

Fault-Tolerant Quantum Simulations of Chemistry in First Quantization

Yuan Su^{1,2,*}, Dominic W. Berry^{3,†}, Nathan Wiebe^{4,5,1,‡}, Nicholas Rubin¹, and Ryan Babbush^{1,§}


¹Google Quantum AI, Venice, California 90291, USA

²Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA

³Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia

⁴Department of Computer Science, University of Toronto, Ontario M5S 2E4, Canada

⁵Pacific Northwest National Laboratory, Richland, Washington 99354, USA

 (Received 31 May 2021; accepted 4 October 2021; published 11 November 2021; corrected 21 September 2022)

Quantum simulations of chemistry in first quantization offer some important advantages over approaches in second quantization including faster convergence to the continuum limit and the opportunity for practical simulations outside of the Born-Oppenheimer approximation. However, since all prior work on quantum simulation of chemistry in first quantization has been limited to asymptotic analysis, it has been impossible to directly compare the resources required for these approaches to those required for the more commonly studied algorithms in second quantization. Here, we compile, optimize, and analyze the finite resources required to implement two first quantized quantum algorithms for chemistry from Babbush *et al.* [Npj Quantum Inf. 5, 92 (2019)] that realize block encodings for the qubitization and interaction-picture frameworks of Low *et al.* [Quantum 3, 163 (2019), arXiv:1805.00675 (2018)]. The two algorithms we study enable simulation with gate complexities of $\tilde{O}(\eta^{8/3}N^{1/3}t + \eta^{4/3}N^{2/3}t)$ and $\tilde{O}(\eta^{8/3}N^{1/3}t)$ where η is the number of electrons, N is the number of plane-wave basis functions, and t is the duration of time evolution (t is linearly inverse to target precision when the goal is to estimate energies). In addition to providing the first explicit circuits and constant factors for any first quantized simulation, and then introducing improvements, which reduce circuit complexity by about a thousandfold over naive implementations for modest sized systems, we also describe new algorithms that asymptotically achieve the same scaling in a real-space representation. Finally, we assess the resources required to simulate various molecules and materials and conclude that the qubitized algorithm will often be more practical than the interaction-picture algorithm. We demonstrate that our qubitization algorithm often requires much less surface-code space-time volume for simulating millions of plane waves than the best second quantized algorithms require for simulating hundreds of Gaussian orbitals.

DOI: [10.1103/PRXQuantum.2.040332](https://doi.org/10.1103/PRXQuantum.2.040332)

I. INTRODUCTION

The last several years have seen monumental developments in quantum algorithms, resulting in significant progress towards the goal of showing that a modest sized quantum computer can provide a decisive advantage for scientifically relevant problems. The expectation that quantum simulation might enable an exponential speedup

for physical problems stretches back to Feynman's original proposal for quantum computers [1,2] and was first concretely proposed for solving the the central problem of quantum chemistry by Aspuru-Guzik *et al.* [3]. Since then, progress has accelerated on multiple fronts: improvements in general purpose simulation frameworks [4–14], more efficient algorithms for quantum chemistry [15–28], more advantageous-to-quantum-simulate representations of molecular Hamiltonians [29–37], and an improved understanding of the problems that genuinely require a quantum computer versus those that can be classically computed [38,39].

However, a problem lurks within this discussion that has not received as much attention as it deserves. Specifically, most existing methods for simulating quantum chemistry (especially those leveraging simple basis functions like plane waves) become impractical as we scale to the continuum limit. This is because the second quantized

*buptsuyuan@gmail.com

†dominic.berry@mq.edu.au

‡nathanwiebe@gmail.com

§ryanbabbush@gmail.com

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

simulations that are deployed in the vast majority of quantum-computing algorithms for chemistry require a number of qubits that scales linearly with the number of spin-orbital basis functions. Dating back to early work by Kassal *et al.* [40], first quantized simulations of chemistry have been proposed as a method to ameliorate this problem. The central idea behind such simulations is to track the basis state that each particle is in rather than storing the occupancy of each basis state. This has two significant benefits: the number of qubits needed to represent the state scales logarithmically with the number of basis functions (although linearly with the number of particles) and also, such simulations are easy to adapt to cases where the entanglement between the electronic and nuclear subsystems is non-negligible. Despite the promise of these methods, no prior work investigating first quantized approaches for chemistry has managed to rigorously estimate all of the constant factors associated with the realization of these algorithms [41].

Our aim in this paper is to address this shortcoming by providing a full costing of the two leading quantum-simulation algorithms for chemistry in a first quantized representation, both first described by Babbush *et al.* [42]. These algorithms comprise asymptotically efficient schemes for realizing the oracles required by the qubitization and interaction-picture frameworks of Low *et al.* [4,5]. Here, we compile and optimize these algorithms within an error-correctable gate set. We develop new techniques to reduce the complexity, resulting in about 3 orders of magnitude improvement over a naive implementation for even modest sized systems. Our detailed cost analysis reveals that despite the improved scaling of the interaction-picture-based method, the qubitization-based method often proves to be more practical. We conduct numerical experiments to characterize the overhead of both approaches and find that they are highly competitive with second quantized approaches, often requiring significantly fewer resources to reach comparable accuracy.

The layout of this paper is as follows. In the remainder of this section, we provide the formalism and context needed to understand the two most efficient schemes for simulating the molecular Hamiltonian in first quantization. Section II develops and analyzes our optimized implementation of the qubitization-based algorithm in first quantization. Section III contains a similar analysis for the interaction-picture-based algorithm. Finally, Sec. IV numerically compares these algorithms (to each other and to other methods in the literature), estimating the resources required to simulate real systems in a first quantized basis.

