only ranges over constant number of binary vectors. This approximation is accurate for a short-time evolution. To evolve for a longer time, we divide the evolution into r steps and apply the triangle inequality to obtain

$$\left\|\mathscr{S}_p^r(t/r) - e^{-itH}\right\| \le r \left\|\mathscr{S}_p(t/r) - e^{-i\frac{t}{r}H}\right\| = \mathcal{O}\left(\max_\gamma \left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\| \frac{t^{p+1}}{r^p}\right). \tag{23}$$

It thus suffices to choose

$$r = \mathcal{O}\left(\frac{\left(\max_\gamma \left\| \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right] \right\|\right)^{1/p} t^{1+1/p}}{\epsilon^{1/p}}\right) \tag{24}$$

to ensure that the error of simulation is no more than ϵ.

The above analysis is versatile for computing the commutator dependence of Trotter error. Unfortunately, the resulting bound does not use prior knowledge of the initial state and will in particular be loose if the initial state lies within a low-energy subspace. On the other hand, recent work of Şahinoğlu and Somma proposed a Trotterization approach for simulating low-energy initial states but the commutativity of Hamiltonian was ignored in their analysis [21]. Here, we address this by simultaneously using the commutativity of Hamiltonian and the prior knowledge of initial state to improve quantum simulation of a class of correlated electrons. In the following, we introduce preliminaries about the second-quantization representation (Section 2.2) and the notion of fermionic seminorm (Section 2.3), on which our analysis will be based.

2.2 Second-quantization representation

In this section, we review several facts about the second-quantization representation that are relevant to our analysis. We refer the reader to the book of Helgaker, Jørgensen, and Olsen [25] for a detailed discussion of this topic.

We use the abstract Fock space to represent electronic Hamiltonians. Specifically, for a system of n spin orbitals, we construct a 2^n-dimensional Fock space $\mathrm{span}\{|\ell_1, \ell_2, \ldots, \ell_n\rangle\}$ spanned by the binary vectors $|\ell_1, \ell_2, \ldots, \ell_n\rangle$, where $\ell_j = 1$ represents that mode j is occupied and $\ell_j = 0$ otherwise. General vectors in the Fock space, denoted by $|\psi\rangle$ or $|\phi\rangle$, are then given by linear combinations of these orthonormal basis vectors. We define the *η-electron subspace* as $\mathrm{span}\{|\ell_1, \ell_2, \ldots, \ell_n\rangle, \sum_j \ell_j = \eta\}$. By considering all $0 \leq \eta \leq n$, we obtain the decomposition

$$\mathrm{span}\,\{|\ell_1, \ell_2, \ldots, \ell_n\rangle\} = \bigodot_{\eta=0}^{n} \mathrm{span}\left\{|\ell_1, \ell_2, \ldots, \ell_n\rangle, \sum_j \ell_j = \eta\right\}, \tag{25}$$

where $\odot$ denotes the orthogonal direct sum. We say that normalized vectors in the η-electron subspace form the *η-electron manifold* and denote an arbitrary such vector by $|\psi_\eta\rangle$ or $|\phi_\eta\rangle$.

The n elementary fermionic *creation operators* are defined through the relations

$$\begin{aligned} A_j^\dagger |\ell_1, \ell_2, \ldots, 0_j, \ldots, \ell_n\rangle &= (-1)^{\sum_{k=1}^{j-1} \ell_k} |\ell_1, \ell_2, \ldots, 1_j, \ldots, \ell_n\rangle, \\ A_j^\dagger |\ell_1, \ell_2, \ldots, 1_j, \ldots, \ell_n\rangle &= 0, \end{aligned} \tag{26}$$

whereas the fermionic *annihilation operators* are defined by

$$\begin{aligned} A_j |\ell_1, \ell_2, \ldots, 0_j, \ldots, \ell_n\rangle &= 0, \\ A_j |\ell_1, \ell_2, \ldots, 1_j, \ldots, \ell_n\rangle &= (-1)^{\sum_{k=1}^{j-1} \ell_k} |\ell_1, \ell_2, \ldots, 0_j, \ldots, \ell_n\rangle. \end{aligned} \tag{27}$$

The use of $\dagger$ is justified by the fact that $A_j^\dagger$ is indeed the Hermitian adjoint of A_j with respect to the inner product in the Fock space. We also introduce the *occupation-number operators* $N_j = A_j^\dagger A_j$ and add them together to get the *particle-number operator* $N = \sum_{j=1}^n N_j$.

Fermionic creation and annihilation operators satisfy the canonical anticommutation relations

$$A_j^\dagger A_k^\dagger + A_k^\dagger A_j^\dagger = A_j A_k + A_k A_j = 0, \qquad A_j^\dagger A_k + A_k A_j^\dagger = \delta_{j,k} I, \tag{28}$$

where the Kronecker-delta function $\delta_{j,k}$ is one if $j = k$ and zero otherwise. Applying these, we obtain the following commutation relations of second-quantized fermionic operators.

Proposition 2 (Commutation relations of fermionic operators)**.** *The following commutation relations hold for second-quantized fermionic operators:*

1. $\left[A_l^\dagger A_m, A_j^\dagger\right] = \delta_{j,m} A_l^\dagger$;
2. $\left[A_l^\dagger A_m, A_k\right] = -\delta_{k,l} A_m$;
3. $\left[N_l, A_j^\dagger\right] = \delta_{l,j} A_j^\dagger$, $\left[N, A_j^\dagger\right] = A_j^\dagger$;
4. $[N_l, A_k] = -\delta_{l,k} A_k$, $[N, A_k] = -A_k$;
5. $[N_l, N_m] = 0$.

We say a fermionic operator is *number-preserving* if every η-electron subspace is invariant under the action of this operator. Equivalently, operator X is number-preserving if and only if it commutes with the particle-number operator, i.e., $[N, X] = 0$. Yet another equivalent definition is based on the notion of *η-electron projections*: letting Π_η be orthogonal projections onto the η-electron subspaces, then X is number-preserving if and only if it commutes with every Π_η, namely, $[\Pi_\eta, X] = 0$. In the matrix representation, X is block-diagonalized by the set of η-electron projections $\{\Pi_\eta\}$.

A special example of number-preserving operator is the particle-number operator N, which acts as a scalar multiplication by η within the η-electron subspace. Other examples include *excitation operators* $A_j^\dagger A_k$, occupation-number operators N_l, and elementary exponentials $e^{-it\sum_{j,k}\tau_{j,k}A_j^\dagger A_k}$ that appear in the Trotterization algorithm. In fact, the following lemma shows that the set of number-preserving operators are closed under linear combination, multiplication, Hermitian conjugation, and taking limit.

Proposition 3 (Number-preserving operators as a closed †-subalgebra)**.** *The following operators are respectively number-preserving:*

1. *$\lambda X + \mu Y$, if X and Y are number-preserving, and λ and μ are complex numbers;*
2. *XY, if X and Y are number-preserving;*
3. *$X^\dagger$, if X is number-preserving;*
4. *$\lim_{i\to\infty} X_i$, if X_i are number-preserving and the limit exists.*

2.3 Fermionic seminorm

We now introduce the notion of fermionic seminorm, which we use to quantify the error of the Trotterization algorithm that takes the prior knowledge of initial state into consideration.

For any number-preserving operator X and $0 \le \eta \le n$, we define the *fermionic η-seminorm* as the maximum transition amplitude within the η-electron manifold:

$$\|X\|_\eta := \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle|, \tag{29}$$

where $|\psi_\eta\rangle, |\phi_\eta\rangle$ are quantum states containing η electrons.[2] When there is no ambiguity, we drop the dependence on η and call $\|X\|_\eta$ the fermionic seminorm of X. As the name suggests and the

following proposition confirms, the fermionic seminorm is indeed a seminorm defined on the closed †-subalgebra of number-preserving operators.

Proposition 4 (Seminorm properties)**.** *The following properties hold for the fermionic seminorm:*

1. $\|\lambda X\|_\eta = |\lambda| \, \|X\|_\eta$*, if* X *is number-preserving and* λ *is a complex number;*
2. $\|X + Y\|_\eta \leq \|X\|_\eta + \|Y\|_\eta$*, if* X *and* Y *are number-preserving;*
3. $\|XY\|_\eta \leq \|X\|_\eta \, \|Y\|_\eta$*, if* X *and* Y *are number-preserving;*
4. $\|I\|_\eta = 1$*;*
5. $\|UXW\|_\eta = \|X\|_\eta$*, if* U, X, W *are number-preserving and* U, W *are unitary;*
6. $\left\|X^\dagger\right\|_\eta = \|X\|_\eta$*, if* X *is number-preserving;*

Proof. We will only prove the third statement, as the remaining follow directly from the definition of the fermionic seminorm. We consider

$$\begin{aligned} \|XY\|_\eta &= \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} |\langle \phi_\eta | XY | \psi_\eta \rangle| \\ &= \max_{|\psi_\eta\rangle, |\phi_\eta\rangle} |\langle \phi_\eta | X \Pi_\eta \Pi_\eta Y | \psi_\eta \rangle| \\ &\leq \max_{|\phi_\eta\rangle} \left\| \Pi_\eta X^\dagger |\phi_\eta\rangle \right\| \max_{|\psi_\eta\rangle} \|\Pi_\eta Y |\psi_\eta\rangle\| \,, \end{aligned} \tag{30}$$

where the last step follows from the Cauchy-Schwarz inequality. To proceed, we optimize over arbitrary state $|\varphi\rangle$ to get

$$\begin{aligned} \left\| \Pi_\eta X^\dagger |\phi_\eta\rangle \right\| &= \max_{|\varphi\rangle} \left| \langle \varphi | \Pi_\eta X^\dagger | \phi_\eta \rangle \right| \\ &= \max_{|\varphi\rangle} \|\Pi_\eta |\varphi\rangle\| \left| \frac{\langle \varphi | \Pi_\eta}{\|\Pi_\eta |\varphi\rangle\|} X^\dagger |\phi_\eta\rangle \right| \\ &\leq \left\| X^\dagger \right\|_\eta = \|X\|_\eta \end{aligned} \tag{31}$$

assuming $\Pi_\eta |\varphi\rangle \neq 0$, as the case $\Pi_\eta |\varphi\rangle = 0$ never leads to maximality. But on the other hand,

$$\begin{aligned} \|X\|_\eta = \left\| X^\dagger \right\|_\eta &= \max_{|\phi_\eta\rangle, |\varphi_\eta\rangle} \left| \langle \varphi_\eta | X^\dagger | \psi_\eta \rangle \right| \\ &= \max_{|\phi_\eta\rangle, |\varphi_\eta\rangle} \left| \langle \varphi_\eta | \Pi_\eta X^\dagger | \psi_\eta \rangle \right| \\ &\leq \max_{|\phi_\eta\rangle, |\varphi\rangle} \left| \langle \varphi | \Pi_\eta X^\dagger | \psi_\eta \rangle \right| = \max_{|\phi_\eta\rangle} \left\| \Pi_\eta X^\dagger |\phi_\eta\rangle \right\| , \end{aligned} \tag{32}$$

implying $\max_{|\phi_\eta\rangle} \left\| \Pi_\eta X^\dagger |\phi_\eta\rangle \right\| = \|X\|_\eta$. Similarly, we have $\max_{|\psi_\eta\rangle} \|\Pi_\eta Y |\psi_\eta\rangle\| = \|Y\|_\eta$. This completes the proof of the third statement. □

[2]Note that it is possible to extend this to define $\|\cdot\|_{\eta\to\xi}$ for operators that map the η-electron subspace to ξ-electron subspace, although this is not needed in our analysis and will not be further pursued here.

The fermionic seminorm, as defined in (29) by the maximum transition amplitude within an η-electron manifold, provides a reasonable metric for quantifying the error of quantum simulation with initial-state constraints. Indeed, a seminorm similar to our definition was used by Somma [41] for analyzing quantum simulation of bosonic Hamiltonians. However, we point out that this is not the only error metric that takes the prior knowledge of initial state into account. Recent work [21] analyzed the low-energy simulation of k-local frustration-free Hamiltonians by computing the spectral norm of Trotter error projected on the low-energy subspace. However, the following proposition shows that these two error metrics are the same for fermionic systems.

Proposition 5 (Fermionic seminorm as a projected spectral norm)**.** *For number-preserving operator X, it holds that*

$$\|X\|_\eta = \max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| = \|X\Pi_\eta\|\,. \tag{33}$$

Proof. The underlying idea behind this proposition is already hinted in the proof of Proposition 4. We have

$$\begin{aligned}\max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| &= \max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|\Pi_\eta X\Pi_\eta|\psi_\eta\rangle| \\ &\leq \max_{|\psi\rangle,|\phi\rangle} |\langle\phi|\Pi_\eta X\Pi_\eta|\psi\rangle| = \|\Pi_\eta X\Pi_\eta\|\,.\end{aligned} \tag{34}$$

But on the other hand,

$$\begin{aligned}\|\Pi_\eta X\Pi_\eta\| &= \max_{|\psi\rangle,|\phi\rangle} |\langle\phi|\Pi_\eta X\Pi_\eta|\psi\rangle| \\ &= \max_{|\psi\rangle,|\phi\rangle} \|\Pi_\eta|\phi\rangle\|\,\|\Pi_\eta|\psi\rangle\| \left|\frac{\langle\phi|\Pi_\eta}{\|\Pi_\eta|\phi\rangle\|} X \frac{\Pi_\eta|\psi\rangle}{\|\Pi_\eta|\psi\rangle\|}\right| \\ &\leq \max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle|\end{aligned} \tag{35}$$

assuming $\Pi_\eta|\phi\rangle \neq 0$ and $\Pi_\eta|\psi\rangle \neq 0$, as the zero vector will not lead to maximality. The proposition then follows since number-preserving operator X commutes with the η-electron projection Π_η. □

Another common approach to quantify the simulation error is to take the expectation within the η-electron subspace. This approach is used by previous work [5, 40] and appears to give a natural metric when quantum simulation is used as a subroutine in phase estimation. We show that this only differs from our definition (29) by at most a constant factor, reaffirming the fermionic seminorm as a proper error metric for simulating fermionic systems.

Proposition 6 (Transition amplitude and expectation)**.** *For number-preserving operator X, the following statements hold:*

1. $\max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| = \max_{|\psi_\eta\rangle} |\langle\psi_\eta|X|\psi_\eta\rangle|$, *if X is Hermitian;*
2. $\max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| = \max_{|\psi_\eta\rangle} \sqrt{\langle\psi_\eta|X^\dagger X|\psi_\eta\rangle}$;
3. $\max_{|\psi_\eta\rangle} |\langle\psi_\eta|X|\psi_\eta\rangle| \leq \max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| \leq 2\max_{|\psi_\eta\rangle} |\langle\psi_\eta|X|\psi_\eta\rangle|$.

Proof. The first statement follows from the fact that $\Pi_\eta X\Pi_\eta$ is Hermitian and that the spectral norm of a Hermitian operator is its largest eigenvalue in absolute value. For the second statement,

$$\begin{aligned}\max_{|\psi_\eta\rangle,|\phi_\eta\rangle} |\langle\phi_\eta|X|\psi_\eta\rangle| &= \|X\|_\eta = \|X\Pi_\eta\| \\ &= \sqrt{\|\Pi_\eta X^\dagger X\Pi_\eta\|} \\ &= \max_{|\psi_\eta\rangle} \sqrt{\langle\psi_\eta|X^\dagger X|\psi_\eta\rangle}.\end{aligned} \tag{36}$$

The first inequality of Statement 3 is trivial. For the second inequality, we apply the polarization identity

$$\begin{aligned}\langle\phi_\eta|X|\psi_\eta\rangle = \frac{1}{4}\Big(&\left(\langle\phi_\eta| + \langle\psi_\eta|\right) X \left(|\phi_\eta\rangle + |\psi_\eta\rangle\right) - \left(\langle\phi_\eta| - \langle\psi_\eta|\right) X \left(|\phi_\eta\rangle - |\psi_\eta\rangle\right) \\ &- i\left(\langle\phi_\eta| - i\langle\psi_\eta|\right) X \left(|\phi_\eta\rangle + i|\psi_\eta\rangle\right) + i\left(\langle\phi_\eta| + i\langle\psi_\eta|\right) X \left(|\phi_\eta\rangle - i|\psi_\eta\rangle\right)\Big)\end{aligned} \tag{37}$$

to obtain

$$\begin{aligned}&|\langle\phi_\eta|X|\psi_\eta\rangle| \\ &\leq \frac{\max_{|\varphi_\eta\rangle}|\langle\varphi_\eta|X|\varphi_\eta\rangle|}{4}\left(\||\phi_\eta\rangle + |\psi_\eta\rangle\|^2 + \||\phi_\eta\rangle - |\psi_\eta\rangle\|^2 + \||\phi_\eta\rangle + i|\psi_\eta\rangle\|^2 + \||\phi_\eta\rangle - i|\psi_\eta\rangle\|^2\right) \\ &= 2\max_{|\varphi_\eta\rangle}|\langle\varphi_\eta|X|\varphi_\eta\rangle|,\end{aligned} \tag{38}$$

from which the claimed inequality follows by maximizing over $|\psi_\eta\rangle$ and $|\phi_\eta\rangle$. □

We now apply Proposition 1 to compute the fermionic seminorm of Trotter error, obtaining

$$\left\|\mathscr{S}_p(t) - e^{-itH}\right\|_\eta = \mathcal{O}\left(\max_\gamma \left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\|_\eta t^{p+1}\right). \tag{39}$$

We find that the resulting error bound depends on the fermionic seminorm of nested commutators, and the performance of quantum simulation can thus be potentially improved by simultaneously exploiting the commutativity of Hamiltonian and the prior knowledge of initial state. However, the main technical challenge here is to give a tight estimate of $\left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\|_\eta$, which has not been addressed in previous literature. To this end, we develop two approaches for bounding the expectation/transition amplitude of general fermionic operators in Section 3 and Section 4 to prove our main result Theorem 1, establish the tightness of our bound in Section 5, and discuss applications and further implications of our result in Section 6 and Section 7.

3 Bounding expectation of fermionic operators by recursion

In this section, we present the first approach for bounding the expectation of fermionic operators, and thereby bounding the fermionic seminorm of Trotter error. We introduce in Section 3.1 the main techniques used in our approach, including an operator Cauchy-Schwarz inequality, a diagonalization procedure, and a Hölder-type inequality for the expectation value. We then describe our approach in detail and apply it to prove Eq. (2) of our main result Theorem 1. The proof is based on induction: we analyze the base case in Section 3.2 and the inductive step in Section 3.3, respectively.

3.1 Main techniques

Recall that the main technical challenge to estimate the simulation error of the electronic Hamiltonian (1) is to bound the fermionic seminorm $\left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\|_\eta$, where $\gamma = 0, 1$, $H_0 = V$ and $H_1 = T$. Applying the commutation relations in Proposition 2, we see that we need to analyze a general fermionic operator of the form

$$X = \sum_{j,k,l} w_{j,k,l} \cdots A^\dagger_{j_x} \cdots A_{k_y} \cdots N_{l_z} \cdots \tag{40}$$

Our first approach starts by reexpressing the fermionic seminorm of X using expectation of $X^\dagger X$:

$$\|X\|_\eta = \max_{|\psi_\eta\rangle} \sqrt{\langle\psi_\eta|X^\dagger X|\psi_\eta\rangle}. \tag{41}$$

We note that $X^\dagger X$ is a positive semidefinite operator, and an upper bound of it with respect to the partial ordering of positive semidefiniteness will therefore give a bound on the expectation value. We achieve this by contracting the corresponding indices in X and $X^\dagger$, using either an operator Cauchy-Schwarz inequality (Lemma 1) or diagonalization (Lemma 2).