A. Background on representing molecular Hamiltonians in a first quantized plane-wave basis

The accurate simulation of quantum systems was Feynman’s original vision for quantum computing [1]. There

is perhaps no more natural application of this idea than the simulation of atoms and molecules, governed by the interactions of electrons and nuclei. Such systems give rise to the properties of most matter, determining everything from the rates of chemical reactions to the conductivity of materials. In the nonrelativistic case, the dynamics of these systems are governed by the Coulomb Hamiltonian:

$$H = \underbrace{-\sum_{i=1}^{\eta} \frac{\nabla_i^2}{2}}_T - \underbrace{\sum_{i=1}^{\eta} \sum_{\ell=1}^L \frac{\zeta_{\ell}}{\|R_{\ell} - r_i\|}}_U + \underbrace{\sum_{i \neq j=1}^{\eta} \frac{1}{2 \|r_i - r_j\|}}_V - \underbrace{\sum_{\ell=1}^L \frac{\nabla_{\ell}^2}{2m_{\ell}}}_{T_{\text{nuclei}}} + \underbrace{\sum_{\ell \neq \kappa=1}^L \frac{\zeta_{\ell} \zeta_{\kappa}}{2 \|R_{\ell} - R_{\kappa}\|}}_{V_{\text{nuclei}}}. \quad (1)$$

Here we use atomic units where \hbar as well as the mass and charge of the electron are unity, and $\|\cdot\|$ denotes the 2-norm. In the above expression i, j represent electronic degrees of freedom and ℓ, κ represent nuclear degrees of freedom; thus, r_i represent the positions of electrons whereas R_{ℓ} represent the positions of nuclei, m_{ℓ} the atomic masses of nuclei, and ζ_{ℓ} the atomic numbers of nuclei. Throughout this work we use η to denote the number of electrons in our simulation and L to denote the number of nuclei in our simulation. We keep a catalog of the different symbols used throughout this paper in Appendix A.

The methods we analyze and develop in this work are useful for simulations of the dynamics of Eq. (1) (treating both nuclear and electronic degrees of freedom as fully quantum) with only minor modifications. But in order to simplify and reduce our analysis, we primarily focus on simulating Eq. (1) under the Born-Oppenheimer approximation: i.e., $H(R) = T + U(R) + V + C(R)$ where $C(R)$ is a constant given by the Coulomb repulsion of the point charges of “classical” nuclei with locations R_{ℓ} . Thus, the Born-Oppenheimer approximation assumes that one can decouple nuclear and electronic degrees of freedom, treating the former as essentially “classical” [43,44]. This is appropriate for many systems near room temperature because the electronic degrees of freedom are thousands of times faster than the nuclear degrees of freedom (due to the disparity in mass between electrons and nuclei) and hence, electrons often relax nearly instantaneously to their quantum ground states whereas the nuclei are often heavy enough to be modeled as classical objects responding to forces arising from the electronic interactions. However, the Born-Oppenheimer approximation is known to break down in some circumstances, including at low temperatures or when chemical bonds involve low atomic mass nuclei (e.g., due to the tunneling of hydrogen atoms); hence, methods for performing simulations of Eq. (1),

referred to as “non-Born-Oppenheimer dynamics” are also of great interest [45–48].

The frequently studied “electronic structure problem” is to solve for the ground-state energies of electrons interacting in the external potential of nuclei under the Born-Oppenheimer approximation [i.e., to solve for the ground-state energies of $H(R)$]. As a function of the nuclear coordinates R , these energies define an energy surface and those energy surfaces can be extremely helpful for understanding the mechanisms and dynamics of chemical reactions as well as material properties. For example, energy surfaces are commonly sought after so that one may simulate the “classical” dynamics of nuclei moving on these “quantum” energy surfaces by integrating Newton’s equations of motion for the nuclei. When the trajectories of these classical nuclei responding to quantum electronic energy surfaces are coupled to a finite temperature bath so that the system may sample the canonical or grand canonical ensemble (rather than just Hamiltonian dynamics), these simulations are referred to as “*ab initio* molecular dynamics” [49–51]. While the electronic energy surfaces (referred to in the molecular dynamics community as “force-fields”) are often determined empirically or by crude calculations that introduce large errors into the end result, molecular dynamics simulations are among the most common calculations in scientific computing and can provide invaluable insights into the thermodynamic properties of real materials and chemical reactions.

In order to simulate molecular systems on a computer one must discretize them in some fashion. For reasons discussed in more depth in Appendix B, most molecular modeling employs a Galerkin representation, which involves projecting the system onto some well behaved set of basis functions $\{\phi_p(r)\}$. The matrix elements of the Hamiltonian operators are then given by the following integrals over these basis functions:

$$T_{pq}^{(m)} = \int dr \phi_p^*(r) \left(-\frac{\nabla^2}{2m} \right) \phi_q(r), \quad (2)$$

$$U_{pq} = \sum_{\ell=1}^L \int dr \phi_p^*(r) \left(\frac{\zeta_{\ell}}{\|r - R_{\ell}\|} \right) \phi_q(r), \quad (3)$$

$$V_{pqrs}^{(\alpha,\beta)} = \int dr_1 dr_2 \phi_p^*(r_1) \phi_q^*(r_2) \left(\frac{\alpha\beta}{\|r_1 - r_2\|} \right) \times \phi_r(r_2) \phi_s(r_1). \quad (4)$$

We note that simple grid representations (as opposed to discrete-variable-representation-like grids, e.g., those discussed for quantum simulation in Ref. [29]) are incompatible with Galerkin discretizations. For example, if one associates a grid with compact basis functions having disjoint support (e.g., delta functions or step functions) then evaluation of the above integrals would give both

potential operators and kinetic operators that are simultaneously diagonal. For more discussion on real-space representations see Appendix K.

A key consideration for Galerkin discretizations is the compactness of the basis set for the states of interest (often the ground state); i.e., how many basis functions are required to converge the discretization of states of interest to within some threshold ϵ of the continuum limit. Throughout this paper we use N to refer to the number of basis functions used to discretize the system. For most reasonable choices of basis sets, the asymptotic behavior is that $\epsilon = \mathcal{O}(1/N)$ [52–58]. However, the constant factors in this scaling can differ considerably depending on the context of the simulation and the choice of basis. The primary factor limiting the convergence of these discretizations is the resolution of “cusps” that are known to be features of eigenstates of molecular Hamiltonians [59]. Wave-function cusps appear at all points in space where particles overlap. Away from these cusps, the wave function is generally smooth and easier to converge. When working within the Born-Oppenheimer approximation, the nuclei have fixed position and there is a clear cusp in the electronic charge density at the location of those nuclei. The electronic density tends to be sharply peaked at these nuclei and then falls off exponentially away from the nuclei. To capture this structure, one of the most common types of basis sets used for electronic structure simulations are nuclei centered Gaussian orbitals. Often, one uses these “primitive” orbitals to construct even more compact numerical orbitals; e.g., by using the primitive orbitals to discretize the diagonalization of an approximate single-particle description of the physics. When the single-particle description is a mean-field model of molecule (usually Hartree-Fock) these numerically optimized orbitals are referred to as “molecular orbitals.”

Aside from Gaussian orbitals, the other most commonly used class of basis functions are plane waves. Plane waves are the type of basis function that we focus on in this paper. Plane-wave basis functions are eigenstates of the linear momentum operator (and thus, also the operator T), expressed as

$$\phi_p(r) = \sqrt{\frac{1}{\Omega}} e^{-i k_p \cdot r}, \quad (5)$$

where r is a position vector in real space, Ω is the computational cell volume, and k_p is a reciprocal lattice vector in three dimensions. In this paper we focus on methods to perform simulations with plane waves defined over a cubic reciprocal lattice so that

$$k_p = \frac{2\pi p}{\Omega^{1/3}} \quad p \in G \quad G = \left[-\frac{N^{1/3} - 1}{2}, \frac{N^{1/3} - 1}{2} \right]^3 \subset \mathbb{Z}^3. \quad (6)$$

We note that while this reciprocal lattice is appropriate for simulating nonperiodic systems or those with cubic periodicity, to simulate systems with other crystal symmetries one may need to generalize the above equation to the case of nonorthogonal Bravais vectors (we leave this to future work) [60].

Plane waves are a proper basis (unlike a grid) but, like a grid, provide an unbiased discretization of space that is especially well suited to representing dynamics. Due to their regularity, plane waves also give very systematic convergence to the continuum limit, allowing one to accurately estimate properties in the continuum by extrapolating from a series of calculations with progressively more plane waves (the extrapolation results in higher accuracy predictions than even the largest simulation used in the extrapolation). But from a quantum algorithms perspective, perhaps the most useful property of plane waves is that the integrals defining their Galerkin representation have a convenient closed form (arising from the Fourier transform of the Coulomb kernel) [60]:

$$T_{pq}^{(m)} = \delta_{pq} \frac{\|k_p\|^2}{2m}, \quad U_{pq} = \frac{4\pi}{\Omega} \sum_{\ell=1}^L \zeta_{\ell} \frac{e^{ik_{q-p} \cdot R_{\ell}}}{\|k_{p-q}\|^2},$$

$$V_{pqrs}^{(\alpha,\beta)} = \delta_{p-s,r-q} \frac{4\pi}{\Omega} \frac{\alpha\beta}{\|k_v\|^2}, \quad (7)$$

where $v = p - s = r - q \neq 0$ in the last expression. As we see, this closed form enables some especially efficient quantum algorithms.

The structure of the plane-wave basis can be used to give quantum algorithms with lower asymptotic gate complexity compared to algorithms with Gaussian orbitals. One reason for this is that most methods for simulating second quantized Hamiltonians tend to have some cost that scales with the number of terms in the Hamiltonian (or specifically, the number of terms in the two-body operator since that is the difficult part). A second quantized plane-wave Hamiltonian has $\mathcal{O}(N^3)$ terms in the two-body operator [less than the $\mathcal{O}(N^4)$ terms of a molecular orbital representation by a factor of N due to conservation of linear momentum] but only $\mathcal{O}(N^2)$ terms in the dual basis obtained by Fourier transforming the basis. This argument was first made in the context of second quantized quantum simulations in Ref. [29]. Today, the most efficient second quantized algorithm for simulating plane-wave electronic structure Hamiltonians with N plane waves and η electrons has Toffoli complexity $(N^{5/3}/\eta^{2/3} + N^{4/3}\eta^{2/3})N^{o(1)}$ with space complexity $\mathcal{O}(N \log N)$ [28], where this complexity assumes that particle density is held fixed as the system size grows. We compare the scaling of all prior quantum algorithms for plane-wave electronic structure in more detail in Table I. By contrast, the best second quantized algorithms for simulating arbitrary basis (e.g., molecular orbital) electronic structure Hamiltonians with N orbitals have Toffoli complexity that is roughly $\tilde{\mathcal{O}}(N^3)$ with space complexity $\mathcal{O}(N \log N)$ [27,61]. Thus, the best plane-wave algorithms have better asymptotic scaling than the best arbitrary basis algorithms. However, while the basis set discretization error of both plane waves and Gaussians is asymptotically refined as $\epsilon = \mathcal{O}(1/N)$, in practice one often requires far fewer Gaussian orbitals than plane

TABLE I. Best quantum algorithms for phase estimating chemistry in a plane-wave basis. N is number of basis functions, $\eta < N$ is the number of electrons, Ω is the computational cell volume, and ϵ is target precision. In some of these papers it is assumed that $\Omega \propto N$ or that $\Omega \propto \eta$, but here we report the complexities without any such assumptions. Unlike other entries here, the scaling of the Kivlichan *et al.* [24] algorithm is empirically observed for specific systems (with some systems scaling somewhat better and some systems scaling somewhat worse) as opposed to rigorous upper bounds. The $\mathcal{O}(N \log N)$ space complexity of Refs. [5,11,28] results from a Fourier-transform-based method of computing the potential operator discussed in Ref. [5]. If those papers were instead to use the Trotter steps introduced in Ref. [23] or Ref. [24], the algorithms would have $N + \mathcal{O}(\log N)$ space complexity but gate complexity that is worse by approximately a factor of N . Algorithms with $\mathcal{O}(N)$ space complexity are likely impractical for any real problems in chemistry when using plane waves due to the large number of plane waves required to reach chemical accuracy; thus, $\mathcal{O}(N \log N)$ space complexity is even less viable. We note that Appendix K introduces new algorithms defined in real space that match the scaling of those from Babbush *et al.* [42].