Lemma 1 (Operator Cauchy-Schwarz inequality [37, Proposition 3.4])**.** *For any finite lists of operators $\{B_j\}$ and $\{C_j\}$ with the same cardinality, we have*

$$-\sum_{j,k} B_j^\dagger C_k^\dagger C_k B_j \leq \sum_{j,k} B_j^\dagger C_k^\dagger C_j B_k \leq \sum_{j,k} B_j^\dagger C_k^\dagger C_k B_j, \tag{42}$$

where Hermitian operators are partially ordered according to the positive semidefiniteness.

Proof. We have

$$\begin{aligned} 0 &\leq \sum_{j,k} \left(C_k B_j \mp C_j B_k\right)^\dagger \left(C_k B_j \mp C_j B_k\right) \\ &= \sum_{j,k} \left(B_j^\dagger C_k^\dagger C_k B_j \mp B_k^\dagger C_j^\dagger C_k B_j \mp B_j^\dagger C_k^\dagger C_j B_k + B_k^\dagger C_j^\dagger C_j B_k\right) \\ &= 2\sum_{j,k} B_j^\dagger C_k^\dagger C_k B_j \mp 2\sum_{j,k} B_j^\dagger C_k^\dagger C_j B_k. \end{aligned} \tag{43}$$

This implies

$$\pm\sum_{j,k} B_j^\dagger C_k^\dagger C_j B_k \leq \sum_{jk} B_j^\dagger C_k^\dagger C_k B_j, \tag{44}$$

from which the claimed inequality follows. □

Lemma 2 (Diagonalization)**.** *For any finite list of operators $\{B_j\}$ and Hermitian coefficient matrix μ, we have*

$$-\|\mu\|\sum_j B_j^\dagger B_j \leq \sum_{j,k} \mu_{j,k} B_j^\dagger B_k \leq \|\mu\|\sum_j B_j^\dagger B_j, \tag{45}$$

where Hermitian operators are partially ordered according to the positive semidefiniteness.

Proof. Since μ is Hermitian, we may diagonalize it to $\widetilde{\mu}$ by unitary transformation w as

$$\mu = w^\dagger \widetilde{\mu} w, \tag{46}$$

where $\widetilde{\mu}$ is a diagonal matrix with all eigenvalues of μ as the diagonal elements. We then define $\widetilde{B}_l := \sum_k w_{l,k} B_k$ so that

$$\sum_{j,k} \mu_{j,k} B_j^\dagger B_k = \sum_l \widetilde{\mu}_l \widetilde{B}_l^\dagger \widetilde{B}_l, \tag{47}$$

which implies

$$-\|\mu\|\sum_l \widetilde{B}_l^\dagger \widetilde{B}_l \leq \sum_{j,k} \mu_{j,k} B_j^\dagger B_k \leq \|\mu\|\sum_l \widetilde{B}_l^\dagger \widetilde{B}_l. \tag{48}$$

But $\sum_l \widetilde{B}_l^\dagger \widetilde{B}_l$ has identity as the coefficient matrix which is invariant under a change of basis:

$$\sum_l \widetilde{B}_l^\dagger \widetilde{B}_l = \sum_j B_j^\dagger B_j. \tag{49}$$

This completes the proof. □

By applying Lemma 1 or Lemma 2, we can get a bound of $X^\dagger X$ with respect to the partial ordering of positive semidefiniteness, with one pair of the corresponding indices in X and $X^\dagger$ contracted. The following result then allows us to perform this contraction recursively to get a desired bound on the expectation value.

Lemma 3 (Hölder-type inequality for expectation)**.** *For any finite lists of fermionic operators $\{B_j\}$ and $\{C_k\}$ with the same cardinality,*

$$\max_{|\psi_\eta\rangle} \langle\psi_\eta| \sum_j B_j^\dagger C_j^\dagger C_j B_j |\psi_\eta\rangle \leq \max_{|\psi_\eta\rangle} \langle\psi_\eta| \sum_j B_j^\dagger B_j |\psi_\eta\rangle \max_{k,|\phi_\xi\rangle} \langle\phi_\xi| C_k^\dagger C_k |\phi_\xi\rangle, \tag{50}$$

where we assume each B_j maps the η-electron subspace to the ξ-electron subspace. In terms of the fermionic seminorm, we have

$$\left\| \sum_j B_j^\dagger C_j^\dagger C_j B_j \right\|_\eta \leq \left\| \sum_j B_j^\dagger B_j \right\|_\eta \max_k \left\| C_k^\dagger C_k \right\|_\xi. \tag{51}$$

Proof. The claimed inequality follows from

$$\begin{aligned} \left\| \sum_j B_j^\dagger C_j^\dagger C_j B_j \right\|_\eta &= \left\| \sum_j B_j^\dagger \Pi_\xi C_j^\dagger C_j \Pi_\xi B_j \right\|_\eta \\ &\leq \left\| \sum_j \left\| \Pi_\xi C_j^\dagger C_j \Pi_\xi \right\| B_j^\dagger B_j \right\|_\eta \\ &\leq \left\| \sum_j B_j^\dagger B_j \right\|_\eta \max_k \left\| \Pi_\xi C_k^\dagger C_k \Pi_\xi \right\| \\ &= \left\| \sum_j B_j^\dagger B_j \right\|_\eta \max_k \left\| C_k^\dagger C_k \right\|_\xi. \end{aligned} \tag{52}$$

□

Using the above lemmas, we can now prove Eq. (2) of our main result Theorem 1 by induction. We analyze the base case in Section 3.2 and the inductive step in Section 3.3.

3.2 Single-layer commutator

We now prove Eq. (2) of our main result Theorem 1 by induction. In the base case, we consider simulating the correlated-electronic Hamiltonian (1) using the first-order formula $\mathscr{S}_1(t)$. We know from (18) that

$$\left\| \mathscr{S}_1(t) - e^{-itH} \right\|_\eta \leq \frac{t^2}{2} \left\| [T, V] \right\|_\eta, \tag{53}$$

where $T = \sum_{j,k} \tau_{j,k} A_j^\dagger A_k$ and $V = \sum_{l,m} \nu_{l,m} N_l N_m$. Our goal is to show that

$$\left\| [T, V] \right\|_\eta = \mathcal{O}\left(\|\tau\| \, \|\nu\|_{\max} \, \eta^2 \right). \tag{54}$$

To this end, we apply Proposition 2 to expand the single-layer commutator $[T, V]$ into linear combinations of fermionic creation, annihilation, and occupation-number operators. We have

$$\begin{aligned}
[T, V] &= \sum_{j,k,l,m} \tau_{j,k}\nu_{l,m} \left[A_j^\dagger A_k, N_l N_m\right] \\
&= \sum_{j,k,l,m} \tau_{j,k}\nu_{l,m} A_j^\dagger \left[A_k, N_l N_m\right] + \sum_{j,k,l,m} \tau_{j,k}\nu_{l,m} \left[A_j^\dagger, N_l N_m\right] A_k \\
&= \sum_{j,k,m} \tau_{j,k}\nu_{k,m} A_j^\dagger A_k N_m + \sum_{j,k,l} \tau_{j,k}\nu_{l,k} A_j^\dagger N_l A_k \\
&\quad - \sum_{j,k,m} \tau_{j,k}\nu_{j,m} A_j^\dagger N_m A_k - \sum_{j,k,l} \tau_{j,k}\nu_{l,j} N_l A_j^\dagger A_k.
\end{aligned} \tag{55}$$

At this stage, it is possible to directly bound the terms in the last equality using Lemma 1, Lemma 2, and Lemma 3 from the previous subsection. However, we will further commute the occupation-number operator in between the creation and annihilation operators, obtaining

$$\begin{aligned}
[T, V] &= \sum_{j,k,m} \tau_{j,k}\nu_{k,m} A_j^\dagger N_m A_k + \sum_{j,k} \tau_{j,k}\nu_{k,k} A_j^\dagger A_k + \sum_{j,k,l} \tau_{j,k}\nu_{l,k} A_j^\dagger N_l A_k \\
&\quad - \sum_{j,k,m} \tau_{j,k}\nu_{j,m} A_j^\dagger N_m A_k - \sum_{j,k} \tau_{j,k}\nu_{j,j} A_j^\dagger A_k - \sum_{j,k,l} \tau_{j,k}\nu_{l,j} A_j^\dagger N_l A_k.
\end{aligned} \tag{56}$$

This additional commutation leads to an error bound with the same asymptotic scaling but a slightly larger prefactor. The benefit is that the analysis can be directly extended to handle the inductive step in the next subsection.

We now bound the asymptotic scaling for each of the six terms in the commutator expansion.

Proposition 7 (Structure of single-layer commutator)**.** *Let $H = T + V = \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} \nu_{l,m} N_l N_m$ be a correlated-electronic Hamiltonian (1). Then, the commutator $[T, V]$ has the expansion (56), where*

1. $\left\| \sum_{j,k,m} \tau_{j,k}\nu_{k,m} A_j^\dagger N_m A_k \right\|_\eta \leq \|\tau\| \, \|\nu\|_{\max} \, \eta^2$;
2. $\left\| \sum_{j,k} \tau_{j,k}\nu_{k,k} A_j^\dagger A_k \right\|_\eta \leq \|\tau\| \, \|\nu\|_{\max} \, \eta$;
3. $\left\| \sum_{j,k,l} \tau_{j,k}\nu_{l,k} A_j^\dagger N_l A_k \right\|_\eta \leq \|\tau\| \, \|\nu\|_{\max} \, \eta^2$;
4. $\left\| \sum_{j,k,m} \tau_{j,k}\nu_{j,m} A_j^\dagger N_m A_k \right\|_\eta \leq \|\tau\| \, \|\nu\|_{\max} \, \eta^2$;
5. $\left\| \sum_{j,k} \tau_{j,k}\nu_{j,j} A_j^\dagger A_k \right\|_\eta \leq \|\tau\| \, \|\nu\|_{\max} \, \eta$;
6. $\left\| \sum_{j,k,l} \tau_{j,k}\nu_{l,j} A_j^\dagger N_l A_k \right\|_\eta \leq \|\tau\| \, \|\nu\|_{\max} \, \eta^2$.

Proof. We describe the proof of the first two statements here. The remaining justifications proceed in a similar way and are left to Appendix A.

Letting $X = \sum_{j,k,m} \tau_{j,k}\nu_{k,m}A_j^\dagger N_m A_k$, we have $\|X\|_\eta = \sqrt{\|X^\dagger X\|_\eta}$. Now,

$$\begin{aligned} X^\dagger X &= \sum_{j_1,k_1,m_1,j_2,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_2,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}A_{j_1}A_{j_2}^\dagger N_{m_2}A_{k_2} \\ &= \sum_{j_1,k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_1,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}N_{m_2}A_{k_2} \\ &\quad - \sum_{j_1,k_1,m_1,j_2,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_2,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}A_{j_2}^\dagger A_{j_1}N_{m_2}A_{k_2}, \end{aligned} \tag{57}$$

where we have used the anti-commutation relation $A_{j_1}A_{j_2}^\dagger + A_{j_2}^\dagger A_{j_1} = \delta_{j_1,j_2}I$. For the second term, we define $B_{j_1}^\dagger = \sum_{k_1,m_1} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}A_{k_1}^\dagger N_{m_1}$ and apply the operator Cauchy-Schwarz inequality (Lemma 1):

$$\begin{aligned} &- \sum_{j_1,k_1,m_1,j_2,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_2,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}A_{j_2}^\dagger A_{j_1}N_{m_2}A_{k_2} \\ &= -\sum_{j_1,j_2} B_{j_1}^\dagger A_{j_2}^\dagger A_{j_1}B_{j_2} \le \sum_{j_1,j_2} B_{j_1}^\dagger A_{j_2}^\dagger A_{j_2}B_{j_1} = \sum_{j_1} B_{j_1}^\dagger N B_{j_1} \\ &= \sum_{j_1}\sum_{k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_1,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}NN_{m_2}A_{k_2} \\ &= \sum_{j_1}\sum_{k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_1,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}N_{m_2}A_{k_2}N \\ &\quad - \sum_{j_1}\sum_{k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_1,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}N_{m_2}A_{k_2}. \end{aligned} \tag{58}$$

This implies

$$\begin{aligned} X^\dagger X &\le \sum_{j_1}\sum_{k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{k_1,m_1}\tau_{j_1,k_2}\nu_{k_2,m_2}A_{k_1}^\dagger N_{m_1}N_{m_2}A_{k_2}N \\ &= \sum_{k_1,k_2}\left(\sum_{j_1}\bar{\tau}_{j_1,k_1}\tau_{j_1,k_2}\right)\left(\sum_{m_1}\bar{\nu}_{k_1,m_1}A_{k_1}^\dagger N_{m_1}\right)\left(\sum_{m_2}\nu_{k_2,m_2}N_{m_2}A_{k_2}\right)N. \end{aligned} \tag{59}$$

Note that $\sum_{j_1}\bar{\tau}_{j_1,k_1}\tau_{j_1,k_2}$ gives the (k_1,k_2) matrix element of $\tau^\dagger\tau$. Then, we define $C_{k_1}^\dagger = \sum_{m_1}\bar{\nu}_{k_1,m_1}A_{k_1}^\dagger N_{m_1}$ and perform diagonalization (Lemma 2):

$$\begin{aligned} &\sum_{k_1,k_2}\left(\sum_{j_1}\bar{\tau}_{j_1,k_1}\tau_{j_1,k_2}\right)\left(\sum_{m_1}\bar{\nu}_{k_1,m_1}A_{k_1}^\dagger N_{m_1}\right)\left(\sum_{m_2}\nu_{k_2,m_2}N_{m_2}A_{k_2}\right) \\ &= \sum_{k_1,k_2}\left(\tau^\dagger\tau\right)_{k_1,k_2}C_{k_1}^\dagger C_{k_2} \le \left\|\tau^\dagger\tau\right\|\sum_{k_1}C_{k_1}^\dagger C_{k_1} \\ &= \left\|\tau^\dagger\tau\right\|\sum_{k_1}\sum_{m_1,m_2}\bar{\nu}_{k_1,m_1}\nu_{k_1,m_2}A_{k_1}^\dagger N_{m_1}N_{m_2}A_{k_1}. \end{aligned} \tag{60}$$

Now that the indices k_1 and k_2 are contracted, we can apply the Hölder-type inequality for

expectation (Lemma 3). To this end, we let $D_{k_1}^\dagger = \sum_{m_1} \bar{\nu}_{k_1,m_1} A_{k_1}^\dagger N_{m_1}$ and compute

$$\begin{aligned}\left\| \sum_{k_1} \sum_{m_1,m_2} \bar{\nu}_{k_1,m_1} \nu_{k_1,m_2} A_{k_1}^\dagger N_{m_1} N_{m_2} A_{k_1} \right\|_\eta &= \left\| \sum_{k_1} A_{k_1}^\dagger D_{k_1}^\dagger D_{k_1} A_{k_1} \right\|_\eta \\ &\leq \left\| \sum_{k_1} A_{k_1}^\dagger A_{k_1} \right\|_\eta \max_{k_1} \left\| D_{k_1}^\dagger D_{k_1} \right\|_{\eta-1}.\end{aligned} \tag{61}$$

The first factor can be directly bounded as

$$\left\| \sum_{k_1} A_{k_1}^\dagger A_{k_1} \right\|_\eta = \|N\|_\eta = \eta. \tag{62}$$

For the second factor, we have

$$\begin{aligned}D_{k_1}^\dagger D_{k_1} &= \sum_{m_1,m_2} \bar{\nu}_{k_1,m_1} \nu_{k_1,m_2} N_{m_1} N_{m_2} = \sum_{m_1,m_2} \bar{\nu}_{k_1,m_1} \nu_{k_1,m_2} N_{m_1} N_{m_2} N_{m_1} \\ &\leq \|\nu\|_{\max}^2 \sum_{m_1,m_2} N_{m_1} N_{m_2} N_{m_1} = \|\nu\|_{\max}^2 N^2,\end{aligned} \tag{63}$$

which implies

$$\left\| D_{k_1}^\dagger D_{k_1} \right\|_{\eta-1} \leq \|\nu\|_{\max}^2 \eta^2. \tag{64}$$

Combining (59), (60), (61), (62), and (64) establishes the first statement.

For the second statement, we let $X = \sum_{j,k} \tau_{j,k} \nu_{k,k} A_j^\dagger A_k$ and compute

$$\begin{aligned}X^\dagger X &= \sum_{j_1,k_1,j_2,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_2,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{j_1} A_{j_2}^\dagger A_{k_2} \\ &= \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_2,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{k_2} - \sum_{j_1,k_1,j_2,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_2,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{j_2}^\dagger A_{j_1} A_{k_2}.\end{aligned} \tag{65}$$

Applying Lemma 1,

$$\begin{aligned}X^\dagger X &\leq \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_2,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{k_2} + \sum_{j_1,k_1,j_2,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_1,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{j_2}^\dagger A_{j_2} A_{k_2} \\ &= \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_2,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{k_2} + \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_1,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger N A_{k_2} \\ &= \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{k_1,k_1} \tau_{j_1,k_2} \nu_{k_2,k_2} A_{k_1}^\dagger A_{k_2} N \\ &= \sum_{k_1,k_2} \left(\sum_{j_1} \bar{\tau}_{j_1,k_1} \tau_{j_1,k_2} \right) \bar{\nu}_{k_1,k_1} \nu_{k_2,k_2} A_{k_1}^\dagger A_{k_2} N.\end{aligned} \tag{66}$$

Performing diagonalization using Lemma 2, we have

$$X^\dagger X \leq \|\tau\|^2 \sum_{k_1} \bar{\nu}_{k_1,k_1} \nu_{k_1,k_1} A_{k_1}^\dagger A_{k_1} N. \tag{67}$$

Note that we could directly bound the above operators as $\|\tau\|^2 \|\nu\|_{\max}^2 N^2$ and thereby completes the proof. But we choose to instead apply Lemma 3 so that the analysis can be directly extended to analyze multilayer nested commutators. We have

$$\begin{aligned}\left\|X^\dagger X\right\|_\eta &\le \left\|\|\tau\|^2 \sum_{k_1} \bar{\nu}_{k_1,k_1}\nu_{k_1,k_1} A_{k_1}^\dagger A_{k_1} N\right\|_\eta = \|\tau\|^2 \eta \left\|\sum_{k_1} \bar{\nu}_{k_1,k_1}\nu_{k_1,k_1} A_{k_1}^\dagger A_{k_1}\right\|_\eta \\ &\le \|\tau\|^2 \eta \left\|\sum_{k_1} A_{k_1}^\dagger A_{k_1}\right\|_\eta \max_{k_1} \|\bar{\nu}_{k_1,k_1}\nu_{k_1,k_1} I\|_{\eta-1} \le \|\tau\|^2 \|\nu\|_{\max}^2 \eta^2.\end{aligned} \tag{68}$$

The proof of the second statement is now completed. See Appendix A for the proof of the remaining statements. □

3.3 Multilayer nested commutators

We now analyze the error of simulating the correlated-electronic Hamiltonian (1) using a general pth-order formula $\mathscr{S}_p(t)$. We know from (39) that

$$\left\|\mathscr{S}_p(t) - e^{-itH}\right\|_\eta = \mathcal{O}\left(\max_\gamma \left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\|_\eta t^{p+1}\right), \tag{69}$$

where $H_0 = V = \sum_{l,m} \nu_{l,m} N_l N_m$ and $H_1 = T = \sum_{j,k} \tau_{j,k} A_j^\dagger A_k$. Our goal is to show that

$$\left\|\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]\right\|_\eta = \mathcal{O}\left(\left(\|\tau\| + \|\nu\|_{\max}\eta\right)^{p-1} \|\tau\| \|\nu\|_{\max} \eta^2\right) \tag{70}$$

for each multilayer nested commutator $\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]$.

To this end, we assume that $\left[H_{\gamma_p}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]$ is expressed as a fermionic operator of the form

$$\sum_{j,k,l} w_{j,k,l} \cdots A_{j_x}^\dagger \cdots A_{k_y} \cdots N_{l_z} \cdots \tag{71}$$

and analyze its commutator with either T or V. For the commutator with T, we have from Proposition 2

$$\left[A_j^\dagger A_k, A_{j_x}^\dagger\right] = \delta_{k,j_x} A_j^\dagger, \quad \left[A_j^\dagger A_k, A_{k_y}\right] = -\delta_{k_y,j} A_k, \quad \left[A_j^\dagger A_k, N_{l_z}\right] = \delta_{k,l_z} A_j^\dagger A_k - \delta_{j,l_z} A_j^\dagger A_k. \tag{72}$$

To develop some intuitions about these commutations, we introduce the notion of *fermionic chain*, which refers to a product of fermionic operators that has a creation operator on the left and an annihilation operator on the right. Then, the above commutations either extend an existing fermionic chain (in the case where commutator is taken with $A_{j_x}^\dagger$ or A_{k_y}), or create a new chain (in the case where commutator is taken with N_{l_z}). On the other hand, we can also apply Proposition 2 to compute the commutator with V:

$$\begin{aligned}\left[N_l N_m, A_{j_x}^\dagger\right] &= \delta_{m,j_x} N_l A_{j_x}^\dagger + \delta_{l,j_x} A_{j_x}^\dagger N_m = \delta_{m,j_x} A_{j_x}^\dagger N_l + \delta_{l,j_x} A_{j_x}^\dagger N_m + \delta_{m,j_x}\delta_{l,j_x} A_{j_x}^\dagger, \\ [N_l N_m, A_{k_x}] &= -\delta_{m,k_x} N_l A_{k_x} - \delta_{l,k_x} A_{k_x} N_m = -\delta_{m,k_x} N_l A_{k_x} - \delta_{l,k_x} N_m A_{k_x} - \delta_{m,k_x}\delta_{l,k_x} A_{k_x}.\end{aligned} \tag{73}$$

Unlike the commutator with T, these commutations do not extend an existing chain or create a new chain. Rather, their effect is to append V to an existing chain.

We now apply (72) and (73) iteratively to compute a general multilayer nested commutator $\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]$. We summarize the structure of the resulting operator in the following proposition.

Proposition 8 (Structure of multilayer nested commutators)**.** *Let* $H = T + V = \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} \nu_{l,m} N_l N_m$ *be a correlated-electronic Hamiltonian (1). Then, each* $\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]$ *where* $H_0 = V$ *and* $H_1 = T$ *is a linear combination of* fermionic chains*:*

$$X = \sum_{j,k} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot A_{j_q}^\dagger \prod_{x=1}^{q} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}. \tag{74}$$

Comment: add for some $q < p$*. Here, each* $B_{x,y}$ *either equals* $\sum_l \nu_{l,j_x} N_l$*,* $\sum_m \nu_{j_x,m} N_m$*,* $\nu_{j_x,j_x} I$*, or defines a* fermionic subchain*:*

$$B_{x,u} = \sum_{\widetilde{j},\widetilde{k}} \prod_{\widetilde{x}=1}^{\widetilde{q}} \tau_{\widetilde{j}_{\widetilde{x}},\widetilde{k}_{\widetilde{x}}} \prod_{\widetilde{x}=1}^{\widetilde{q}-1} \delta_{\widetilde{k}_{\widetilde{x}+1},\widetilde{j}_{\widetilde{x}}} \cdot A_{\widetilde{j}_{\widetilde{q}}}^\dagger \prod_{\widetilde{x}=1}^{\widetilde{q}} \left(\prod_{\widetilde{y}} \widetilde{B}_{\widetilde{x},\widetilde{y}} \prod_{\widetilde{z}} \widetilde{C}_{\widetilde{x},\widetilde{z}} \right) A_{\widetilde{k}_1}, \tag{75}$$

similar to the definition of fermionic chain, except one and only one $\widetilde{B}_{\widetilde{x},\widetilde{y}}$ *equals* $\nu_{\widetilde{j}_{\widetilde{x}},j_x} I$*,* $\nu_{j_x,\widetilde{j}_{\widetilde{x}}} I$*, or* $\widetilde{C}_{\widetilde{x},\widetilde{z}}$ *equals* $\nu_{\widetilde{k}_{\widetilde{x}},j_x} I$*,* $\nu_{j_x,\widetilde{k}_{\widetilde{x}}} I$*. Operators* $C_{x,z}$ *are defined similarly. Furthermore, all operators* $B_{x,y}$*,* $C_{x,z}$ *and hence the entire chain* X *are number-preserving. It holds*

$$\|X\|_\eta \leq \|\tau\|^q \, \eta \prod_{x=1}^{q} \left(\prod_y \max_{j_x} \|B_{x,y}\|_{\eta-1} \prod_z \max_{k_x} \|C_{x,z}\|_{\eta-1} \right). \tag{76}$$

Proof. We will analyze the structure of multilayer nested commutators by induction. In the base case where $p = 1$, we have $[H_{\gamma_2}, H_{\gamma_1}] = [T, V]$. This commutator has the expansion (56), where each term is indeed a fermionic chain and number-preserving. This completes the proof of the base case.

Assuming the claim holds for the nested commutator $\left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]$, we now consider the structure of $\left[H_{\gamma_{p+2}}, \cdots [H_{\gamma_2}, H_{\gamma_1}]\right]$. By induction, this nested commutator is a linear combination of either

$$\begin{aligned}
&\sum_{j,k} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot \left[T, A_{j_q}^\dagger\right] \prod_{x=1}^{q} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}, \\
&\sum_{j,k} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot A_{j_q}^\dagger \prod_{x=1}^{q} \left(\prod_y [T, B_{x,y}] \prod_z C_{x,z} \right) A_{k_1}, \\
&\sum_{j,k} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot A_{j_q}^\dagger \prod_{x=1}^{q} \left(\prod_y B_{x,y} \prod_z [T, C_{x,z}] \right) A_{k_1}, \\
&\sum_{j,k} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot A_{j_q}^\dagger \prod_{x=1}^{q} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) [T, A_{k_1}],
\end{aligned} \tag{77}$$

or

$$\begin{aligned}
&\sum_{j,k}\prod_{x=1}^{q}\tau_{j_x,k_x}\prod_{x=1}^{q-1}\delta_{k_{x+1},j_x}\cdot\left[V,A_{j_q}^\dagger\right]\prod_{x=1}^{q}\left(\prod_y B_{x,y}\prod_z C_{x,z}\right)A_{k_1},\\
&\sum_{j,k}\prod_{x=1}^{q}\tau_{j_x,k_x}\prod_{x=1}^{q-1}\delta_{k_{x+1},j_x}\cdot A_{j_q}^\dagger\prod_{x=1}^{q}\left(\prod_y [V,B_{x,y}]\prod_z C_{x,z}\right)A_{k_1},\\
&\sum_{j,k}\prod_{x=1}^{q}\tau_{j_x,k_x}\prod_{x=1}^{q-1}\delta_{k_{x+1},j_x}\cdot A_{j_q}^\dagger\prod_{x=1}^{q}\left(\prod_y B_{x,y}\prod_z [V,C_{x,z}]\right)A_{k_1},\\
&\sum_{j,k}\prod_{x=1}^{q}\tau_{j_x,k_x}\prod_{x=1}^{q-1}\delta_{k_{x+1},j_x}\cdot A_{j_q}^\dagger\prod_{x=1}^{q}\left(\prod_y B_{x,y}\prod_z C_{x,z}\right)[V,A_{k_1}].
\end{aligned}\tag{78}$$

In each case, we see from (72) and (73) that the result is again a fermionic chain. Specifically, commutators $\left[T,A_{j_q}^\dagger\right]$ and $[T,A_{k_1}]$ increase the "length" of the current fermionic chain from q to $q+1$; commutators $[T,B_{x,y}]$ and $[T,C_{x,z}]$ either create a fermionic subchain or give zero operator, or they can be computed recursively when $B_{x,y}$ and $C_{x,z}$ are fermionic subchains; commutators $\left[V,A_{j_q}^\dagger\right]$ and $[V,A_{k_1}]$ do not increase the length q of the current fermionic chain, but they increase the number of $B_{x,y}$ and $C_{x,z}$ by one; commutators $[V,B_{x,y}]$ and $[V,C_{x,z}]$ either give zero operator, or they can be computed recursively if $B_{x,y}$ and $C_{x,z}$ are fermionic subchains. The claim about the number preservation can be verified directly. This completes the inductive step.

It remains to bound $\|X\|_\eta$ for an arbitrary fermionic chain X in (74). This can be done in a similar way as in the proof of Proposition 7, using Lemma 1, Lemma 2, and Lemma 3. We leave the detailed justification to Appendix B. □

We now apply Proposition 8 to analyze each nested commutator $\left\|\left[H_{\gamma_{p+1}},\cdots[H_{\gamma_2},H_{\gamma_1}]\right]\right\|_\eta$. We already have the correct spectral-norm $\|\tau\|$ scaling from Proposition 8. To proceed, we need to further bound each $\|B_{x,y}\|_{\eta-1}$ and $\|C_{x,z}\|_{\eta-1}$ separately. We have

$$\begin{aligned}
\left\|\sum_l \nu_{l,j_x}N_l\right\|_{\eta-1} &= \sqrt{\left\|\sum_{l_1,l_2}\bar{\nu}_{l_1,j_{x_1}}\nu_{l_2,j_{x_2}}N_{l_1}N_{l_2}\right\|_{\eta-1}} \le \|\nu\|_{\max}\eta,\\
\left\|\sum_m \nu_{j_x,m}N_m\right\|_{\eta-1} &= \sqrt{\left\|\sum_{m_1,m_2}\bar{\nu}_{j_{x_1},m_1}\nu_{j_{x_2},m_2}N_{m_1}N_{m_2}\right\|_{\eta-1}} \le \|\nu\|_{\max}\eta,\\
\|\nu_{j_x,j_x}I\|_{\eta-1} &\le \|\nu\|_{\max}
\end{aligned}\tag{79}$$

for $B_{x,y}$ and similar estimates hold for $C_{x,z}$. In the case where $B_{x,y}$ or $C_{x,z}$ creates a fermionic subchain, we can bound the subchain recursively using Proposition 8. In particular, we will introduce a factor of $\|\nu\|_{\max}\eta$ each time a subchain is created. This completes the proof of Eq. (2) of our main result Theorem 1.

4 Bounding expectation of fermionic operators by path counting

We now present an alternative strategy for bounding the Trotterization error of Fermionic simulation. We will illustrate this strategy by proving a bound for the d-sparse Fermionic Hamiltonian,

$$H = T + V = \sum_{j,k}\tau_{j,k}A_j^\dagger A_k + \sum_{l,m}\nu_{l,m}N_lN_m,\tag{80}$$

where each column and row of τ, ν has at most d non-zero elements. Recall from (39) that

$$\left\| \mathscr{S}_p(t) - e^{-itH} \right\|_\eta = \mathcal{O}\left(\max_\gamma \left\| \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \right] \right\|_\eta t^{p+1} \right), \tag{81}$$

where $H_0 = V = \sum_{l,m} \nu_{l,m} N_l N_m$ and $H_1 = T = \sum_{j,k} \tau_{j,k} A_j^\dagger A_k$. Hence to estimate the simulation error of electronic Hamiltonian (1) using a general pth-order formula $\mathscr{S}_p(t)$, we need to bound the fermionic seminorm

$$\left\| \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \right] \right\|_\eta, \tag{82}$$

where $\gamma_p \in \{0, 1\}$, $H_0 = V$ and $H_1 = T$.

4.1 Path counting bound

Using Proposition 6, we can bound the transition amplitude between any two states by at most two times the expectation value.

$$\left\| \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \right] \right\|_\eta \leq 2 \max_{|\psi_\eta\rangle} |\langle \psi_\eta | \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \right] |\psi_\eta\rangle|. \tag{83}$$

We now aim to bound the term

$$X = |\langle \psi_\eta | \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \right] |\psi_\eta\rangle| \tag{84}$$

for any $|\psi_\eta\rangle$. First, we write out the superposition of $|\psi_\eta\rangle$ in the second-quantization picture:

$$|\psi_\eta\rangle = \sum_{C \in \{0,1\}^n, |C| = \eta} \alpha_C |C\rangle, \tag{85}$$

where C is a configuration of the occupation number with η electrons, and we represent the number of ones in C as $|C|$. We also define the following notations

$$\mu^0 = \nu, \quad \mu^1 = \tau, \quad h^0_{jk} = N_j N_k, \quad h^1_{jk} = A_j^\dagger A_k. \tag{86}$$

We can expand everything to give

$$X = \left| \sum_{j_{p+1}, k_{p+1}} \cdots \sum_{j_1, k_1} \sum_{C_1} \sum_{C_2} \alpha^*_{C_1} \alpha_{C_2} \mu^{\gamma_{p+1}}_{j_{p+1} k_{p+1}} \cdots \mu^{\gamma_1}_{j_1 k_1} \langle C_1 | \left[h^{\gamma_{p+1}}_{j_{p+1} k_{p+1}}, \ldots \left[h^{\gamma_2}_{j_2 k_2}, h^{\gamma_1}_{j_1 k_1} \right] \right] |C_2\rangle \right| \tag{87}$$

$$\leq \|\tau\|_{\max}^{\sum_{q=1}^{p+1} \gamma_q} \|\nu\|_{\max}^{\sum_{q=1}^{p+1} (1-\gamma_q)} \sum_{C_1} \sum_{C_2} |\alpha_{C_1}| |\alpha_{C_2}| \times \tag{88}$$

$$\sum_{\langle j_{p+1}, k_{p+1} \rangle} \cdots \sum_{\langle j_1, k_1 \rangle} \left| \langle C_1 | \left[h^{\gamma_{p+1}}_{j_{p+1} k_{p+1}}, \ldots \left[h^{\gamma_2}_{j_2 k_2}, h^{\gamma_1}_{j_1 k_1} \right] \right] |C_2\rangle \right|, \tag{89}$$

where C_1, C_2 are configurations of the occupation number with η electrons, and $\langle j, k \rangle$ only sum over indices such that the corresponding $\mu_{j,k} \neq 0$ (could be τ or ν depending on γ).

Using the commutation relations in Equation (72) and (73), we know that the nested commutator $\left[h^{\gamma_{p+1}}_{j_{p+1} k_{p+1}}, \ldots \left[h^{\gamma_2}_{j_2 k_2}, h^{\gamma_1}_{j_1 k_1} \right] \right]$ can be written as a sum of

$$(-1)^a \ldots A_j^\dagger \ldots A_k \ldots N_l \ldots, \tag{90}$$

for some $a \in \{0,1\}$ and a sequence of Fermionic operators. We call each term a Fermionic path P and write $P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1k_1}\right)$ if P is one of the terms in the expansion of the nested commutator. If the nested commutator evaluates to zero, then we consider the set

$$\left\{P \text{ such that } P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1k_1}\right)\right\} \tag{91}$$

to be an empty set. The detailed expansion of the nested commutator is presented in Proposition 8. This allows us to make a further expansion to yield

$$X \leq c_{\tau\nu} \sum_{C_1} \sum_{C_2} |\alpha_{C_1}| \, |\alpha_{C_2}| \sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1k_1}\right)} |\langle C_1|P|C_2\rangle| \,, \tag{92}$$

where $c_{\tau\nu} = \|\tau\|_{\max}^{\sum_{q=1}^{p+1}\gamma_q} \|\nu\|_{\max}^{\sum_{q=1}^{p+1}(1-\gamma_q)}$. We use the following proposition to characterize $|\langle C_1|P|C_2\rangle|$.

Proposition 9. *For any computational basis state $|C\rangle$, where C represents the second-quantized configuration for whether an electron occupies each of the n sites, and Fermionic path*

$$P = (-1)^a \ldots A_j^\dagger \ldots A_k \ldots N_l \ldots, \tag{93}$$

for some $a \in \{0,1\}$ and a sequence of Fermionic operators. We have $P|C\rangle$ is a computational basis state with some phase ± 1 or $P|C\rangle = 0$.

Proof. The proof follows from a simple induction. For the base case, we have $P = (-1)^a$ without any Fermionic operator, so $P|C\rangle$ is a computational basis state with some phase ± 1. Now we consider the three cases: $P = N_l P'$, $P = A_k P'$, or $P = A_j^\dagger P'$. By induction, we have $P'|C\rangle$ is a computational basis state $|C'\rangle$ with some phase ± 1 or $P'|C\rangle = 0$. The latter is trivial. For the former case, we go through the following three cases.

- If N_l is applied on $|C'\rangle$, we check if site-l has an electron in configuration C'. If site-l has an electron, then $N_l|C'\rangle = |C'\rangle$. If site-$l$ does not have an electron, then $N_l|C'\rangle = 0$.
- If A_k is applied on $|C'\rangle$, we check if site-k has an electron in configuration C'. If site-k has an electron, then $A_k|C'\rangle$ will remove the site-k electron and add some phase according to the rule given in Equation (27). If site-k does not have an electron, then $A_k|C'\rangle = 0$.
- If $A_j^\dagger$ is applied on $|C'\rangle$, we check if site-j has an electron in configuration C'. If site-j does not have an electron, then $A_j^\dagger|C'\rangle$ will create an electron at site-j and add some phase according to the rule given in Equation (26). If site-j does have an electron, then $A_j^\dagger|C'\rangle = 0$.