Year	Reference	Primary innovation	Logical qubits	Toffoli/T complexity
2017	Babbush <i>et al.</i> [29]	Using plane waves with Trotter	$N + \mathcal{O}(\log N)$	$\tilde{\mathcal{O}}[\eta^2 N^{17/6} \sqrt{1 + \eta \Omega^{1/3}/N^{1/3}} / (\Omega^{5/6} \epsilon^{3/2})]$
2017	Babbush <i>et al.</i> [29]	Using plane waves with LCU	$N + \mathcal{O}(\log N)$	$\tilde{\mathcal{O}}[(N^4/\Omega^{1/3} + N^{11/3}/\Omega^{2/3})/\epsilon]$
2018	Babbush <i>et al.</i> [20]	Linear scaling quantum walks	$N + \mathcal{O}(\log N)$	$\tilde{\mathcal{O}}[(N^{10/3}/\Omega^{1/3} + N^{8/3}/\Omega^{2/3})/\epsilon]$
2018	Low <i>et al.</i> [5]	Using the interaction pic.	$\mathcal{O}(N \log N)$	$\tilde{\mathcal{O}}[N^{8/3}/(\Omega^{2/3} \epsilon)]$
2018	Babbush <i>et al.</i> [42]	First quantized qubitization	$\mathcal{O}(\eta \log N)$	$\tilde{\mathcal{O}}[(\eta^3 N^{1/3}/\Omega^{1/3} + \eta^2 N^{2/3}/\Omega^{2/3})/\epsilon]$
2018	Babbush <i>et al.</i> [42]	First quantized interaction pic.	$\mathcal{O}(\eta \log N)$	$\tilde{\mathcal{O}}[\eta^3 N^{1/3}/(\Omega^{1/3} \epsilon)]$
2019	Kivlichan <i>et al.</i> [24]	Better Trotter steps	$N + \mathcal{O}(\log N)$	$\tilde{\mathcal{O}}[N^3/(\Omega^{2/3} \epsilon^{3/2})]$
2019	Childs <i>et al.</i> [11]	Tighter Trotter bounds	$\mathcal{O}(N \log N)$	$N^{7/3+o(1)}/(\Omega^{1/3} \epsilon^{1+o(1)})$
2020	Su <i>et al.</i> [28]	Yet tighter Trotter bounds	$\mathcal{O}(N \log N)$	$(\eta/\Omega^{1/3} + N^{1/3}/\Omega^{2/3})N^{4/3+o(1)}/\epsilon^{1+o(1)}$

waves in order to achieve “chemically accurate” electronic structure simulations [62]. For some systems, one could easily require thousands of times more plane waves to reach chemical accuracy.

Plane waves are especially well suited to treat regular solid-state systems (like crystals) since the basis is naturally periodic. Furthermore, there are some special systems, such as the uniform electron gas, for which plane waves are essentially the most compact basis. Thus, several papers [20,24,63] suggest that using a plane-wave basis in second quantization one can encounter interesting and challenging instances of electronic structure that require fewer resources to solve than any second quantized Gaussian orbital representation. However, for most chemical calculations on a quantum computer, a second quantized plane-wave representation is not realistic to use and this is why the vast majority of work on quantum-computing quantum chemistry focuses on using Gaussian orbitals. Even though second quantized plane-wave representations have asymptotically lower gate complexity than second quantized Gaussian basis algorithms, the space complexity required for accurate calculations using second quantized plane waves is daunting. For example, if a calculation requires 100 Gaussian orbitals but 100 000 plane waves that means in second quantization the former would require hundreds of logical system qubits (reasonable) whereas the latter would require hundreds of thousands of logical system qubits (an extravagant cost). Thus, for chemistry calculations in second quantization, Gaussian orbitals are more practical than plane waves.

However, in first quantization the story is rather different. Because the space complexity of first quantized representations is only $\mathcal{O}(\eta \log N)$ one can easily simulate systems with a very large number of plane waves without the number of qubits becoming too extravagant. As we later discuss, the gate complexity of first quantized plane-wave algorithms can also have sublinear scaling in N ; for example, the algorithm of Ref. [42] has gate complexity $\tilde{\mathcal{O}}(\eta^{8/3} N^{1/3})$. Sublinear scaling in N would be impossible in second quantization as each of N qubits must be acted upon at least once. In principle one can also perform simulations of Gaussian orbitals in first quantization but due to the lack of structure in the integrals this would likely give a relatively inefficient gate complexity. The work of Ref. [64] does something that is very close to this; in that work, the authors simulate a fixed particle-number manifold of the second quantized molecular orbital Hamiltonian (this is known in chemistry as the “configuration interaction” representation). Their approach has space complexity $\mathcal{O}(\eta \log N)$ but it is technically a second quantized approach in a fixed particle manifold rather than a first quantized approach because the symmetries are still encoded by how the Hamiltonian operator acts on the state rather than the algorithm requiring antisymmetrized initial states. However, the gate complexity of that approach is

$\tilde{\mathcal{O}}(\eta^2 N^3)$, which is still relatively high in N , meaning that one cannot refine the basis to extremely large sizes the way that one can with first quantized plane-wave approaches.

We thus argue that plane waves are especially well suited for use within first quantized quantum simulations whereas Gaussian orbitals are especially well suited for use within second quantized quantum simulations. The relative merits of these two representations is a fairly subtle issue because they have different properties. One requires more plane waves than Gaussian orbitals to converge a calculation but the first quantized plane-wave algorithm asymptotically scales better for essentially all values of η and N . That is because $\eta < N$, so the $\tilde{\mathcal{O}}(\eta^{8/3} N^{1/3})$ Toffoli complexity of the best first quantized plane algorithm [42] reduces to the roughly $\tilde{\mathcal{O}}(N^3)$ Toffoli complexity of the best second quantized arbitrary basis algorithm [27] in the worst case that $\eta = \Theta(N)$. There is also a substantial difference in how the algorithms approach the continuum limit. For example, the Toffoli complexity to achieve basis error ϵ scales as $\tilde{\mathcal{O}}(1/\epsilon^3)$ with $\tilde{\mathcal{O}}(1/\epsilon)$ space complexity for the arbitrary basis algorithm but the Toffoli complexity is $\tilde{\mathcal{O}}(1/\epsilon^{1/3})$ with space complexity $\mathcal{O}[\log(1/\epsilon)]$ for the plane-wave algorithm. This enormous difference—a polynomial scaling difference by a power of 9—suggests a strong preference for the first quantized plane-wave algorithm when targeting high precision [65]. Furthermore, the plane-wave representation is the clear preference when performing simulations beyond the Born-Oppenheimer approximation. The reason for this is because the advantage of Gaussian basis sets lies in their ability to resolve cusps in the wave function centered on the nuclei. But outside of the Born-Oppenheimer approximation the nuclei are treated explicitly and thus one cannot center Gaussians on the nuclei because the nuclei do not have a fixed position that is known in advance. Thus, for simulations of nuclear quantum dynamics plane waves are going to be the better choice.

By plugging the plane-wave integrals given in Eq. (7) into the definitions of the Born-Oppenheimer and non-Born-Oppenheimer (nBO) first quantized Galerkin discretizations discussed in Appendix B, we obtain

$$H_{\text{BO}} = T + U + V + \frac{1}{2} \sum_{\ell \neq \kappa=1}^L \frac{\zeta_\ell \zeta_\kappa}{\|R_\ell - R_\kappa\|}, \quad (8)$$

$$H_{\text{nBO}} = T + T_{\text{nuc}} + U_{\text{nBO}} + V + V_{\text{nuc}}, \quad (9)$$

$$T = \sum_{i=1}^{\eta} \sum_{p \in G} \frac{\|k_p\|^2}{2} |p\rangle \langle p|_i$$

$$T_{\text{nuc}} = \sum_{\ell=\eta+1}^{L+\eta} \sum_{p \in G} \frac{\|k_p\|^2}{2 m_\ell} |p\rangle \langle p|_\ell, \quad (10)$$

$$U = -\frac{4\pi}{\Omega} \sum_{\ell=1}^L \sum_{i=1}^{\eta} \sum_{\substack{p,q \in G \\ p \neq q}} \left(\zeta_{\ell} \frac{e^{ik_{q-p} \cdot R_{\ell}}}{\|k_{p-q}\|^2} \right) |p\rangle\langle q|_i, \quad (11)$$

$$U_{\text{nBO}} = -\frac{4\pi}{\Omega} \sum_{i=1}^{\eta} \sum_{\ell=\eta+1}^{L+\eta} \sum_{p,q \in G} \sum_{\substack{v \in G_0 \\ (p+v) \in G \\ (q-v) \in G}} \frac{1}{\|k_v\|^2} \cdot |p+v\rangle\langle p|_i |q-v\rangle\langle q|_{\ell}, \quad (12)$$

$$V = \frac{2\pi}{\Omega} \sum_{i \neq j=1}^{\eta} \sum_{p,q \in G} \sum_{\substack{v \in G_0 \\ (p+v) \in G \\ (q-v) \in G}} \frac{1}{\|k_v\|^2} \cdot |p+v\rangle\langle p|_i |q-v\rangle\langle q|_j, \quad (13)$$

$$V_{\text{nuc}} = \frac{2\pi}{\Omega} \sum_{\ell \neq \kappa=\eta+1}^{L+\eta} \sum_{p,q \in G} \sum_{\substack{v \in G_0 \\ (p+v) \in G \\ (q-v) \in G}} \frac{\zeta_{\ell} \zeta_{\kappa}}{\|k_v\|^2} \cdot |p+v\rangle\langle p|_{\ell} |q-v\rangle\langle q|_{\kappa}, \quad (14)$$

where $G_0 = [-N^{1/3}, N^{1/3}]^3 \subset \mathbb{Z}^3 \setminus \{(0, 0, 0)\}$ is the set of the difference of frequencies from G excluding the singular zero mode, and $|p\rangle\langle q|_i$ is a shorthand for $I_1 \otimes \cdots \otimes |p\rangle\langle q|_i \otimes \cdots \otimes I_{\eta}$ where I_i is the identity operator on the $\log N$ bit register associated with electron i . We store our wave function by having a computational basis that encodes configurations of the electrons in N basis functions such that a configuration is specified as $|\phi_1 \phi_2 \cdots \phi_{\eta}\rangle$ where each ϕ encodes the index of an occupied basis function.