Therefore, $P|C\rangle$ will be a computational basis state with some phases ± 1 or otherwise $P|C\rangle = 0$. □

Corollary 1. *$|\langle C_1|P|C_2\rangle|$ is either 0 or 1.*

Corollary 2. *For any set S of configurations, we have $\sum_{C_1 \in S} |\langle C_1|P|C_2\rangle| \leq \|P|C_2\rangle\|$.*

Next, we define a graph $G = (V, E)$ where the vertices V are the second-quantized configurations with η electrons, and the weighted adjacency matrix for the edges E is defined as

$$w_{C_1,C_2} = \sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1k_1}\right)} |\langle C_1|P|C_2\rangle| \,. \tag{94}$$

The weight w_{C_1,C_2} counts the number of Fermionic paths that can take $|C_2\rangle$ to $|C_1\rangle$. Note that this graph may contain self-loop (equivalent to $w_{C_1,C_1} > 0$) as there are Fermionic path that leaves $|C_1\rangle$ unchanged or simply add a phase of -1. We now define the degree of C_2 as

$$\deg(C_2) = \sum_{C_1} \frac{w_{C_1,C_2} + w_{C_2,C_1}}{2} \tag{95}$$

$$= \sum_{C_1} \sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} \frac{|\langle C_1|P|C_2\rangle| + |\langle C_2|P|C_1\rangle|}{2} \tag{96}$$

$$= \sum_{C_1} \sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} \frac{|\langle C_1|P|C_2\rangle| + \left|\langle C_1|P^\dagger|C_2\rangle\right|}{2} \tag{97}$$

$$\leq \sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} \frac{\|P|C_2\rangle\| + \left\|P^\dagger|C_2\rangle\right\|}{2}, \tag{98}$$

which is equivalent to counting how many Fermionic paths exist that do not evaluate to zero on the initial state $|C_2\rangle$. The last inequality follows from Corollary 2.

Lemma 4. *For any symmetric, nonnegative matrix $W \in \mathbb{R}^{k\times k}$, we have*

$$\sum_{ij} W_{ij} v_i v_j \leq \max_i \sum_j W_{ij}, \tag{99}$$

for all $v \in \mathbb{R}^k$ with $v^T v = 1$.

Proof. Let us consider the eigenvector u_1 corresponding to the largest eigenvalue λ_1 of W. Because

$$u_1 = \operatorname*{argmax}_{u\in\mathbb{R}^k, u^T u=1} u^T W u, \tag{100}$$

we have $v^T W v \leq u_1^T W u_1$. Let us consider $i^* = \operatorname{argmax}_j (u_1)_j$. We can always multiply u_1 by -1 and still remains as the eigenvector corresponding to eigenvalue λ_1. Hence, we can always assume $(u_1)_{i^*} > 0$. Therefore, we have

$$u_1^T W u_1 = \lambda_1 = \frac{(Wu_1)_{i^*}}{(u_1)_{i^*}} \leq \frac{\sum_j W_{i^*j}(u_1)_j}{(u_1)_{i^*}} \leq \sum_j W_{i^*j} \leq \max_i \sum_j W_{ij}. \tag{101}$$

This concludes the proof. □

Using Lemma 4, we can now simplify the upper bound of X to obtain

$$X \leq c_{\tau\nu} \sum_{C_1}\sum_{C_2} |\alpha_{C_1}|\,|\alpha_{C_2}| \sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} |\langle C_1|P|C_2\rangle| \tag{102}$$

$$= c_{\tau\nu} \sum_{C_1}\sum_{C_2} w_{C_1,C_2}\,|\alpha_{C_1}|\,|\alpha_{C_2}| \tag{103}$$

$$= c_{\tau\nu} \sum_{C_1}\sum_{C_2} \frac{w_{C_1,C_2} + w_{C_2,C_1}}{2}\,|\alpha_{C_1}|\,|\alpha_{C_2}| \tag{104}$$

$$\leq c_{\tau\nu} \max_{C_2} \sum_{C_1} \frac{w_{C_1,C_2} + w_{C_2,C_1}}{2} \tag{105}$$

$$\leq c_{\tau\nu} \max_C \deg(C). \tag{106}$$

Hence the error of Fermionic simulation can be upper bounded by the maximum degree of the graph G. Finally, we arrive at the following proposition by combining with Equation (98).

Proposition 10. *We have the following path counting bound,*

$$\left\| \left[H_{\gamma_{p+1}}, \cdots [H_{\gamma_2}, H_{\gamma_1}] \right] \right\|_\eta \leq c_{\tau\nu} \max_{C_\eta} \deg\left(C_\eta\right), \tag{107}$$

where the degree is defined as

$$\deg\left(C_\eta\right) = \sum_{\langle j_{p+1}, k_{p+1}\rangle} \cdots \sum_{\langle j_1, k_1\rangle} \sum_{P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)} \frac{1}{2}(\|P|C_\eta\rangle\| + \|P^\dagger|C_\eta\rangle\|). \tag{108}$$

Recall the following definitions, $c_{\tau\nu} = \|\tau\|_{\max}^{\sum_{q=1}^{p+1}\gamma_q} \|\nu\|_{\max}^{\sum_{q=1}^{p+1}(1-\gamma_q)}$. *$C_\eta$ is any one of the $\binom{n}{\eta}$ second-quantized occupation number configurations with η electrons. For all $q = 1, \ldots, p+1$, $\langle j_q, k_q\rangle$ sum over indices j_q, k_q such that*

$$\begin{cases} \tau_{j_q,k_q} \neq 0 & \text{if } \gamma_q = 1, \\ \nu_{j_q,k_q} \neq 0 & \text{if } \gamma_q = 0. \end{cases} \tag{109}$$

We use the following short-hand notation for local Fermionic terms: $h^0_{jk} = N_j N_k, h^1_{jk} = A^\dagger_j A_k$. *Finally, we denote* $P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)$ *if*

$$P = (-1)^a \ldots A^\dagger_j \ldots A_k \ldots N_l \ldots \tag{110}$$

is one of the terms in the expansion of the nested commutator $\left[h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots \left[h^{\gamma_2}_{j_2k_2}, h^{\gamma_1}_{j_1k_1}\right]\right]$. *We refer to P as a Fermionic path.*

4.2 Counting the number of Fermionic paths in d-sparse Hamiltonians

As an illustrative example, let us consider an upper bound of $\max_C \deg(C)$ for d-sparse Hamiltonians. We will use the commutation relation in Equation (72) and (73), restated below

$$\left[A^\dagger_j A_k, A^\dagger_{j_x}\right] = \delta_{k,j_x} A^\dagger_j, \quad \left[A^\dagger_j A_k, A_{k_y}\right] = -\delta_{k_y,j} A_k, \quad \left[A^\dagger_j A_k, N_{l_z}\right] = \delta_{k,l_z} A^\dagger_j A_k - \delta_{j,l_z} A^\dagger_j A_k. \tag{111}$$

$$\left[N_l N_m, A^\dagger_{j_x}\right] = \delta_{m,j_x} N_l A^\dagger_{j_x} + \delta_{l,j_x} A^\dagger_{j_x} N_m, \quad [N_l N_m, A_{k_x}] = -\delta_{m,k_x} N_l A_{k_x} - \delta_{l,k_x} A_{k_x} N_m. \tag{112}$$

We start with an intuitive argument. For every $q = 2, \ldots, p+1$, we have

$$\left[h^{\gamma_q}_{j_q k_q}, \ldots \left[h^{\gamma_2}_{j_2k_2}, h^{\gamma_1}_{j_1k_1}\right]\right] = \sum_{P \rhd \left(h^{\gamma_{q-1}}_{j_{q-1}k_{q-1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)} \left[h^{\gamma_q}_{j_q k_q}, P\right]. \tag{113}$$

Because P only contains Fermionic operator acting on sites $j_1, k_1, \ldots, j_{q-1}, k_{q-1}$, from the commutation relation, we see that at least one of j_q, k_q must be equal to one of the indices $j_1, k_1, \ldots, j_{q-1}, k_{q-1}$. Furthermore, for every j_q, there are at most d k_q's that have non-zero coefficient in τ_{j_q,k_q} (for $\gamma_q = 1$) or ν_{j_q,k_q} (for $\gamma_q = 0$). Hence, we have the following bound

$$\sum_{\langle j_{p+1}, k_{p+1}\rangle} \cdots \sum_{\langle j_1, k_1\rangle} \sum_{P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)} 1 \leq \mathcal{O}(nd^{p+1}). \tag{114}$$

The n factor comes from the fact that only one index can freely choose between $1, \ldots, n$. And for any two indices j_q, k_q, one of them has to be the same as the previous indices, while the other one can only choose from the d indices as required when summing over $\langle j_q, k_q \rangle$. Hence we have the d^p contribution for the rest of the indices. Furthermore, we will later show that the rightmost Fermionic operator in P will never be a creation operator $A^\dagger$. Hence, when we include the term $\|P|C_\eta\rangle\|$, the Fermionic path P must begin with a Fermionic operator that acts on one of the η sites with an electron in the configuration C_η. Therefore, we have

$$\sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} \|P|C_\eta\rangle\| \leq \mathcal{O}(\eta d^{p+1}). \tag{115}$$

Similarly, we have the following bound

$$\sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P^\dagger \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} \|P^\dagger|C_\eta\rangle\| \leq \mathcal{O}(\eta d^{p+1}). \tag{116}$$

Combining with Trotter simulation bound (39) and the path counting bound (107), we have obtained the result

$$\left\|\mathscr{S}_p(t) - e^{-itH}\right\|_\eta = \mathcal{O}\left(c_{\tau\nu} d^{p+1} \eta t^{p+1}\right). \tag{117}$$

Finally, because $1 \leq \sum_{q=1}^{p+1} \gamma_q \leq p$, we also have

$$c_{\tau\nu} = \|\tau\|_{\max}^{\sum_{q=1}^{p+1}\gamma_q} \|\nu\|_{\max}^{\sum_{q=1}^{p+1}(1-\gamma_q)} \leq \left(\|\tau\|_{\max} + \|\nu\|_{\max}\right)^p \|\tau\|_{\max} \|\nu\|_{\max}. \tag{118}$$

This gives an overview for proving the scaling stated in Equation (3) of Theorem 1. We now present a rigorous proof by induction.

Proposition 11. *When each column and row of τ, ν has at most d non-zero elements, for any constant integer $p \geq 1$, we have*

$$\sum_{\langle j_{p+1},k_{p+1}\rangle} \cdots \sum_{\langle j_1,k_1\rangle} \sum_{P \triangleright \left(h^{\gamma_{p+1}}_{j_{p+1}k_{p+1}},\ldots,h^{\gamma_1}_{j_1k_1}\right)} \|P|C_\eta\rangle\| \leq \mathcal{O}(\eta d^{p+1}). \tag{119}$$

Proof. We will prove the following induction hypothesis on $q = 2, \ldots, p+1$.

- All Fermionic path P will start with either N or A, but will never start with $A^\dagger$ (we refer to the rightmost operator as the starting point).
- All Fermionic path P will have at most $q+1$ Fermionic operators.
- The total number of Fermionic path P is at most $(\Pi_{z=2}^q 2z) d^q n$.
- The number of Fermionic path P that start with a Fermionic operator acting on site i is at most $(\Pi_{z=2}^q 2z) d^q$.

The base case $q = 2$ can be easily verified by noting that we only need to consider $[T, V]$ or $[V, T]$. Using the commutation relations given in Equation (111) and (112), we can see that both cases result in at most $4d^2 n$ different Fermionic paths and all start with either N or A. For every site i, there are at most $4d^2$ Fermionic paths starting with site i. Furthermore, every Fermionic path consists of 3 Fermionic operators. These results established the induction hypothesis (all four bullet points) for the base case of $q = 2$.

For every $q > 2$, we can use the induction hypothesis for $q-1$ to prove the desired result. If $\gamma_q = 1$, then we will be adding another commutator with $T = \sum_{j_q,k_q} \tau_{j_q,k_q} A^\dagger_{j_q} A_{k_q}$. We can see that all Fermionic path $P \rhd \left(h^{\gamma_q}_{j_q k_q}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)$ comes from the expansion of

$$[A^\dagger_{j_q} A_{k_q}, P'], \forall \langle j_q, k_q \rangle, \forall P' \rhd \left(h^{\gamma_{q-1}}_{j_{q-1} k_{q-1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right). \tag{120}$$

Note that the commutator of $[X, Y_1 \ldots Y_K] = \sum_{k=1}^K Y_1 \ldots Y_{k-1} [X, Y_k] Y_{k+1} \ldots Y_K$. Using this rule, we can easily show that the induction hypothesis holds for q. First of all, if all the Fermionic path $P' \rhd \left(h^{\gamma_{q-1}}_{j_{q-1} k_{q-1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)$ start with either N or A, then all the Fermionic path $P \rhd \left(h^{\gamma_q}_{j_q k_q}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)$ will start with either N or A. This is because of the commutation relations: $[A^\dagger_j A_k, A_{k_q}] = -\delta_{k_y,j} A_k$ and $\left[A^\dagger_j A_k, N_{l_z}\right] = \delta_{k,l_z} A^\dagger_j A_k - \delta_{j,l_z} A^\dagger_j A_k$. Furthermore, because P' have at most $(q-1)+1 = q$ Fermionic operators, the expansion of $[A^\dagger_{j_q} A_{k_q}, P']$ will have at most $q+1$ Fermionic operators. Using the fact that the commutation relation always create a delta function between j_q or k_q with an existing index in P' and the fact that every j_q connects to at most d k_q's and every k_q connects to at most d j_q's, there can only be $2dq$ different Fermionic paths for a fixed P'. Hence, there will be at most a total of

$$2dq \times \left(\prod_{z=2}^{q-1} 2z\right) d^{q-1} n = \left(\prod_{z=2}^{q} 2z\right) d^q n \tag{121}$$

Fermionic paths for q. We now obtain an upper bound for the Fermionic paths starting from some site i. If we take the commutator of $A^\dagger_{j_q} A_{k_q}$ with a Fermionic operator that is not the starting operator in P, then the starting operator is not affected. Because of the sparsity constraint and the delta function created by the commutation relation in Equation (111), we have created at most $2d(q-1)\times$ more Fermionic paths starting with site i. Now if we take the commutator of $A^\dagger_{j_q} A_{k_q}$ with the starting operator A_{k_y} (for some index k_y) in the Fermionic path P', then the starting operator becomes A_{k_q} and we have an additional delta function δ_{k_y,j_q}. In this case, k_q can start from any site, but there will be at most d choices of j_q, hence d choices of k_y. This means we have created at most $2d\times$ more Fermionic paths starting with each site. Together, we have created at most $2dq\times$ more Fermionic paths starting with each site. This leads to an upper bound of

$$2dq \times \left(\prod_{z=2}^{q-1} 2z\right) d^{q-1} = \left(\prod_{z=2}^{q} 2z\right) d^q \tag{122}$$

Fermionic paths starting with each site. The inductive step for $\gamma_q = 0$ follows from a similar argument. Hence, the induction hypothesis holds for q.

Performing the induction over q from 2 to $p+1$ shows that the number of Fermionic paths starting with site i is at most

$$\left(\prod_{z=2}^{p+1} 2z\right) d^{p+1} = \mathcal{O}(d^{p+1}). \tag{123}$$

Because P start with either A or N, $\|P|C_\eta\rangle\|$ would be nonzero if the starting Fermionic operator acts on one of the η sites with an occupying electron in the configuration C_η. Hence there are at most $\eta\mathcal{O}(d^{p+1})$ Fermionic paths with non-zero $\|P|C_\eta\rangle\|$. Finally, recall from Proposition 9 that $\|P|C_\eta\rangle\|$ is either 0 or 1. Hence

$$\sum_{\langle j_{p+1}, k_{p+1}\rangle} \cdots \sum_{\langle j_1, k_1\rangle} \sum_{P \rhd \left(h^{\gamma_{p+1}}_{j_{p+1} k_{p+1}}, \ldots, h^{\gamma_1}_{j_1 k_1}\right)} \|P|C_\eta\rangle\| \leq \mathcal{O}(\eta d^{p+1}), \tag{124}$$

which is the desired bound on the number of Fermionic paths for d-sparse Hamiltonians. □

5 Tightness

We have already established in Theorem 1 multiple bounds on the fermionic seminorm of Trotter error. However, a common issue with the Trotterization algorithm is that its error estimate can be very loose for simulating specific systems. Here, we prove Theorem 2 that demonstrates the tightness of our analysis.

Specifically, we construct concrete examples of correlated-electronic Hamiltonian $H = T + V$ and lower-bound the fermionic seminorm of nested commutators: $\|[T, \dots [T, V]]\|_\eta$ in Section 5.1 and $\|[V, \dots [V, T]]\|_\eta$ in Section 5.2. We show that the results almost match the upper bounds in Theorem 1. Since Trotter error depends on these nested commutators, this shows that our result is nearly tight modulo an application of the triangle inequality.

5.1 Lower-bounding $\|[T, \dots [T, V]]\|_\eta$

We construct the electronic Hamiltonian $H = T + V$, where

$$T = \sum_{j,k=0}^{n-1} A_j^\dagger A_k, \qquad V = \sum_{u,v=0}^{\frac{n}{2}-1} N_u N_v. \tag{125}$$

Comparing with the definition of correlated-electronic model (1), we see that the coefficient matrix τ is an all-ones matrix with spectral norm $\|\tau\| = n$, whereas ν contains an all-ones submatrix on the top left corner with max-norm $\|\nu\|_{\max} = 1$. Our goal is to lower-bound the fermionic seminorm $\|[T, \dots [T, V]]\|_\eta$.

Due to the complicated commutation relations between T and V, a direct computation of $[T, \dots [T, V]]$ seems technically challenging. Instead, we perform a change of basis by applying the fermionic Fourier transform

$$\mathrm{FFFT}^\dagger \cdot A_j^\dagger \cdot \mathrm{FFFT} = \frac{1}{\sqrt{n}} \sum_l e^{-\frac{2\pi i j l}{n}} A_l^\dagger, \qquad \mathrm{FFFT}^\dagger \cdot A_k \cdot \mathrm{FFFT} = \frac{1}{\sqrt{n}} \sum_m e^{\frac{2\pi i k m}{n}} A_m. \tag{126}$$

This gives

$$\begin{aligned} \widetilde{T} &= \mathrm{FFFT}^\dagger \cdot T \cdot \mathrm{FFFT} = n N_0, \\ \widetilde{V} &= \mathrm{FFFT}^\dagger \cdot V \cdot \mathrm{FFFT} = \frac{1}{n^2} \sum_{j,k,l,m} \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u (k-j)}{n}} \right) \left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v (m-l)}{n}} \right) A_j^\dagger A_k A_l^\dagger A_m. \end{aligned} \tag{127}$$

We also define the η-electron states

$$|\widetilde{\psi}_\eta\rangle = \frac{|010\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle + |100\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle}{\sqrt{2}}, \qquad |\widetilde{\phi}_\eta\rangle = \frac{|010\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle + i|100\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle}{\sqrt{2}}. \tag{128}$$

The following proposition shows that the above choice of operators and states almost saturates the fermionic seminorm of nested commutators.

Proposition 12. *Define $\widetilde{T}$, $\widetilde{V}$ as in (127) and $|\widetilde{\psi}_\eta\rangle$, $|\widetilde{\phi}_\eta\rangle$ as in (128). Then,*

$$\left.\begin{array}{ll} |\langle\widetilde{\psi}_\eta|\overbrace{[\widetilde{T},\ldots[\widetilde{T}}^{p},\widetilde{V}]]|\widetilde{\psi}_\eta\rangle|, & p \text{ odd} \\ |\langle\widetilde{\phi}_\eta|\underbrace{[\widetilde{T},\ldots[\widetilde{T}}_{p},\widetilde{V}]]|\widetilde{\phi}_\eta\rangle|, & p \text{ even} \end{array}\right\} = \frac{n^p\eta}{\pi} + \mathcal{O}\left(n^p + n^{p-1}\eta\right). \tag{129}$$

A proof of this proposition is given in Appendix C. By rescaling the Hamiltonian constructed in (127), we can demonstrate the tightness of our bound as follows. For any $s, w > 0$, we define the rescaled Hamiltonian

$$T = \frac{s}{n}\sum_{j,k=0}^{n-1} A_j^\dagger A_k, \qquad V = w\sum_{u,v=0}^{\frac{n}{2}-1} N_u N_v. \tag{130}$$

Comparing with the definition of the correlated-electronic model (1), we see that $\|\tau\| = s$ and $\|\nu\|_{\max} = w$. The above proposition then shows that

$$\left\|\underbrace{[T,\ldots[T}_{p},V]]\right\|_\eta = \left\|\underbrace{[\widetilde{T},\ldots[\widetilde{T}}_{p},\widetilde{V}]]\right\|_\eta = \Omega\left(s^p w\eta\right), \tag{131}$$

where we have used the unitary invariance of the fermionic seminorm in the first equality. This establishes the claimed tightness result (5) of Theorem 2.

Note that a similar example can be constructed to demonstrate the tightness of our bound for simulating sparse electronic Hamiltonians. Specifically, for $t, w > 0$ and positive integer $1 \leq d \leq \eta$, we define

$$T = t\sum_{j,k=0}^{d-1} A_j^\dagger A_k, \qquad V = w\sum_{u,v=0}^{\frac{d}{2}-1} N_u N_v. \tag{132}$$

Comparing with the definition of the correlated-electronic model (1), we see that $\|\tau\|_{\max} = t$ and $\|\nu\|_{\max} = w$. We also perform a fermionic Fourier transform to define $\widetilde{T}$ and $\widetilde{V}$, but only to the first d spin orbitals

$$\begin{aligned} \mathrm{FFFT}^\dagger \cdot A_j^\dagger \cdot \mathrm{FFFT} &= \begin{cases} \frac{1}{\sqrt{d}}\sum_{l=0}^{d-1} e^{-\frac{2\pi i j l}{n}} A_l^\dagger, & 0 \leq j \leq d-1, \\ A_j^\dagger, & j \geq d. \end{cases} \\ \mathrm{FFFT}^\dagger \cdot A_k \cdot \mathrm{FFFT} &= \begin{cases} \frac{1}{\sqrt{d}}\sum_{m=0}^{d-1} e^{\frac{2\pi i k m}{n}} A_m, & 0 \leq k \leq d-1, \\ A_k, & k \geq d. \end{cases} \end{aligned} \tag{133}$$

The above proposition then shows that

$$\left\|\underbrace{[T,\ldots[T}_{p},V]]\right\|_\eta = \left\|\underbrace{[\widetilde{T},\ldots[\widetilde{T}}_{p},\widetilde{V}]]\right\|_\eta = \Omega\left((td)^p w\eta\right). \tag{134}$$

This proves the claimed tightness result (6) of Theorem 2.

5.2 Lower-bounding $\|[V,\ldots[V,T]]\|_\eta$

Recall from the previous section that we will consider the electronic Hamiltonian (125). Comparing to the definition of correlated-electronic model (1), we see that the coefficient matrix τ has spectral

norm $\|\tau\| = n$, whereas coefficient matrix ν has max-norm $\|\nu\|_{\max} = 1$. Our goal is to lower-bound the fermionic seminorm $\|[V, \dots [V, T]]\|_\eta$. To this end, we define the η-electron states

$$\begin{aligned} |\psi_\eta\rangle &= \frac{|0\overbrace{1\cdots1}^{\frac{n}{2}}\underbrace{0\cdots0}_{\eta-1}10\cdots0\rangle + i|1\overbrace{1\cdots1}^{\frac{n}{2}}\underbrace{0\cdots0}_{\eta-1}00\cdots0\rangle}{\sqrt{2}}, \\ |\phi_\eta\rangle &= \frac{|0\overbrace{1\cdots1}^{\frac{n}{2}}\underbrace{0\cdots0}_{\eta-1}10\cdots0\rangle + |1\overbrace{1\cdots1}^{\frac{n}{2}}\underbrace{0\cdots0}_{\eta-1}00\cdots0\rangle}{\sqrt{2}}. \end{aligned} \tag{135}$$

Similar to the previous section, we have the following proposition showing that the fermionic seminorm of nested commutators is nearly attained.

Proposition 13. *Define T, V as in (125) and $|\psi_\eta\rangle$, $|\phi_\eta\rangle$ as in (135). Then,*

$$\left.\begin{matrix} |\langle\psi_\eta|\overbrace{[V, \dots [V}^{p}, T]]|\psi_\eta\rangle|, & p \text{ odd} \\ |\langle\phi_\eta|\underbrace{[V, \dots [V}_{p}, T]]|\phi_\eta\rangle|, & p \text{ even} \end{matrix}\right\} = 2^p\eta^p + \mathcal{O}\left(\eta^{p-1}\right). \tag{136}$$

A proof of this proposition is given in Appendix D. By rescaling the Hamiltonian constructed in (127), we can demonstrate the tightness of our bound as follows. For any $s, w > 0$, we define the rescaled Hamiltonian as in (130). Comparing with the definition of the correlated-electronic model (1), we see that $\|\tau\| = s$ and $\|\nu\|_{\max} = w$. The above proposition then shows that

$$\left\| \underbrace{[V, \dots [V}_{p}, T]] \right\|_\eta = \Omega\left((w\eta)^p s/n\right). \tag{137}$$

This establishes the claimed tightness result (5) of Theorem 2.

Note that a similar example can be constructed to demonstrate the tightness of our bound for simulating sparse electronic Hamiltonians. Specifically, for $t, w > 0$ and positive integer $1 \leq d \leq \eta$, we define the electronic Hamiltonian as in (132). Comparing with the definition of the correlated-electronic model (1), we see that $\|\tau\|_{\max} = t$ and $\|\nu\|_{\max} = w$. We may then use the states

$$\begin{aligned} |\psi_{\eta,d}\rangle &= \frac{|\overbrace{01\cdots1}^{d}10\cdots0\,0\cdots0\underbrace{1\cdots1}_{\eta-\frac{d}{2}}\rangle + i|\overbrace{11\cdots1}^{d}00\cdots0\,0\cdots0\underbrace{1\cdots1}_{\eta-\frac{d}{2}}\rangle}{\sqrt{2}}, \\ |\phi_{\eta,d}\rangle &= \frac{|\overbrace{01\cdots1}^{d}10\cdots0\,0\cdots0\underbrace{1\cdots1}_{\eta-\frac{d}{2}}\rangle + |\overbrace{11\cdots1}^{d}00\cdots0\,0\cdots0\underbrace{1\cdots1}_{\eta-\frac{d}{2}}\rangle}{\sqrt{2}} \end{aligned} \tag{138}$$

to show that

$$\left\| \underbrace{[V, \dots [V}_{p}, T]] \right\|_\eta = \Omega\left((wd)^p t\right). \tag{139}$$

This proves the claimed tightness result (6) of Theorem 2.

6 Applications

The class of correlated electronic Hamiltonian (1) encompasses various quantum systems arising in physics and chemistry, for which the performance of quantum simulation can be improved using our result. As for illustration, we consider improving quantum simulation of plane-wave-basis electronic structure in Section 6.1 and the Fermi-Hubbard model in Section 6.2.

6.1 Plane-wave-basis electronic structure

Simulating the electronic-structure Hamiltonians is one of the most promising applications of digital quantum computers. Recall that in the second-quantized plane-wave basis, such a Hamiltonian takes the form

$$\begin{aligned} H = & \frac{1}{2n} \sum_{j,k,\mu} \kappa_\mu^2 \cos[\kappa_\mu \cdot r_{k-j}] A_j^\dagger A_k \\ & - \frac{4\pi}{\omega} \sum_{j,\iota,\mu\neq 0} \frac{\zeta_\iota \cos[\kappa_\mu \cdot (\widetilde{r}_\iota - r_j)]}{\kappa_\mu^2} N_j + \frac{2\pi}{\omega} \sum_{\substack{j\neq k \\ \mu\neq 0}} \frac{\cos[\kappa_\mu \cdot r_{j-k}]}{\kappa_\mu^2} N_j N_k, \end{aligned} \tag{140}$$

where ω is the volume of the computational cell, $\kappa_\mu = 2\pi\mu/\omega^{1/3}$ are n vectors of plane-wave frequencies, μ 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, and $\widetilde{r}_\iota$ are the nuclear coordinates. We further rewrite the second term as

$$-\frac{4\pi}{\omega} \sum_{j,\iota,\mu\neq 0} \frac{\zeta_\iota \cos[\kappa_\mu \cdot (\widetilde{r}_\iota - r_j)]}{\kappa_\mu^2} N_j = -\frac{4\pi}{\omega\eta} \sum_{j,k,\iota,\mu\neq 0} \frac{\zeta_\iota \cos[\kappa_\mu \cdot (\widetilde{r}_\iota - r_j)]}{\kappa_\mu^2} N_j N_k, \tag{141}$$

which is valid since we estimate the simulation error within the η-electron manifold. Comparing with the definition of correlated electronic model (1), we see that

$$\begin{aligned} \tau_{j,k} &= \frac{1}{2n} \sum_{\mu} \kappa_\mu^2 \cos[\kappa_\mu \cdot r_{k-j}], \\ \nu_{l,m} &= -\frac{4\pi}{\omega\eta} \sum_{\iota,\mu\neq 0} \frac{\zeta_\iota \cos[\kappa_\mu \cdot (\widetilde{r}_\iota - r_l)]}{\kappa_\mu^2} + \frac{2\pi}{\omega} \sum_{\mu\neq 0} \frac{\cos[\kappa_\mu \cdot r_{l-m}]}{\kappa_\mu^2} (1 - \delta_{l,m}). \end{aligned} \tag{142}$$

To proceed, we need to bound the spectral norm $\|\tau\|$ and the max-norm $\|\nu\|_{\max}$ of the coefficient matrices. We have

$$\|\tau\| = \mathcal{O}\left(\frac{n^{2/3}}{\omega^{2/3}}\right), \qquad \|\nu\|_{\max} = \mathcal{O}\left(\frac{n^{1/3}}{\omega^{1/3}}\right), \tag{143}$$

where the first equality follows from [5, Eq. (F10)] and the second equality follows from [5, Eq. (F7) and (F9)]. We also consider a constant system density $\eta = \mathcal{O}(\omega)$ following the setting of [5]. Applying Theorem 1, we find that a pth-order formula $\mathscr{S}_p(t)$ can approximates the evolution of electronic-structure Hamiltonian with Trotter error

$$\begin{aligned} \left\| \mathscr{S}_p(t) - e^{-itH} \right\|_\eta &= \mathcal{O}\left((\|\tau\| + \|\nu\|_{\max} \eta)^{p-1} \|\tau\| \|\nu\|_{\max} \eta^2 t^{p+1} \right) \\ &= \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3} \right)^p n^{1/3}\eta^{2/3} t^{p+1} \right). \end{aligned} \tag{144}$$

This approximation is accurate for sufficiently small t. To evolve for a longer time, we divide the evolution into r steps and use $\mathscr{S}_p(t/r)$ within each step, which gives an approximation with error

$$\left\|\mathscr{S}_p^r(t/r) - e^{-itH}\right\|_\eta \leq r \left\|\mathscr{S}_p(t/r) - e^{-i\frac{t}{r}H}\right\|_\eta = \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3}\right)^p n^{1/3}\eta^{2/3}\frac{t^{p+1}}{r^p}\right). \quad (145)$$

To simulate with accuracy ϵ, it suffices to choose

$$r = \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3}\right) n^{1/3p}\eta^{2/3p}\frac{t^{1+1/p}}{\epsilon^{1/p}}\right). \quad (146)$$

To simplify our discussion, we choose the order p sufficiently large and consider quantum simulation with constant time and accuracy, obtaining

$$r = \mathcal{O}\left(\left(\frac{n^{2/3}}{\eta^{2/3}} + n^{1/3}\eta^{2/3}\right) n^{o(1)}\right). \quad (147)$$

We further implement each Trotter step using the approach of [34, Sect. 5], and obtain a quantum circuit with gate complexity

$$g = \mathcal{O}\left(\left(\frac{n^{5/3}}{\eta^{2/3}} + n^{4/3}\eta^{2/3}\right) n^{o(1)}\right). \quad (148)$$

In the little-o limit, this gate complexity improves the best previous result of electronic-structure simulation in the second-quantized plane-wave basis. This is because our approach improves the performance of quantum simulation by simultaneously exploiting the commutativity of Hamiltonian and the prior knowledge of initial state, whereas previous results were only able to employ at most one of these information. Indeed, previous work [5] gave a simulation with cost $\mathcal{O}\left(\left(n^{5/3}\eta^{1/3} + n^{4/3}\eta^{5/3}\right) n^{o(1)}\right)$ by computing the Trotter error within the η-electron manifold, but the commutativity of Hamiltonian was ignored in their analysis. On the other hand, the work [19] used the commutativity of Hamiltonian to give a Trotterization algorithm with complexity $\mathcal{O}\left(\frac{n^{7/3}}{\eta^{1/3}} n^{o(1)}\right)$, whereas [34] gave an interaction-picture approach with cost $\mathcal{O}\left(\frac{n^{8/3}}{\eta^{2/3}}\mathrm{polylog}(n)\right)$. Our new result matches these when η and n are comparable to each other, but can be much more efficient in the regime where η is much smaller than n.

Interestingly, our result remains conditionally advantageous even when compared with the first-quantized simulations. There, the best previous approach is the interaction-picture approach [3] with gate complexity $\mathcal{O}\left(n^{\frac{1}{3}}\eta^{\frac{8}{3}}\mathrm{polylog}(n)\right)$, larger than our new complexity in the little-o limit when $\eta = \Omega\left(\sqrt{n}\right)$. A related approach was described in [3] based on qubitization, which has a similar performance comparison with our result. See Table 1 for details.

We mention however that there is one caveat when taking the little-o limit in our above discussion. This limit is achieved by choosing the order p of Trotterization sufficiently large, which can result in a gate complexity with an unrealistically large prefactor due to the definition (17) of higher-order formulas. Nevertheless, recent work suggests that Trotterization remains advantageous for simulating the plane-wave-basis electronic structure even with a low-order formula [26], to which our paper provides new theoretical insights.

6.2 Fermi-Hubbard model

We also consider applications of our result to the simulation of Fermi-Hubbard Hamiltonian, which models many important properties of correlated electrons. This Hamiltonian is defined as

$$H = -s \sum_{\langle j,k\rangle,\sigma} \left(A_{j,\sigma}^\dagger A_{k,\sigma} + A_{k,\sigma}^\dagger A_{j,\sigma} \right) + v \sum_j N_{j,0} N_{j,1}, \tag{149}$$

where $\langle j,k\rangle$ denotes a summation over nearest-neighbor lattice sites and $\sigma \in \{0,1\}$.

We note that this Hamiltonian can be represented in terms of a sparse correlated Hamiltonian. Indeed, in the one-dimensional case, we have

$$H = -s \sum_{j,\sigma} \left(A_{j,\sigma}^\dagger A_{j+1,\sigma} + A_{j+1,\sigma}^\dagger A_{j,\sigma} \right) + v \sum_j N_{j,0} N_{j,1}, \tag{150}$$

where $j = 0, 1, \ldots, n-1$ and $\sigma = 0, 1$. Comparing with the definition of correlated electronic model (1), we see that

$$\tau = -s \sum_j \left(|j\rangle\langle j+1| + |j+1\rangle\langle j| \right) \otimes \left(|0\rangle\langle 0| + |1\rangle\langle 1| \right), \qquad \nu = \frac{v}{2} \sum_j |j\rangle\langle j| \otimes \left(|0\rangle\langle 1| + |1\rangle\langle 0| \right), \tag{151}$$

so the coefficient matrices τ and ν are indeed 2-sparse. Similar analysis holds for the higher-dimensional Fermi-Hubbard model, with the sparsity $d = 2^m$ where m is the dimensionality of the lattice.

We can therefore apply Theorem 1 to conclude that a pth-order formula $\mathscr{S}_p(t)$ approximates the evolution of Fermi-Hubbard Hamiltonian with Trotter error

$$\left\| \mathscr{S}_p(t) - e^{-itH} \right\|_\eta = \mathcal{O}\left((s+v)^{p-1} s v 2^{mp} \eta t^{p+1} \right) = \mathcal{O}\left(\eta t^{p+1} \right), \tag{152}$$

assuming s, v, and m are constant. For r steps of Trotterization, we apply the triangle inequality to get

$$\left\| \mathscr{S}_p^r(t/r) - e^{-itH} \right\|_\eta \leq r \left\| \mathscr{S}_p(t/r) - e^{-i\frac{t}{r}H} \right\|_\eta = \mathcal{O}\left(\eta \frac{t^{p+1}}{r^p} \right). \tag{153}$$

To simulate with constant time and accuracy, it thus suffices to choose

$$r = \mathcal{O}\left(\eta^{1/p} \right), \tag{154}$$

giving gate complexity

$$g = \mathcal{O}\left(n \eta^{1/p} \right). \tag{155}$$

The Fermi-Hubbard model only contains nearest-neighbor interactions and, according to [18], can be near optimally simulated with $\mathcal{O}\left(n^{1+1/p}\right)$ gates. On the other hand, recent work [20] shows that Trotterization algorithm has gate complexity $\mathcal{O}\left(n\eta^{1+1/p}\right)$ when restricted to the η-electron manifold. Our result again improves over these previous work by combining the commutativity of Hamiltonian and the initial-state information.

7 Discussion

We have given improved quantum simulations using Trotterization for a class of correlated electrons, by simultaneously exploiting the commutativity of Hamiltonian, the sparsity of interactions, and the prior knowledge of initial state. We identified applications to simulating the plane-wave-basis electronic structure, improving the best previous result in second quantization up to a negligible factor while conditionally outperforming the first-quantized simulation. We obtained further speedups when the electronic Hamiltonian has d-sparse interactions, using which we gave faster Trotterization of the Fermi-Hubbard model. We constructed concrete electronic systems for which our bounds are almost saturated, providing a provable guarantee on the tightness of our analysis.

Our focus has been on the asymptotic performance of quantum simulation throughout this paper. However, we believe that the techniques we have developed can also be used to give quantum simulations with low constant-prefactor overhead. Such improvements would especially benefit the simulation of plane-wave-basis electronic structure, where many pairs of Hamiltonian terms commute and the number of spin orbitals can be significantly larger than the electron number. Previous work on such problems almost exclusively used the second-order Suzuki formula [4, 26, 47], and we hope future work could consider other low-order Trotterizations that are still easy to implement but may offer advantages over the second-order formula in practice.

Our analysis is applicable to a class of electronic Hamiltonians of the form $H = \sum_{j,k} \tau_{j,k} A_j^\dagger A_k + \sum_{l,m} \nu_{l,m} N_l N_m$. By imposing further constraints on the coefficients, we may somewhat sacrifice this generality but instead get further improvement on the simulation performance. One possibility is to consider the subclass of systems that are translation-invariant, i.e., $\tau_{j,k} = \tau_{j+q,k+q}$ and $\nu_{l,m} = \nu_{l+q,m+q}$. This translational invariance is used in the circuit implementations for both our applications (electronic-structure Hamiltonians and Fermi-Hubbard model), but is nevertheless ignored in the proof of our upper bounds (Theorem 1) and tightness result (Theorem 2). By incorporating additional features of the Hamiltonian such as translational invariance, it is plausible that our current complexity estimate can be further improved.