The representation of the Hamiltonian we show discretized above is in ‘‘first quantization.’’ The distinction between ‘‘first’’ and ‘‘second’’ quantization is fundamentally about how one deals with the symmetries of electrons and nuclei. Electrons (and some nuclei) are fermions. Some nuclei are bosons. Whereas identical fermions must be antisymmetric with respect to particle label exchange, identical bosons must be symmetric with respect to particle label exchange. For example, if our state corresponds to a valid first quantized wave function for identical particles then the system must have the property that

$$|\psi\rangle = \sum_{\phi_p \in \{\phi\}} a_{\phi_1 \dots \phi_{\eta}} |\phi_1 \cdots \phi_i \cdots \phi_j \cdots \phi_{\eta}\rangle = (-1)^{\pi} \times \sum_{\phi_p \in \{\phi\}} a_{\phi_1 \dots \phi_{\eta}} |\phi_1 \cdots \phi_j \cdots \phi_i \cdots \phi_{\eta}\rangle, \quad (15)$$

where $\pi = 1$ if fermions and $\pi = 0$ if bosons and $a_{\phi_1 \dots \phi_{\eta}}$ are the amplitudes of the associated computational basis state. No special symmetries need to exist between distinguishable particles. The work of Ref. [21] shows that

one can symmetrize (or antisymmetrize) any set of η registers, which each index N basis functions with a quantum algorithm that has gate complexity $\mathcal{O}(\eta \log \eta \log N)$. Once symmetrized (or antisymmetrized) any simulation under the Hamiltonian will maintain that property as a consequence of Eq. (1) commuting with the particle permutation operator between any identical particles. Since gate complexity of $\mathcal{O}(\eta \log \eta \log N)$ is significantly less than the cost of time evolving these systems, it can be regarded as a negligible additive cost to the overall simulation.

Finally, one might notice that the spin degree of freedom does not appear in the Hamiltonians written above, nor does it appear in Eq. (1). A spin is associated with each of the η registers but because there are not magnetic field terms in Eq. (1) that would explicitly interact with these spins, it is not necessary to include the spin label and so we avoid doing so for simplicity. However, because particles of different spins are distinguishable, this should be accounted for during the antisymmetrization of the initial wave function.

B. Overview of the Hamiltonian simulation frameworks deployed in this work

In the prior section we described the full non-Born-Oppenheimer Hamiltonian in a plane-wave basis and suggested that plane waves would be an ideal representation in which to perform nonadiabatic dynamics. However, the algorithms we analyze in detail in this paper focus on the Born-Oppenheimer Hamiltonian. Still, we emphasize that the non-Born-Oppenheimer application helps to motivate our work since the algorithms used to simulate the Born-Oppenheimer Hamiltonian apply to the non-Born-Oppenheimer Hamiltonian with very minor modifications; the only additional terms in the non-Born-Oppenheimer Hamiltonian are U_{nBO} (identical to V up to different charges) and T_{nuc} (identical to T up to different masses).

Because we simulate the Born-Oppenheimer Hamiltonian, which is typically of interest in electronic structure, our focus is preparing eigenstates rather than effecting dynamics. Specifically, we use the phase estimation algorithm to sample in the eigenbasis of the Hamiltonian [66,67]. Among several possible uses of this technique is the refinement of molecular ground states from initial guess states with nonvanishing support on the exact ground state [3]. The requirement for such algorithms is that we are able to synthesize a unitary circuit encoding eigenvalues of the Born-Oppenheimer Hamiltonian as a simple function of their eigenphases. In this paper, we study two different algorithms for this task based on the qubitization [4] and interaction-picture simulation [5] algorithms of Low *et al.*, with subroutines adapted for simulating first quantized chemistry by Babbush *et al.* [42]. We now outline these approaches at a high level and establish conventions useful for later (more technical) sections.

Both of the algorithms we explore are based (at least in part) on block encoding the Hamiltonian using qubitization operators [4]. This method begins from the observation that we express our target Hamiltonian as a linear combination of unitaries (often referred to here as ‘‘LCU’’) [68], i.e.,

$$H = \sum_{\ell} \alpha_{\ell} H_{\ell}, \quad (16)$$

where H_{ℓ} are unitary operators and α_{ℓ} are positive coefficients (by convention we choose to incorporate any phases into the H_{ℓ}). Using this notation we then define the quantum operators PREP_H and SEL_H as

$$\begin{aligned} \text{PREP}_H |0\rangle &= \sum_{\ell} \sqrt{\frac{\alpha_{\ell}}{\lambda}} |\ell\rangle, \quad \text{SEL}_H |\ell\rangle = |\ell\rangle \otimes H_{\ell}, \\ \lambda &= \sum_{\ell} \alpha_{\ell} \end{aligned} \quad (17)$$

and we notice that $\langle 0 | \text{PREP}_H^{\dagger} \cdot \text{SEL}_H \cdot \text{PREP}_H |0\rangle = H/\lambda$ [5]. Because of this property, if we can realize PREP_H and SEL_H as quantum circuits then by calling them in a certain sequence it is possible to ‘‘block encode’’ the Hamiltonian in a particular subspace of a qubitization operator that we define as

$$Q = (2 |0\rangle\langle 0| - I) \cdot \text{PREP}_H^{\dagger} \cdot \text{SEL}_H \cdot \text{PREP}_H. \quad (18)$$

What this means is that Q acts as a quantum walk [69] with eigenvalues $e^{\pm i \arccos(E_k/\lambda)}$ [4], related to the energy E_k of the target Hamiltonian H . The work of Refs. [8,21] showed that we can use these quantum walks to project to an eigenstate of the Hamiltonian and estimate its eigenvalue by performing phase estimation on Q and applying the cosine function on the resulting phase. To estimate with accuracy ϵ , we need $\mathcal{O}(\lambda/\epsilon)$ uses of controlled- Q

operations. Alternatively, the resource requirements of performing phase estimation can be reduced by using the multiplexed qubitization operation [20]:

$$|0\rangle\langle 0| \otimes Q^{\dagger} + |1\rangle\langle 1| \otimes Q, \quad (19)$$

instead of the controlled- Q operation. This amplifies the difference between the eigenvalues by a factor of 2, reducing the complexity by a factor of 2. Provided SEL_H is self-inverse, this control can be achieved simply by making the reflection controlled by a qubit from the control register for phase estimation, as per Fig. 2 of Ref. [27]. This control only adds a cost of one Toffoli per step.