A natural problem that has yet to be addressed is the simulation of electronic-structure Hamiltonians in a more compact molecular basis. Such Hamiltonians typically take the form $H = \sum_{j,k} h_{j,k} A_j^\dagger A_k + \sum_{j,k,l,m} h_{j,k,l,m} A_j^\dagger A_k A_l^\dagger A_m$, more complex than the electronic model (1) considered here. In this case, the exponential of the two-body terms $\sum_{j,k,l,m} h_{j,k,l,m} A_j^\dagger A_k A_l^\dagger A_m$ does not have a convenient circuit implementation and our current approach is not directly applicable. This may motivate further developments of *hybrid quantum simulation*, in which Trotterization is combined with more advanced quantum algorithms to speed up quantum simulation. We leave a detailed study of such problems as a subject for future work.

More generally, we could consider quantum simulations of other physical systems. For example, previous work by Somma considered simulating bosonic systems by combining the commutativity of Hamiltonian and the initial-state information, although his solution seems to have a divergent issue in general. We hope our techniques could offer insights to such issues and find further applications in quantum simulation beyond what have been discussed here.

Acknowledgements

YS thanks Nathan Wiebe, Guang Hao Low, Ryan Babbush, Kunal Sharma, John Preskill, and Andrew Childs for helpful discussions. We thank Fernando Brandão for inspiring discussions during the initial stages of this work.

A Analysis of single-layer commutators

In this appendix, we complete the proof of Proposition 7 on bounding the terms arising in the commutator analysis of first-order formula.

For the third statement of Proposition 7, we let $X = \sum_{j,k,l} \tau_{j,k} \nu_{l,k} A_j^\dagger N_l A_k$ and compute

$$\begin{aligned} X^\dagger X &= \sum_{j_1,k_1,l_1,j_2,k_2,l_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{l_1,k_1} \tau_{j_2,k_2} \nu_{l_2,k_2} A_{k_1}^\dagger N_{l_1} A_{j_1} A_{j_2}^\dagger N_{l_2} A_{k_2} \\ &= \sum_{j_1,k_1,l_1,k_2,l_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{l_1,k_1} \tau_{j_1,k_2} \nu_{l_2,k_2} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} \\ &\quad - \sum_{j_1,k_1,l_1,j_2,k_2,l_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{l_1,k_1} \tau_{j_2,k_2} \nu_{l_2,k_2} A_{k_1}^\dagger N_{l_1} A_{j_2}^\dagger A_{j_1} N_{l_2} A_{k_2}. \end{aligned} \tag{156}$$

Applying the operator Cauchy-Schwarz inequality Lemma 1,

$$\begin{aligned} X^\dagger X &\leq \sum_{j_1,k_1,l_1,k_2,l_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{l_1,k_1} \tau_{j_1,k_2} \nu_{l_2,k_2} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} \\ &\quad + \sum_{j_1,k_1,l_1,j_2,k_2,l_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{l_1,k_1} \tau_{j_1,k_2} \nu_{l_2,k_2} A_{k_1}^\dagger N_{l_1} A_{j_2}^\dagger A_{j_2} N_{l_2} A_{k_2} \\ &= \sum_{j_1,k_1,l_1,k_2,l_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{l_1,k_1} \tau_{j_1,k_2} \nu_{l_2,k_2} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} N. \end{aligned} \tag{157}$$

We now perform diagonalization using Lemma 2, obtaining

$$X^\dagger X \leq \|\tau\|^2 \sum_{k_1,l_1,l_2} \bar{\nu}_{l_1,k_1} \nu_{l_2,k_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_1} N. \tag{158}$$

Using the Hölder-type inequality for expectation Lemma 3, we have

$$\begin{aligned} \left\| X^\dagger X \right\|_\eta &\leq \left\| \|\tau\|^2 \sum_{k_1,l_1,l_2} \bar{\nu}_{l_1,k_1} \nu_{l_2,k_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_1} N \right\|_\eta = \|\tau\|^2 \eta \left\| \sum_{k_1,l_1,l_2} \bar{\nu}_{l_1,k_1} \nu_{l_2,k_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_1} \right\|_\eta \\ &\leq \|\tau\|^2 \eta \left\| \sum_{k_1} A_{k_1}^\dagger A_{k_1} \right\|_\eta \max_{k_1} \left\| \sum_{l_1,l_2} \bar{\nu}_{l_1,k_1} \nu_{l_2,k_1} N_{l_1} N_{l_2} \right\|_{\eta-1}, \end{aligned} \tag{159}$$

where $\left\| \sum_{k_1} A_{k_1}^\dagger A_{k_1} \right\|_\eta = \eta$ and $\left\| \sum_{l_1,l_2} \bar{\nu}_{l_1,k_1} \nu_{l_2,k_1} N_{l_1} N_{l_2} \right\|_{\eta-1} \leq \|\nu\|_{\max}^2 \eta^2$. This completes the proof of the third statement of Proposition 7.

For the fourth statement, we let $X = \sum_{j,k,m} \tau_{j,k} \nu_{j,m} A_j^\dagger N_m A_k$ and compute

$$\begin{aligned} X^\dagger X &= \sum_{j_1,k_1,m_1,j_2,k_2,m_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{j_1,m_1} \tau_{j_2,k_2} \nu_{j_2,m_2} A_{k_1}^\dagger N_{m_1} A_{j_1} A_{j_2}^\dagger N_{m_2} A_{k_2} \\ &= \sum_{j_1,k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{j_1,m_1} \tau_{j_1,k_2} \nu_{j_1,m_2} A_{k_1}^\dagger N_{m_1} N_{m_2} A_{k_2} \\ &\quad - \sum_{j_1,k_1,m_1,j_2,k_2,m_2} \bar{\tau}_{j_1,k_1} \bar{\nu}_{j_1,m_1} \tau_{j_2,k_2} \nu_{j_2,m_2} A_{k_1}^\dagger N_{m_1} A_{j_2}^\dagger A_{j_1} N_{m_2} A_{k_2}. \end{aligned} \tag{160}$$

Applying the operator Cauchy-Schwarz inequality Lemma 1,

$$\begin{aligned} X^\dagger X \leq & \sum_{j_1,k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,m_1}\tau_{j_1,k_2}\nu_{j_1,m_2}A^\dagger_{k_1}N_{m_1}N_{m_2}A_{k_2} \\ & + \sum_{j_1,k_1,m_1,j_2,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,m_1}\tau_{j_1,k_2}\nu_{j_1,m_2}A^\dagger_{k_1}N_{m_1}A^\dagger_{j_2}A_{j_2}N_{m_2}A_{k_2} \\ = & \sum_{j_1,k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,m_1}\tau_{j_1,k_2}\nu_{j_1,m_2}A^\dagger_{k_1}N_{m_1}N_{m_2}A_{k_2}N. \end{aligned} \tag{161}$$

We now use the Hölder-type inequality for expectation Lemma 3 to get

$$\begin{aligned} \left\|X^\dagger X\right\|_\eta &\leq \left\| \sum_{j_1,k_1,m_1,k_2,m_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,m_1}\tau_{j_1,k_2}\nu_{j_1,m_2}A^\dagger_{k_1}N_{m_1}N_{m_2}A_{k_2}N \right\|_\eta \\ &\leq \eta \left\| \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\tau_{j_1,k_2}A^\dagger_{k_1}A_{k_2} \right\|_\eta \max_{j_1} \left\| \sum_{m_1,m_2} \bar{\nu}_{j_1,m_1}\nu_{j_1,m_2}N_{m_1}N_{m_2} \right\|_{\eta-1}. \end{aligned} \tag{162}$$

The second fermionic seminorm can be directly bounded as $\left\|\sum_{m_1,m_2} \bar{\nu}_{j_1,m_1}\nu_{j_1,m_2}N_{m_1}N_{m_2}\right\|_{\eta-1} \leq \|\nu\|^2_{\max}\eta^2$, whereas the first seminorm can be bounded using diagonalization Lemma 2

$$\left\| \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\tau_{j_1,k_2}A^\dagger_{k_1}A_{k_2} \right\|_\eta \leq \left\| \sum_{k_1} \left\|\tau^\dagger\tau\right\| A^\dagger_{k_1}A_{k_1} \right\|_\eta \leq \left\|\tau^\dagger\tau\right\| \eta. \tag{163}$$

This completes the proof of the fourth statement of Proposition 7.

For the fifth statement, we let $X = \sum_{j,k} \tau_{j,k}\nu_{j,j}A^\dagger_j A_k$ and compute

$$\begin{aligned} X^\dagger X &= \sum_{j_1,k_1,j_2,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_2,k_2}\nu_{j_2,j_2}A^\dagger_{k_1}A_{j_1}A^\dagger_{j_2}A_{k_2} \\ &= \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_1,k_2}\nu_{j_1,j_1}A^\dagger_{k_1}A_{k_2} - \sum_{j_1,k_1,j_2,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_2,k_2}\nu_{j_2,j_2}A^\dagger_{k_1}A^\dagger_{j_2}A_{j_1}A_{k_2}. \end{aligned} \tag{164}$$

Applying the operator Cauchy-Schwarz inequality Lemma 1,

$$\begin{aligned} X^\dagger X &\leq \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_1,k_2}\nu_{j_1,j_1}A^\dagger_{k_1}A_{k_2} + \sum_{j_1,k_1,j_2,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_1,k_2}\nu_{j_1,j_2}A^\dagger_{k_1}A^\dagger_{j_2}A_{j_2}A_{k_2} \\ &= \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_1,k_2}\nu_{j_1,j_1}A^\dagger_{k_1}A_{k_2}N. \end{aligned} \tag{165}$$

We now use the Hölder-type inequality for expectation Lemma 3 to get

$$\begin{aligned} \left\|X^\dagger X\right\|_\eta &\leq \left\| \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\bar{\nu}_{j_1,j_1}\tau_{j_1,k_2}\nu_{j_1,j_1}A^\dagger_{k_1}A_{k_2}N \right\|_\eta \\ &= \eta \left\| \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1}\tau_{j_1,k_2}A^\dagger_{k_1}A_{k_2} \right\|_\eta \max_{j_1} \left\|\bar{\nu}_{j_1,j_1}\nu_{j_1,j_1}I\right\|_{\eta-1}. \end{aligned} \tag{166}$$

The second fermionic seminorm can be directly bounded by $\|\nu\|_{\max}^2$, whereas we perform diagonalization to the first seminorm Lemma 2:

$$\left\| \sum_{j_1,k_1,k_2} \bar{\tau}_{j_1,k_1} \tau_{j_1,k_2} A_{k_1}^\dagger A_{k_2} \right\|_\eta \leq \left\| \tau^\dagger \tau \right\| \left\| \sum_{k_1} A_{k_1}^\dagger A_{k_1} \right\|_\eta = \left\| \tau^\dagger \tau \right\| \eta. \tag{167}$$

This completes the proof of the fifth statement of Proposition 7.

For the sixth statement, we let $X = \sum_{j,k,l} \tau_{j,k} \nu_{l,j} A_j^\dagger N_l A_k$ and compute

$$\begin{aligned} X^\dagger X &= \sum_{j_1,k_1,l_1,j_2,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_2,k_2} \nu_{l_2,j_2} A_{k_1}^\dagger N_{l_1} A_{j_1} A_{j_2}^\dagger N_{l_2} A_{k_2} \\ &= \sum_{j_1,k_1,l_1,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_1,k_2} \nu_{l_2,j_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} \\ &\quad - \sum_{j_1,k_1,l_1,j_2,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_2,k_2} \nu_{l_2,j_2} A_{k_1}^\dagger N_{l_1} A_{j_2}^\dagger A_{j_1} N_{l_2} A_{k_2}. \end{aligned} \tag{168}$$

Applying the operator Cauchy-Schwarz inequality Lemma 1,

$$\begin{aligned} X^\dagger X &\leq \sum_{j_1,k_1,l_1,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_1,k_2} \nu_{l_2,j_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} \\ &\quad + \sum_{j_1,k_1,l_1,j_2,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_1,k_2} \nu_{l_2,j_1} A_{k_1}^\dagger N_{l_1} A_{j_2}^\dagger A_{j_2} N_{l_2} A_{k_2} \\ &= \sum_{j_1,k_1,l_1,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_1,k_2} \nu_{l_2,j_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} N. \end{aligned} \tag{169}$$

We now use the Hölder-type inequality for expectation Lemma 3 to get

$$\begin{aligned} \left\| X^\dagger X \right\|_\eta &\leq \left\| \sum_{j_1,k_1,l_1,k_2,l_2} \tau_{j_1,k_1} \nu_{l_1,j_1} \tau_{j_1,k_2} \nu_{l_2,j_1} A_{k_1}^\dagger N_{l_1} N_{l_2} A_{k_2} N \right\|_\eta \\ &= \eta \left\| \sum_{j_1,k_1,k_2} \tau_{j_1,k_1} \tau_{j_1,k_2} A_{k_1}^\dagger A_{k_2} \right\|_\eta \max_{j_1} \left\| \sum_{l_1,l_2} \nu_{l_1,j_1} \nu_{l_2,j_1} N_{l_1} N_{l_2} \right\|_{\eta-1}. \end{aligned} \tag{170}$$

The second fermionic seminorm can be directly bounded by $\|\nu\|_{\max}^2 \eta^2$, whereas we perform diagonalization to the first seminorm Lemma 2:

$$\left\| \sum_{j_1,k_1,k_2} \tau_{j_1,k_1} \tau_{j_1,k_2} A_{k_1}^\dagger A_{k_2} \right\|_\eta \leq \left\| \tau^\dagger \tau \right\| \left\| \sum_{k_1} A_{k_1}^\dagger A_{k_2} \right\|_\eta = \left\| \tau^\dagger \tau \right\| \eta. \tag{171}$$

This completes the proof of the sixth statement of Proposition 7.

B Analysis of multilayer nested commutators

In this appendix, we complete the proof of Proposition 8 on bounding the terms arising in the commutator analysis of pth-order formulas. Recall from Proposition 8 that operator X is called a fermionic chain if

$$X = \sum_{j,k} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot A_{j_q}^\dagger \prod_{x=1}^{q} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}, \tag{172}$$

where all $B_{x,y}$, $C_{x,z}$ and hence the entire chain X are number-preserving. Our goal is to prove the bound

$$\|X\|_\eta \leq \|\tau\|^q \, \eta \prod_{x=1}^{q} \left(\prod_y \max_{j_x} \|B_{x,y}\|_{\eta-1} \prod_z \max_{k_x} \|C_{x,z}\|_{\eta-1} \right). \tag{173}$$

We will prove this bound using Lemma 1, Lemma 2, and Lemma 3 in a similar way as in Proposition 7. Specifically, we write $X = \sum_{j_q} A^\dagger_{j_q} D_{j_q}$, where

$$D_{j_q} = \sum_{\substack{j_1,\ldots,\\ j_{q-1},k}} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot \prod_{x=1}^{q} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}. \tag{174}$$

Then,

$$X^\dagger X = \sum_{j_{q_1},j_{q_2}} D^\dagger_{j_{q_1}} A_{j_{q_1}} A^\dagger_{j_{q_2}} D_{j_{q_2}} = \sum_{j_{q_1}} D^\dagger_{j_{q_1}} D_{j_{q_1}} - \sum_{j_{q_1},j_{q_2}} D^\dagger_{j_{q_1}} A^\dagger_{j_{q_2}} A_{j_{q_1}} D_{j_{q_2}}. \tag{175}$$

Applying the operator Cauchy-Schwarz inequality (Lemma 1), we obtain

$$X^\dagger X \leq \sum_{j_{q_1}} D^\dagger_{j_{q_1}} D_{j_{q_1}} + \sum_{j_{q_1},j_{q_2}} D^\dagger_{j_{q_1}} A^\dagger_{j_{q_2}} A_{j_{q_2}} D_{j_{q_1}} = \sum_{j_{q_1}} D^\dagger_{j_{q_1}} D_{j_{q_1}} N. \tag{176}$$

Next, we write $D_{j_q} = \prod_y B_{q,y} E_{j_q}$, where

$$E_{j_q} = \sum_{\substack{j_1,\ldots,\\ j_{q-1},k}} \prod_{x=1}^{q} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot \prod_z C_{q,z} \prod_{x=1}^{q-1} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}. \tag{177}$$

Invoking the Hölder-type inequality for expectation (Lemma 3), we get

$$\|X\|_\eta = \sqrt{\|X^\dagger X\|_\eta} \leq \eta^{1/2} \sqrt{\left\| \sum_{j_{q_1}} D^\dagger_{j_{q_1}} D_{j_{q_1}} \right\|_\eta} \leq \eta^{1/2} \sqrt{\left\| \sum_{j_{q_1}} E^\dagger_{j_{q_1}} E_{j_{q_1}} \right\|_\eta} \prod_y \max_{j_q} \|B_{q,y}\|_{\eta-1}. \tag{178}$$

We now write $E_{j_q} = \sum_{k_q} \tau_{j_q,k_q} F_{k_q}$, where

$$F_{k_q} = \sum_{\substack{j_1,\ldots,j_{q-1},\\ k_1,\ldots,k_{q-1}}} \prod_{x=1}^{q-1} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot \prod_z C_{q,z} \prod_{x=1}^{q-1} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}. \tag{179}$$

Then,

$$\sum_{j_{q_1}} E^\dagger_{j_{q_1}} E_{j_{q_1}} = \sum_{k_{q_1},k_{q_2}} \left(\sum_{j_{q_1}} \bar{\tau}_{j_{q_1},k_{q_1}} \tau_{j_{q_1},k_{q_2}} \right) F^\dagger_{k_{q_1}} F_{k_{q_2}}. \tag{180}$$

We perform diagonalization using Lemma 2, obtaining

$$\sum_{j_{q_1}} E^\dagger_{j_{q_1}} E_{j_{q_1}} \leq \left\| \tau^\dagger \tau \right\| \sum_{k_{q_1}} F^\dagger_{k_{q_1}} F_{k_{q_1}}. \tag{181}$$

Next, we write $F_{k_q} = \prod_z C_{q,z} G_{k_q}$, where

$$G_{k_q} = \sum_{\substack{j_1,\ldots,j_{q-1},\\ k_1,\ldots,k_{q-1}}} \prod_{x=1}^{q-1} \tau_{j_x,k_x} \prod_{x=1}^{q-1} \delta_{k_{x+1},j_x} \cdot \prod_{x=1}^{q-1} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}. \tag{182}$$

Invoking again the Hölder-type inequality for expectation (Lemma 3), we get

$$\|X\|_\eta \leq \|\tau\|\, \eta^{1/2} \sqrt{\Big\| \sum_{k_{q_1}} G^\dagger_{k_{q_1}} G_{k_{q_1}} \Big\|_\eta} \prod_y \max_{j_q} \|B_{q,y}\|_{\eta-1} \prod_z \max_{k_q} \|C_{q,z}\|_{\eta-1}\,. \tag{183}$$

Note that we can write $G_{k_q} = \sum_{j_{q-1}} \delta_{k_q,j_{q-1}} H_{j_{q-1}}$ with

$$H_{j_{q-1}} = \sum_{\substack{j_1,\ldots,j_{q-2},\\ k_1,\ldots,k_{q-1}}} \prod_{x=1}^{q-1} \tau_{j_x,k_x} \prod_{x=1}^{q-2} \delta_{k_{x+1},j_x} \cdot \prod_{x=1}^{q-1} \left(\prod_y B_{x,y} \prod_z C_{x,z} \right) A_{k_1}, \tag{184}$$

which implies

$$\sum_{k_q} G^\dagger_{k_q} G_{k_q} = \sum_{j_{q-1}} H^\dagger_{j_{q-1}} H_{j_{q-1}}. \tag{185}$$

We can now iterate this procedure q times to get

$$\begin{aligned} \|X\|_\eta &\leq \|\tau\|^q\, \eta^{1/2} \sqrt{\Big\| \sum_{k_1} A^\dagger_{k_1} A_{k_1} \Big\|_\eta} \prod_{x=1}^{q} \left(\prod_y \max_{j_x} \|B_{x,y}\|_{\eta-1} \prod_z \max_{k_x} \|C_{x,z}\|_{\eta-1} \right) \\ &= \|\tau\|^q\, \eta \prod_{x=1}^{q} \left(\prod_y \max_{j_x} \|B_{x,y}\|_{\eta-1} \prod_z \max_{k_x} \|C_{x,z}\|_{\eta-1} \right). \end{aligned} \tag{186}$$

This completes the proof of Proposition 8.