We note that in addition to being useful for phase estimation, one can also realize time evolution under the Born-Oppenheimer Hamiltonian for duration t by invoking these quantum walks Q within the quantum-signal-processing framework of Low *et al.* [70] a number of times scaling as

$$\mathcal{O}\left(\lambda t + \frac{\log(1/\epsilon)}{\log \log(1/\epsilon)}\right). \quad (20)$$

Thus, while our explicit focus is on phase estimation of the Born-Oppenheimer Hamiltonian, the same algorithms can be used to perform time evolution under the non-Born-Oppenheimer Hamiltonian, with trivial modifications and marginal additional cost. The similarity of the algorithms enabling these disparate applications results from our choice to work in a first quantized representation with basis functions that are independent of nuclear coordinates; if we were instead to work in second quantization or use Gaussian orbitals, there would be a substantial gap between the algorithms required for these two applications. Getting back to the implementation of the qubitization algorithm, in order to express the Hamiltonian of Eq. (8) as a linear combination of unitaries we rewrite the T , U , and V operators as follows:

$$\begin{aligned} T &= \frac{2\pi^2}{\Omega^{2/3}} \sum_{j=1}^{\eta} \sum_{p \in G} \sum_{w \in \{x,y,z\}} \sum_{r=0}^{n_p-2} \sum_{s=0}^{n_p-2} 2^{r+s} p_{w,r} p_{w,s} |p\rangle\langle p|_j, \\ U &= \sum_{v \in G_0} \sum_{\ell=1}^L \frac{2\pi \zeta_{\ell}}{\Omega \|k_v\|^2} \sum_{j=1}^{\eta} \sum_{b \in \{0,1\}} \left(-e^{-ik_v \cdot R_{\ell}} \sum_{q \in G} (-1)^{b[(q-v) \notin G]} |q-v\rangle\langle q|_j \right), \\ V &= \sum_{v \in G_0} \frac{\pi}{\Omega \|k_v\|^2} \sum_{i \neq j=1}^{\eta} \sum_{b \in \{0,1\}} \left(\sum_{p,q \in G} (-1)^{b[(p+v) \notin G] \vee [(q-v) \notin G]} |p+v\rangle\langle p|_i \cdot |q-v\rangle\langle q|_j \right), \end{aligned} \quad (21)$$

where

$$n_p = \lceil \log(N^{1/3} + 1) \rceil \quad (22)$$

represents the number of qubits required to store a signed binary representation of one component of the momentum of a single electron. We can see that the potential operators U and V are now expressed as linear combinations of unitaries. To express the T operator as a linear combination of unitaries we note that the quantity $p_{w,r}$ is bit r of component w of p , and the sum of these products of bits over r and s gives $\|p\|^2$; we apply the identity

$$p_{w,r}p_{w,s} = \frac{1 - (-1)^{p_{w,r}p_{w,s}}}{2}, \quad (23)$$

which gives

$$T = \frac{\pi^2}{\Omega^{2/3}} \sum_{j=1}^{\eta} \sum_{w \in \{x,y,z\}} \sum_{r=0}^{n_p-2} \sum_{s=0}^{n_p-2} 2^{r+s} \times \sum_{b \in \{0,1\}} \left(\sum_{p \in G} (-1)^{b(p_{w,r}p_{w,s} \oplus 1)} |p\rangle \langle p|_j \right). \quad (24)$$

With the Hamiltonian components expressed as a linear combination of unitaries we can now discuss the normalizing parameter λ appearing in Eq. (17), which will have significant ramifications for the complexity of both algorithms studied in this work. The parameter λ is an induced 1-norm of the Hamiltonian expressed as a linear combination of unitaries. For the entire Hamiltonian we have $\lambda = \lambda_T + \lambda_U + \lambda_V$ where

$$\begin{aligned} \lambda_T &= \frac{6\eta\pi^2}{\Omega^{2/3}} (2^{n_p-1} - 1)^2 = \mathcal{O}\left(\frac{\eta N^{2/3}}{\Omega^{2/3}}\right), \\ \lambda_U &= \frac{\eta \sum_{\ell} \zeta_{\ell}}{\pi \Omega^{1/3}} \lambda_v = \mathcal{O}\left(\frac{\eta^2 N^{1/3}}{\Omega^{1/3}}\right), \\ \lambda_V &= \frac{\eta(\eta-1)}{2\pi \Omega^{1/3}} \lambda_v = \mathcal{O}\left(\frac{\eta^2 N^{1/3}}{\Omega^{1/3}}\right), \\ \lambda_v &= \sum_{v \in G_0} \frac{1}{\|v\|^2} \leq \int_0^{2\pi} \int_0^{\pi} \int_0^{N^{1/3}} dr d\phi d\theta \left(\frac{1}{r^2}\right) r^2 \sin\theta \\ &= \mathcal{O}(N^{1/3}), \end{aligned} \quad (25)$$

where in the asymptotic scaling for λ_U we use the fact that when including all electrons explicitly in a charge neutral simulation, $\sum_{\ell} \zeta_{\ell} = \eta$. Further, we note that it is often appropriate to take $\Omega \propto \eta$ (e.g., whenever growing condensed phase systems towards their thermodynamic limit), which simplifies the asymptotic complexities to $\lambda_T = \mathcal{O}(\eta^{1/3} N^{2/3})$ and $\lambda_U \propto \lambda_V = \mathcal{O}(\eta^{5/3} N^{1/3})$.