C Lower-bounding $\|[T, \ldots [T, V]]\|_\eta$

In this appendix, we prove Proposition 12 that lower-bounds the fermionic seminorm $\|[T, \ldots [T, V]]\|_\eta$ for the electronic Hamiltonian (125). After the fermionic Fourier transform (126), we have

$$\begin{aligned} \widetilde{T} &= \mathrm{FFFT}^\dagger \cdot T \cdot \mathrm{FFFT} = n N_0, \\ \widetilde{V} &= \mathrm{FFFT}^\dagger \cdot V \cdot \mathrm{FFFT} = \frac{1}{n^2} \sum_{j,k,l,m} \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u(k-j)}{n}} \right) \left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v(m-l)}{n}} \right) A^\dagger_j A_k A^\dagger_l A_m, \end{aligned} \tag{187}$$

which gives the commutator

$$\begin{aligned} \left[\widetilde{T}, \widetilde{V}\right] = \frac{1}{n} \sum_{k,l,m} \tau_{0klm} A^\dagger_0 A_k A^\dagger_l A_m - \frac{1}{n} \sum_{j,l,m} \tau_{j0lm} A^\dagger_j A_0 A^\dagger_l A_m \\ + \frac{1}{n} \sum_{j,k,m} \tau_{jk0m} A^\dagger_j A_k A^\dagger_0 A_m \; - \frac{1}{n} \sum_{j,k,l} \tau_{jkl0} A^\dagger_j A_k A^\dagger_l A_0 \end{aligned} \tag{188}$$

with

$$\tau_{jklm} := \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u(k-j)}{n}} \right) \left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v(m-l)}{n}} \right). \tag{189}$$

We will choose the initial state from the two-dimensional subspace spanned by

$$|\widetilde{\psi}_0\rangle = |010\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle, \qquad |\widetilde{\psi}_1\rangle = |100\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle. \tag{190}$$

Denoting the projection to this subspace as $\widetilde{\Pi} = |\widetilde{\psi_0}\rangle\langle\widetilde{\psi_0}| + |\widetilde{\psi_1}\rangle\langle\widetilde{\psi_1}|$, we have that $\widetilde{\Pi}$ commutes with $\widetilde{T} = nN_0$, which means $\widetilde{\Pi}[\widetilde{T}, \dots [\widetilde{T}, \widetilde{V}]]\widetilde{\Pi} = [\widetilde{T}, \dots \widetilde{\Pi}[\widetilde{T}, \widetilde{V}]\widetilde{\Pi}]$. We simplify the effective commutator $\widetilde{\Pi}[\widetilde{T}, \widetilde{V}]\widetilde{\Pi}$ based on the following observations:

1. $A_0^\dagger A_k A_l^\dagger A_m$: This will always nullify $|\widetilde{\psi_0}\rangle$ from left. For $\langle\widetilde{\psi_1}|A_0^\dagger A_k A_l^\dagger A_m|\widetilde{\psi_1}\rangle$ to be nonzero, we must let one of $\{k, m\}$ be 0, while the other is equal to l. For $\langle\widetilde{\psi_1}|A_0^\dagger A_k A_l^\dagger A_m|\widetilde{\psi_0}\rangle$ to be nonzero, we must let one of $\{k, m\}$ be 1, while the other is equal to l.

2. $A_j^\dagger A_0 A_l^\dagger A_m$: For $\langle\widetilde{\psi_0}|A_j^\dagger A_0 A_l^\dagger A_m|\widetilde{\psi_0}\rangle$ to be nonzero, we must let $l = 0$ and $j = m$. For $\langle\widetilde{\psi_0}|A_j^\dagger A_0 A_l^\dagger A_m|\widetilde{\psi_1}\rangle$ to be nonzero, we must let one of $\{j, l\}$ be 1, while the other is equal to m. For $\langle\widetilde{\psi_1}|A_j^\dagger A_0 A_l^\dagger A_m|\widetilde{\psi_1}\rangle$ to be nonzero, we must let $j = 0$ and $l = m$. For $\langle\widetilde{\psi_1}|A_j^\dagger A_0 A_l^\dagger A_m|\widetilde{\psi_0}\rangle$ to be nonzero, we must let $j = 0$, $l = 0$ and $m = 1$.

3. $A_j^\dagger A_k A_0^\dagger A_m$: For $\langle\widetilde{\psi_0}|A_j^\dagger A_k A_0^\dagger A_m|\widetilde{\psi_0}\rangle$ to be nonzero, we must let $k = 0$ and $j = m$. For $\langle\widetilde{\psi_1}|A_j^\dagger A_k A_0^\dagger A_m|\widetilde{\psi_0}\rangle$ to be nonzero, we must let one of $\{k, m\}$ be 1, while the other is equal to j. For $\langle\widetilde{\psi_1}|A_j^\dagger A_k A_0^\dagger A_m|\widetilde{\psi_1}\rangle$ to be nonzero, we must let $m = 0$ and $j = k$. For $\langle\widetilde{\psi_0}|A_j^\dagger A_k A_0^\dagger A_m|\widetilde{\psi_1}\rangle$ to be nonzero, we must let $m = 0$, $k = 0$ and $j = 1$.

4. $A_j^\dagger A_k A_l^\dagger A_0$: This will always nullify $|\widetilde{\psi_0}\rangle$ from right. For $\langle\widetilde{\psi_1}|A_j^\dagger A_k A_l^\dagger A_0|\widetilde{\psi_1}\rangle$ to be nonzero, we must let one of $\{j, l\}$ be 0, while the other is equal to k. For $\langle\widetilde{\psi_0}|A_j^\dagger A_k A_l^\dagger A_0|\widetilde{\psi_1}\rangle$ to be nonzero, we must let one of $\{j, l\}$ be 1, while the other is equal to k.

After removing double-counting and canceling redundant terms, we obtain

$$
\begin{aligned}
&\widetilde{\Pi}\left[\widetilde{T}, \widetilde{V}\right]\widetilde{\Pi} \\
&= \cancel{\frac{1}{n}\sum_l \tau_{00ll} A_0^\dagger A_0 A_l^\dagger A_l} + \frac{1}{n}\sum_l \tau_{01ll} A_0^\dagger A_1 A_l^\dagger A_l + \cancel{\frac{1}{n}\sum_k \tau_{0kk0} A_0^\dagger A_k A_k^\dagger A_0} + \frac{1}{n}\sum_k \tau_{0kk1} A_0^\dagger A_k A_k^\dagger A_1 \\
&\quad \cancel{-\frac{1}{n}\tau_{0000} A_0^\dagger A_0 A_0^\dagger A_0} - \frac{1}{n}\tau_{0111} A_0^\dagger A_1 A_1^\dagger A_1 \\
&\quad \cancel{-\frac{1}{n}\sum_j \tau_{j00j} A_j^\dagger A_0 A_0^\dagger A_j} - \frac{1}{n}\sum_l \tau_{10ll} A_1^\dagger A_0 A_l^\dagger A_l - \frac{1}{n}\sum_j \tau_{j01j} A_j^\dagger A_0 A_1^\dagger A_j \cancel{- \frac{1}{n}\sum_l \tau_{00ll} A_0^\dagger A_0 A_l^\dagger A_l} - \frac{1}{n}\tau_{0001} A_0^\dagger A_0 A_0^\dagger A_1 \\
&\quad \cancel{+\frac{1}{n}\tau_{0000} A_0^\dagger A_0 A_0^\dagger A_0} \underbrace{+\frac{1}{n}\tau_{1011} A_1^\dagger A_0 A_1^\dagger A_1}_{0} \\
&\quad \cancel{+\frac{1}{n}\sum_j \tau_{j00j} A_j^\dagger A_0 A_0^\dagger A_j} + \frac{1}{n}\sum_j \tau_{jj01} A_j^\dagger A_j A_0^\dagger A_1 + \frac{1}{n}\sum_j \tau_{j10j} A_j^\dagger A_1 A_0^\dagger A_j \cancel{+ \frac{1}{n}\sum_j \tau_{jj00} A_j^\dagger A_j A_0^\dagger A_0} + \frac{1}{n}\tau_{1000} A_1^\dagger A_0 A_0^\dagger A_0 \\
&\quad \cancel{-\frac{1}{n}\tau_{0000} A_0^\dagger A_0 A_0^\dagger A_0} \underbrace{-\frac{1}{n}\tau_{1101} A_1^\dagger A_1 A_0^\dagger A_1}_{0} \\
&\quad \cancel{-\frac{1}{n}\sum_k \tau_{0kk0} A_0^\dagger A_k A_k^\dagger A_0} - \frac{1}{n}\sum_j \tau_{jj10} A_j^\dagger A_j A_1^\dagger A_0 \cancel{- \frac{1}{n}\sum_j \tau_{jj00} A_j^\dagger A_j A_0^\dagger A_0} - \frac{1}{n}\sum_k \tau_{1kk0} A_1^\dagger A_k A_k^\dagger A_0 \\
&\quad \cancel{+\frac{1}{n}\tau_{0000} A_0^\dagger A_0 A_0^\dagger A_0} + \frac{1}{n}\tau_{1110} A_1^\dagger A_1 A_1^\dagger A_0.
\end{aligned}
\tag{191}
$$

We merge the remaining twelve terms into four groups:

1. The first group contains terms

$$\begin{aligned}&\frac{1}{n}\sum_l \tau_{01ll}A_0^\dagger A_1 A_l^\dagger A_l - \frac{1}{n}\sum_l \tau_{10ll}A_1^\dagger A_0 A_l^\dagger A_l + \frac{1}{n}\sum_j \tau_{jj01}A_j^\dagger A_j A_0^\dagger A_1 - \frac{1}{n}\sum_j \tau_{jj10}A_j^\dagger A_j A_1^\dagger A_0\\ =&N\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u}{n}}\right)A_0^\dagger A_1 - N\left(\sum_{u=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i u}{n}}\right)A_1^\dagger A_0.\end{aligned} \tag{192}$$

   We will see that this is the dominant contribution to the effective commutator that is at least $\Omega(n\eta)$.

2. The second group contains terms

$$\begin{aligned}&-\frac{1}{n}\tau_{0111}A_0^\dagger A_1 A_1^\dagger A_1 - \frac{1}{n}\tau_{0001}A_0^\dagger A_0 A_0^\dagger A_1 + \frac{1}{n}\tau_{1000}A_1^\dagger A_0 A_0^\dagger A_0 + \frac{1}{n}\tau_{1110}A_1^\dagger A_1 A_1^\dagger A_0\\ =&-\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u}{n}}\right)A_0^\dagger A_1 + \left(\sum_{u=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i u}{n}}\right)A_1^\dagger A_0.\end{aligned} \tag{193}$$

   These terms have asymptotic scaling $O(n)$ and thus do not dominate the result.

3. The third group contains terms

$$\begin{aligned}&\frac{1}{n}\sum_k \tau_{0kk1}A_0^\dagger A_k A_k^\dagger A_1 - \frac{1}{n}\sum_k \tau_{1kk0}A_1^\dagger A_k A_k^\dagger A_0\\ =&\frac{1}{n}\tau_{0001}A_0^\dagger A_1 + \frac{1}{n}\sum_k \tau_{0kk1}A_k A_k^\dagger A_0^\dagger A_1\\ &-\frac{1}{n}\tau_{1110}A_1^\dagger A_0 - \frac{1}{n}\sum_k \tau_{1kk0}A_k A_k^\dagger A_1^\dagger A_0\\ =&\frac{1}{n}\tau_{0001}A_0^\dagger A_1 + \frac{1}{n}\sum_k \tau_{0kk1}A_0^\dagger A_1 - \frac{1}{n}\sum_k \tau_{0kk1}A_k^\dagger A_k A_0^\dagger A_1\\ &-\frac{1}{n}\tau_{1110}A_1^\dagger A_0 - \frac{1}{n}\sum_k \tau_{1kk0}A_1^\dagger A_0 + \frac{1}{n}\sum_k \tau_{1kk0}A_k^\dagger A_k A_1^\dagger A_0,\end{aligned} \tag{194}$$

   where

$$\frac{1}{n}\tau_{0001}A_0^\dagger A_1 - \frac{1}{n}\tau_{1110}A_1^\dagger A_0 = \frac{1}{2}\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u}{n}}\right)A_0^\dagger A_1 - \frac{1}{2}\left(\sum_{u=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i u}{n}}\right)A_1^\dagger A_0 = \mathcal{O}(n), \tag{195}$$

   and

$$\begin{aligned}&\frac{1}{n}\sum_k \tau_{0kk1}A_0^\dagger A_1 - \frac{1}{n}\sum_k \tau_{1kk0}A_1^\dagger A_0\\ =&\frac{1}{n}\sum_k\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u k}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v(1-k)}{n}}\right)A_0^\dagger A_1 - \frac{1}{n}\sum_k\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u(k-1)}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i v k}{n}}\right)A_1^\dagger A_0\\ =&\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u}{n}}\right)A_0^\dagger A_1 - \left(\sum_{u=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i u}{n}}\right)A_1^\dagger A_0 = \mathcal{O}(n),\end{aligned} \tag{196}$$

We rewrite the remaining terms as

$$
\begin{aligned}
&-\frac{1}{n}\sum_k \tau_{0kk1} A_k^\dagger A_k A_0^\dagger A_1 + \frac{1}{n}\sum_k \tau_{1kk0} A_k^\dagger A_k A_1^\dagger A_0 \\
=&-\frac{1}{n}\sum_k \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u k}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v(1-k)}{n}}\right) A_k^\dagger A_k A_0^\dagger A_1 + \frac{1}{n}\sum_k \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u(k-1)}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i v k}{n}}\right) A_k^\dagger A_k A_1^\dagger A_0.
\end{aligned}
\tag{197}
$$

4. The fourth group contains terms

$$
\begin{aligned}
&-\frac{1}{n}\sum_j \tau_{j01j} A_j^\dagger A_0 A_1^\dagger A_j + \frac{1}{n}\sum_j \tau_{j10j} A_j^\dagger A_1 A_0^\dagger A_j \\
=&-\frac{1}{n}\tau_{1011} A_1^\dagger A_0 + \frac{1}{n}\sum_j \tau_{j01j} A_j^\dagger A_j A_1^\dagger A_0 + \frac{1}{n}\tau_{0100} A_0^\dagger A_1 - \frac{1}{n}\sum_j \tau_{j10j} A_j^\dagger A_j A_0^\dagger A_1.
\end{aligned}
\tag{198}
$$

Similar to the previous case, we have

$$
-\frac{1}{n}\tau_{1011} A_1^\dagger A_0 + \frac{1}{n}\tau_{0100} A_0^\dagger A_1 = \mathcal{O}(n),
\tag{199}
$$

whereas the remaining terms can be rewritten as

$$
\begin{aligned}
&\frac{1}{n}\sum_j \tau_{j01j} A_j^\dagger A_j A_1^\dagger A_0 - \frac{1}{n}\sum_j \tau_{j10j} A_j^\dagger A_j A_0^\dagger A_1 \\
=&\frac{1}{n}\sum_j \left(\sum_{u=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i u j}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v(j-1)}{n}}\right) A_j^\dagger A_j A_1^\dagger A_0 - \frac{1}{n}\sum_j \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u(1-j)}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v j}{n}}\right) A_j^\dagger A_j A_0^\dagger A_1.
\end{aligned}
\tag{200}
$$

To summarize, the effective commutator $\widetilde{\Pi}\left[\widetilde{T},\widetilde{V}\right]\widetilde{\Pi}$ has action

$$
\begin{aligned}
&\widetilde{\Pi}\left[\widetilde{T},\widetilde{V}\right]\widetilde{\Pi} \\
=&N\left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u}{n}}\right) A_0^\dagger A_1 - N\left(\sum_{u=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i u}{n}}\right) A_1^\dagger A_0 \\
&-\frac{2}{n}\sum_k \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u k}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{\frac{2\pi i v(1-k)}{n}}\right) A_k^\dagger A_k A_0^\dagger A_1 + \frac{2}{n}\sum_k \left(\sum_{u=0}^{\frac{n}{2}-1} e^{\frac{2\pi i u(k-1)}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1} e^{-\frac{2\pi i v k}{n}}\right) A_k^\dagger A_k A_1^\dagger A_0.
\end{aligned}
\tag{201}
$$

We now take the expectation of this operator with respect to the state

$$
|\widetilde{\psi}_\eta\rangle = \frac{|010\cdots 0\overbrace{1\cdots 1}^{\eta-1}\rangle + |100\cdots 0\overbrace{1\cdots 1}^{\eta-1}\rangle}{\sqrt{2}}.
\tag{202}
$$

We have

$$\begin{aligned}
&\langle\widetilde{\psi}_\eta|\left(N\left(\sum_{u=0}^{\frac{n}{2}-1}e^{\frac{2\pi iu}{n}}\right)A_0^\dagger A_1-N\left(\sum_{u=0}^{\frac{n}{2}-1}e^{-\frac{2\pi iu}{n}}\right)A_1^\dagger A_0\right)|\widetilde{\psi}_\eta\rangle\\
=&\eta\langle\widetilde{\psi}_\eta|\left(\frac{2}{1-e^{\frac{2\pi i}{n}}}A_0^\dagger A_1-\frac{2}{1-e^{-\frac{2\pi i}{n}}}A_1^\dagger A_0\right)|\widetilde{\psi}_\eta\rangle+\mathcal{O}(n)\\
=&-\frac{n\eta}{\pi i}\langle\widetilde{\psi}_\eta|\left(A_0^\dagger A_1+A_1^\dagger A_0\right)|\widetilde{\psi}_\eta\rangle+\mathcal{O}(n+\eta)=-\frac{n\eta}{\pi i}+\mathcal{O}(n+\eta).
\end{aligned}\tag{203}$$

On the other hand,

$$\begin{aligned}
&\langle\widetilde{\psi}_\eta|\left(-\frac{1}{n}\sum_k\left(\sum_{u=0}^{\frac{n}{2}-1}e^{\frac{2\pi iuk}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1}e^{\frac{2\pi iv(1-k)}{n}}\right)A_k^\dagger A_kA_0^\dagger A_1+\frac{1}{n}\sum_k\left(\sum_{u=0}^{\frac{n}{2}-1}e^{\frac{2\pi iu(k-1)}{n}}\right)\left(\sum_{v=0}^{\frac{n}{2}-1}e^{-\frac{2\pi ivk}{n}}\right)A_k^\dagger A_kA_1^\dagger A_0\right)|\widetilde{\psi}_\eta\rangle\\
=&\langle\widetilde{\psi}_\eta|\left(-\frac{1}{n}\sum_{k=n-\eta+1}^{n-1}\frac{1-e^{\pi ik}}{1-e^{\frac{2\pi ik}{n}}}\frac{1-e^{\pi i(1-k)}}{1-e^{\frac{2\pi i(1-k)}{n}}}A_0^\dagger A_1+\frac{1}{n}\sum_{k=n-\eta+1}^{n-1}\frac{1-e^{\pi i(k-1)}}{1-e^{\frac{2\pi i(k-1)}{n}}}\frac{1-e^{-\pi ik}}{1-e^{-\frac{2\pi ik}{n}}}A_1^\dagger A_0\right)|\widetilde{\psi}_\eta\rangle+\mathcal{O}(n)=\mathcal{O}(n),
\end{aligned}\tag{204}$$

where the last equality holds since for integer k exactly one of k and $k-1$ is even. We have thus proved

$$\langle\widetilde{\psi}_\eta|\left[\widetilde{T},\widetilde{V}\right]|\widetilde{\psi}_\eta\rangle=-\frac{n\eta}{\pi i}+\mathcal{O}(n+\eta).\tag{205}$$

The above argument can be extended to analyze multilayer nested commutators. Indeed, for initial state

$$|\widetilde{\phi}_\eta\rangle=\frac{|010\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle+i|100\cdots0\overbrace{1\cdots1}^{\eta-1}\rangle}{\sqrt{2}},\tag{206}$$

we have

$$\begin{aligned}
\langle\widetilde{\phi}_\eta|\left[\widetilde{T},\left[\widetilde{T},\widetilde{V}\right]\right]|\widetilde{\phi}_\eta\rangle&=-\frac{n^2\eta}{\pi i}\langle\widetilde{\phi}_\eta|\left(A_0^\dagger A_1-A_1^\dagger A_0\right)|\widetilde{\phi}_\eta\rangle+\mathcal{O}\left(n^2+n\eta\right)\\
&=\frac{n^2\eta}{\pi}+\mathcal{O}\left(n^2+n\eta\right),
\end{aligned}\tag{207}$$

and similar results hold for general nested commutators $[\widetilde{T},\dots[\widetilde{T},\widetilde{V}]]$. This completes the proof of Proposition 12.

D Lower-bounding $\|[V,\dots[V,T]]\|_\eta$

In this appendix, we prove Proposition 13 that lower-bounds the fermionic seminorm $\|[V,\dots[V,T]]\|_\eta$ for the electronic Hamiltonian (125). Recall that we have $H=T+V$ with

$$T=\sum_{j,k=0}^{n-1}A_j^\dagger A_k,\qquad V=\sum_{u,v=0}^{\frac{n}{2}-1}N_uN_v,\tag{208}$$

which implies the commutator

$$[V,T]=\sum_{0\le u\le\frac{n}{2}-1}N_u\Bigg(\sum_{\substack{0\le j\le\frac{n}{2}-1\\ \frac{n}{2}\le k\le n-1}}-\sum_{\substack{\frac{n}{2}\le j\le n-1\\ 0\le k\le\frac{n}{2}-1}}\Bigg)A_j^\dagger A_k+\Bigg(\sum_{\substack{0\le j\le\frac{n}{2}-1\\ \frac{n}{2}\le k\le n-1}}-\sum_{\substack{\frac{n}{2}\le j\le n-1\\ 0\le k\le\frac{n}{2}-1}}\Bigg)A_j^\dagger A_k\sum_{0\le v\le\frac{n}{2}-1}N_v.\tag{209}$$

We will choose the initial state from the two-dimensional subspace spanned by

$$|\psi_0\rangle = |0\overbrace{\underbrace{1\cdots 1}_{\eta-1}0\cdots 0}^{\frac{n}{2}}10\cdots 0\rangle, \qquad |\psi_1\rangle = |1\overbrace{\underbrace{1\cdots 1}_{\eta-1}0\cdots 0}^{\frac{n}{2}}00\cdots 0\rangle. \tag{210}$$

Denoting the projection to this subspace as $\Pi = |\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|$, we have that Π commutes with $\sum_{0\leq u\leq \frac{n}{2}-1} N_u$. Meanwhile,

$$\Pi\Bigg(\sum_{\substack{0\leq j\leq \frac{n}{2}-1\\ \frac{n}{2}\leq k\leq n-1}} - \sum_{\substack{\frac{n}{2}\leq j\leq n-1\\ 0\leq k\leq \frac{n}{2}-1}}\Bigg) A_j^\dagger A_k \Pi = A_0^\dagger A_{\frac{n}{2}} - A_{\frac{n}{2}}^\dagger A_0. \tag{211}$$

This shows that the effective commutator $\Pi[V,T]\Pi$ has the action

$$\Pi[V,T]\Pi = \sum_{0\leq u\leq \frac{n}{2}-1} N_u \left(A_0^\dagger A_{\frac{n}{2}} - A_{\frac{n}{2}}^\dagger A_0\right) + \left(A_0^\dagger A_{\frac{n}{2}} - A_{\frac{n}{2}}^\dagger A_0\right) \sum_{0\leq v\leq \frac{n}{2}-1} N_v. \tag{212}$$

We now take the expectation of this operator with respect to the state

$$|\psi_\eta\rangle = \frac{|0\overbrace{\underbrace{1\cdots 1}_{\eta-1}0\cdots 0}^{\frac{n}{2}}10\cdots 0\rangle + i|1\overbrace{\underbrace{1\cdots 1}_{\eta-1}0\cdots 0}^{\frac{n}{2}}00\cdots 0\rangle}{\sqrt{2}}, \tag{213}$$

which gives

$$\begin{aligned}\langle\psi_\eta|[V,T]|\psi_\eta\rangle &= \langle\psi_\eta|\left(\sum_{0\leq u\leq \frac{n}{2}-1} N_u \left(A_0^\dagger A_{\frac{n}{2}} - A_{\frac{n}{2}}^\dagger A_0\right) + \left(A_0^\dagger A_{\frac{n}{2}} - A_{\frac{n}{2}}^\dagger A_0\right) \sum_{0\leq v\leq \frac{n}{2}-1} N_v\right)|\psi_\eta\rangle \\ &= 2\eta\langle\psi_\eta|\left(A_0^\dagger A_{\frac{n}{2}} - A_{\frac{n}{2}}^\dagger A_0\right)|\psi_\eta\rangle + \mathcal{O}(1) = -2i\eta + \mathcal{O}(1).\end{aligned} \tag{214}$$

This proves the desired scaling for the single-layer commutator. This argument can be extended to analyze multilayer nested commutators. Indeed, for initial state

$$|\phi_\eta\rangle = \frac{|0\overbrace{\underbrace{1\cdots 1}_{\eta-1}0\cdots 0}^{\frac{n}{2}}10\cdots 0\rangle + |1\overbrace{\underbrace{1\cdots 1}_{\eta-1}0\cdots 0}^{\frac{n}{2}}00\cdots 0\rangle}{\sqrt{2}}, \tag{215}$$

we have

$$\begin{aligned}\langle\phi_\eta|[V,[V,T]]|\phi_\eta\rangle &= \langle\phi_\eta|\left(\sum_{1\leq p\leq \eta} N_p\right)^2 \left(A_0^\dagger A_{\frac{n}{2}} + A_{\frac{n}{2}}^\dagger A_0\right)|\phi_\eta\rangle \\ &\quad + 2\langle\phi_\eta|\left(\sum_{1\leq p\leq \eta} N_p\right)\left(A_0^\dagger A_{\frac{n}{2}} + A_{\frac{n}{2}}^\dagger A_0\right)\left(\sum_{1\leq q\leq \eta} N_q\right)|\phi_\eta\rangle \\ &\quad + \langle\phi_\eta|\left(A_0^\dagger A_{\frac{n}{2}} + A_{\frac{n}{2}}^\dagger A_0\right)\left(\sum_{1\leq q\leq \eta} N_q\right)^2|\phi_\eta\rangle \\ &= 4\eta^2 + \mathcal{O}(\eta),\end{aligned} \tag{216}$$

and similar results hold for general nested commutators $[V, \ldots [V, T]]$. This completes the proof of Proposition 13.

References

[1] Dorit Aharonov and Amnon Ta-Shma, *Adiabatic quantum state generation and statistical zero knowledge*, Proceedings of the 35th ACM Symposium on Theory of Computing, pp. 20–29, 2003, arXiv:quant-ph/0301023.

[2] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon, *Simulated quantum computation of molecular energies*, Science **309** (2005), no. 5741, 1704–1707, arXiv:quant-ph/0604193.

[3] Ryan Babbush, Dominic W. Berry, Jarrod R. McClean, and Hartmut Neven, *Quantum simulation of chemistry with sublinear scaling in basis size*, npj Quantum Information **5** (2019), no. 1, 92, arXiv:1807.09802.

[4] Ryan Babbush, Jarrod McClean, Dave Wecker, Alán Aspuru-Guzik, and Nathan Wiebe, *Chemical basis of trotter-suzuki errors in quantum chemistry simulation*, Physical Review A **91** (2015), no. 2, 022311, arXiv:1410.8159.

[5] Ryan Babbush, Nathan Wiebe, Jarrod McClean, James McClain, Hartmut Neven, and Garnet Kin-Lic Chan, *Low-depth quantum simulation of materials*, Physical Review X **8** (2018), 011044, arXiv:1706.00023.

[6] Bela Bauer, Sergey Bravyi, Mario Motta, and Garnet Kin Chan, *Quantum algorithms for quantum chemistry and quantum materials science*, 2020, arXiv:2001.03685.

[7] Dominic W Berry, Graeme Ahokas, Richard Cleve, and Barry C Sanders, *Efficient quantum algorithms for simulating sparse hamiltonians*, Communications in Mathematical Physics **270** (2007), no. 2, 359–371, arXiv:quant-ph/0508139.

[8] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma, *Exponential improvement in precision for simulating sparse Hamiltonians*, Proceedings of the 46th Annual ACM Symposium on Theory of Computing, pp. 283–292, 2014, arXiv:1312.1414.

[9] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma, *Simulating Hamiltonian dynamics with a truncated Taylor series*, Physical Review Letters **114** (2015), no. 9, 090502, arXiv:1412.4687.

[10] Dominic W. Berry, Andrew M. Childs, and Robin Kothari, *Hamiltonian simulation with nearly optimal dependence on all parameters*, Proceedings of the 56th IEEE Symposium on Foundations of Computer Science, pp. 792–809, 2015, arXiv:1501.01715.

[11] Dominic W. Berry, Craig Gidney, Mario Motta, Jarrod R. McClean, and Ryan Babbush, *Qubitization of Arbitrary Basis Quantum Chemistry Leveraging Sparsity and Low Rank Factorization*, Quantum **3** (2019), 208, arXiv:1902.02134.

[12] Chris Cade, Lana Mineh, Ashley Montanaro, and Stasja Stanisic, *Strategies for solving the Fermi-Hubbard model on near-term quantum computers*, 2019, arXiv:1912.06007.

[13] Zhenyu Cai, *Resource estimation for quantum variational simulations of the hubbard model*, Physical Review Applied **14** (2020), 014059, arXiv:1910.02719.

[14] Earl Campbell, *Random compiler for fast Hamiltonian simulation*, Physical Review Letters **123** (2019), no. 7, 070503, arXiv:1811.08017.

[15] Yudong Cao, Jonathan Romero, Jonathan P Olson, Matthias Degroote, Peter D Johnson, Mária Kieferová, Ian D Kivlichan, Tim Menke, Borja Peropadre, Nicolas PD Sawaya, et al., *Quantum chemistry in the age of quantum computing*, Chemical reviews **119** (2019), no. 19, 10856–10915, arXiv:1812.09976.

[16] Chi-Fang Chen, Hsin-Yuan Huang, Richard Kueng, and Joel A. Tropp, *Quantum simulation via randomized product formulas: Low gate complexity with accuracy guarantees*, 2020, arXiv:2008.11751.

[17] Andrew M. Childs, Aaron Ostrander, and Yuan Su, *Faster quantum simulation by randomization*, Quantum **3** (2019), 182, arXiv:1805.08385.

[18] Andrew M. Childs and Yuan Su, *Nearly optimal lattice simulation by product formulas*, Physical Review Letters **123** (2019), 050503, arXiv:1901.00564.

[19] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu, *A theory of Trotter error*, 2019, arXiv:1912.08854.

[20] Laura Clinton, Johannes Bausch, and Toby Cubitt, *Hamiltonian simulation algorithms for near-term quantum hardware*, 2020, arXiv:2003.06886.

[21] Burak Şahinoğlu and Rolando D. Somma, *Hamiltonian simulation in the low energy subspace*, 2020, arXiv:2006.02660.

[22] Richard P. Feynman, *Simulating physics with computers*, International Journal of Theoretical Physics **21** (1982), no. 6-7, 467–488.

[23] Jeongwan Haah, Matthew B. Hastings, Robin Kothari, and Guang Hao Low, *Quantum algorithm for simulating real time evolution of lattice Hamiltonians*, Proceedings of the 59th IEEE Symposium on Foundations of Computer Science, pp. 350–360, 2018, arXiv:1801.03922.

[24] Stuart Hadfield and Anargyros Papageorgiou, *Divide and conquer approach to quantum Hamiltonian simulation*, New Journal of Physics **20** (2018), no. 4, 043003.

[25] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen, *Molecular electronic-structure theory*, John Wiley & Sons, 2014.

[26] Ian D Kivlichan, Craig Gidney, Dominic W Berry, Nathan Wiebe, Jarrod McClean, Wei Sun, Zhang Jiang, Nicholas Rubin, Austin Fowler, Alán Aspuru-Guzik, et al., *Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization*, Quantum **4** (2020), 296, arXiv:1902.10673.

[27] Ian D Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush, *Quantum simulation of electronic structure with linear depth and connectivity*, Physical Review Letters **120** (2018), no. 11, 110501, arXiv:1711.04789.

[28] J. P. F. LeBlanc, Andrey E. Antipov, Federico Becca, Ireneusz W. Bulik, Garnet Kin-Lic Chan, Chia-Min Chung, Youjin Deng, Michel Ferrero, Thomas M. Henderson, Carlos A. Jiménez-Hoyos, E. Kozik, Xuan-Wen Liu, Andrew J. Millis, N. V. Prokof'ev, Mingpu Qin, Gustavo E. Scuseria, Hao Shi, B. V. Svistunov, Luca F. Tocchio, I. S. Tupitsyn, Steven R. White, Shiwei Zhang, Bo-Xiao Zheng, Zhenyue Zhu, and Emanuel Gull, *Solutions of the two-dimensional hubbard model: Benchmarks and results from a wide range of numerical algorithms*, Physical Review X **5** (2015), 041041, arXiv:1505.02290.

[29] Seth Lloyd, *Universal quantum simulators*, Science (1996), 1073–1078.

[30] Guang Hao Low and Isaac L. Chuang, *Hamiltonian simulation by uniform spectral amplification*, 2017, arXiv:1707.05391.

[31] Guang Hao Low and Isaac L. Chuang, *Optimal Hamiltonian simulation by quantum signal processing*, Physical Review Letters **118** (2017), 010501, arXiv:1606.02685.

[32] Guang Hao Low and Isaac L. Chuang, *Hamiltonian simulation by qubitization*, Quantum **3** (2019), 163, arXiv:1610.06546.

[33] Guang Hao Low, Vadym Kliuchnikov, and Nathan Wiebe, *Well-conditioned multiproduct Hamiltonian simulation*, 2019, arXiv:1907.11679.

[34] Guang Hao Low and Nathan Wiebe, *Hamiltonian simulation in the interaction picture*, 2018, arXiv:1805.00675.

[35] Sam McArdle, Suguru Endo, Alan Aspuru-Guzik, Simon C Benjamin, and Xiao Yuan, *Quantum computational chemistry*, Reviews of Modern Physics **92** (2020), no. 1, 015003, arXiv:1808.10402.

[36] Richard Meister, Simon C Benjamin, and Earl T Campbell, *Tailoring term truncations for electronic structure calculations using a linear combination of unitaries*, 2020, arXiv:2007.11624.

[37] Peter Otte, *Boundedness properties of fermionic operators*, Journal of Mathematical Physics **51** (2010), no. 8, 083503, arXiv:0911.4438.

[38] Yingkai Ouyang, David R White, and Earl T Campbell, *Compilation by stochastic hamiltonian sparsification*, Quantum **4** (2020), 235, arXiv:1910.06255.

[39] Tianyi Peng, Aram W. Harrow, Maris Ozols, and Xiaodi Wu, *Simulating large quantum circuits on a small quantum computer*, Physical Review Letters **125** (2020), 150504, arXiv:1904.00102.

[40] David Poulin, Matthew B. Hastings, Dave Wecker, Nathan Wiebe, Andrew C. Doherty, and Matthias Troyer, *The Trotter step size required for accurate quantum simulation of quantum chemistry*, Quantum Information and Computation **15** (2015), no. 5-6, 361–384, arXiv:1406.4920.

[41] Rolando D. Somma, *Quantum simulations of one dimensional quantum systems*, 2015, arXiv:1503.06319.

[42] Rolando D. Somma, *A Trotter-Suzuki approximation for Lie groups with applications to Hamiltonian simulation*, Journal of Mathematical Physics **57** (2016), 062202, arXiv:1512.03416.

[43] Masuo Suzuki, *Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics*, Journal of Mathematical Physics **26** (1985), no. 4, 601–612.

[44] Masuo Suzuki, *Fractal decomposition of exponential operators with applications to many-body theories and monte carlo simulations*, Physics Letters A **146** (1990), no. 6, 319–323.

[45] Minh C. Tran, Andrew Y. Guo, Yuan Su, James R. Garrison, Zachary Eldredge, Michael Foss-Feig, Andrew M. Childs, and Alexey V. Gorshkov, *Locality and digital quantum simulation of power-law interactions*, Physical Review X **9** (2019), 031006, arXiv:1808.05225.

[46] Minh C. Tran, Yuan Su, Daniel Carney, and Jacob M. Taylor, *Faster digital quantum simulation by symmetry protection*, 2020, arXiv:2006.16248.

[47] Dave Wecker, Matthew B Hastings, Nathan Wiebe, Bryan K Clark, Chetan Nayak, and Matthias Troyer, *Solving strongly correlated electron models on a quantum computer*, Physical Review A **92** (2015), no. 6, 062318, arXiv:1506.05135.

[48] Shenglong Xu, Leonard Susskind, Yuan Su, and Brian Swingle, *A sparse model of quantum holography*, 2020, arXiv:2008.02303.

[49] Bo-Xiao Zheng, Chia-Min Chung, Philippe Corboz, Georg Ehlers, Ming-Pu Qin, Reinhard M Noack, Hao Shi, Steven R White, Shiwei Zhang, and Garnet Kin-Lic Chan, *Stripe order in the underdoped region of the two-dimensional Hubbard model*, Science **358** (2017), no. 6367, 1155–1160.