The main challenge of realizing an algorithmlike qubitization with low complexity is coming up with efficient circuits for the operators defined in Eq. (17). Describing such algorithms asymptotically for first quantized chemistry is one of the main contributions of the work of Babbush *et al.* [42] and this work will loosely follow the methods outlined therein. However, the fact that this paper is an order of magnitude longer than the former is a testament to the fact that bounding asymptotic complexity is significantly less work than coming up with, improving, and compiling, a practical implementation of that algorithm with low constant factors in a fault-tolerant cost model. Thus, the main technical contribution of Sec. II is to describe a detailed algorithm for realizing the qubitization operators corresponding to the first quantized plane-wave electronic structure Hamiltonian. We introduce various new techniques, including a strategy for controlled swapping the momentum registers (Secs. II B and II D) and a proposal for performing the kinetic term when the state preparation for the potential fails (Appendix D), which lead to a significant reduction in the algorithmic complexity. Ultimately, we show that the operator Q can be realized with a gate complexity of $\tilde{\mathcal{O}}(\eta)$ and that one can implement phase estimation to estimate eigenvalues of Eq. (8) to within precision ϵ with a leading order Toffoli complexity given by Theorem 4.

By combining the $\tilde{\mathcal{O}}(\eta)$ gate complexity of Q with the $\mathcal{O}(\lambda/\epsilon)$ times that Q must be invoked within phase estimation, we see that the asymptotic complexity of the qubitized algorithm for phase estimation is

$$\begin{aligned} \tilde{\mathcal{O}}\left(\frac{(\lambda_T + \lambda_U + \lambda_V)\eta}{\epsilon}\right) &= \tilde{\mathcal{O}}\left(\frac{\eta^2 N^{2/3}}{\epsilon \Omega^{2/3}} + \frac{\eta^3 N^{1/3}}{\epsilon \Omega^{1/3}}\right) \\ &= \tilde{\mathcal{O}}\left(\frac{\eta^{4/3} N^{2/3} + \eta^{8/3} N^{1/3}}{\epsilon}\right) \end{aligned} \quad (26)$$

as reported for the qubitized algorithm complexity in Ref. [42]. While the scaling of this algorithm in terms of η is worse than the scaling in terms of N , for practical applications we may be more concerned about the scaling with N . The reason is because in order to accurately solve the electronic structure problem in a plane-wave basis it will be necessary to use a very large number of plane waves. For example, some enticing applications may require roughly $\eta = 100$ and $N = 10^6$. While such simulations would be impossible within second quantization because that would (in this case) require at least $N = 10^6$ logical qubits, within first quantization the space complexity required for the system register is only $\eta \log N \approx 2 \times 10^3$ qubits. Here and throughout the paper we use the convention that \log indicates a logarithm base 2.

The vast majority of this space complexity arises from representing the simulated system, which has basis states

of the form

$$|p_1\rangle \cdots |p_j\rangle \cdots |p_\eta\rangle. \quad (27)$$

Representing coordinates of the orbitals explicitly, we have

$$|p_{1x}, p_{1y}, p_{1z}\rangle \cdots |p_{jx}, p_{jy}, p_{jz}\rangle \cdots |p_{\eta x}, p_{\eta y}, p_{\eta z}\rangle. \quad (28)$$

As in Eq. (6), these registers hold values in the range $[-(N^{1/3} - 1)/2, (N^{1/3} - 1)/2]$. Thus, we use n_p qubits with $N^{1/3} = 2^{n_p} - 1$ to encode each register. Furthermore, for reasons of algorithmic efficiency described later, we use signed binary numbers to represent each register of length n_p , so

$$|q\rangle = |q_{\text{sign}}, q_{n_p-2}, \dots, q_0\rangle. \quad (29)$$

The total number of qubits in the simulated system is thus

$$n_s = 3 \eta n_p. \quad (30)$$

So while the space complexity is likely not a problem for performing simulations at large N , the gate complexity scaling as $\tilde{O}(N^{2/3})$ is somewhat concerning and we might wonder if it is possible to reduce this scaling. This is the motivation for the second algorithm we explore (also introduced in Ref. [42]), which provides an algorithm that scales linearly in $\lambda_U + \lambda_V$ but avoids scaling polynomially in λ_T (which is where the $N^{2/3}$ scaling was entering previously). The approach is based on the interaction-picture simulation technique introduced by Low *et al.* [5], which begins by supposing that we want to simulate a target Hamiltonian containing two terms A and B , where $\|A\|$ is significantly larger than $\|B\|$ but time evolution of A can be fast forwarded. Then, we can represent the evolution in the interaction picture of the rotating frame of A ,

$$e^{-i\tau(A+B)} = e^{-i\tau A} \mathcal{T} \exp\left(-i \int_0^\tau ds e^{isA} B e^{-isA}\right), \quad (31)$$

where \mathcal{T} is the time-ordering operator. The potential benefit of using this representation is that the larger term A appears only in the exponent of matrix exponentials and its implementation can be efficient even when $\|A\|$ is large. For the first quantized plane wave Hamiltonian in the regime that $N \gg \eta$, we have the norm relation $\|T\| \geq \|U + V\|$ so we choose $A = T$ and $B = U + V$ [42].

To perform quantum simulation in the interaction picture, we truncate and discretize the Dyson series of the time-ordered exponential, obtaining

$$\begin{aligned} e^{-i(A+B)\tau} &= e^{-i\tau A} \mathcal{T} \exp\left(-i \int_0^\tau ds e^{isA} B e^{-isA}\right) \\ &= \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} \int_0^\tau d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{k-1}} d\tau_k e^{-i(\tau-\tau'_k)A} \end{aligned}$$

$$\begin{aligned} &\cdot B e^{-i(\tau'_k - \tau'_{k-1})A} B \cdots B e^{-i(\tau'_2 - \tau'_1)A} B e^{-i\tau_1 A} \\ &\approx \sum_{k=0}^K \frac{(-i\tau)^k}{M^k k!} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \cdots \sum_{m_k=0}^{M-1} e^{-i\tau(M-1/2-m'_k)A/M} \\ &\cdot B e^{-i\tau(m'_k - m'_{k-1})A/M} B \cdots \\ &\cdots B e^{-i\tau(m'_2 - m'_1)A/M} B e^{-i\tau(m'_1 + 1/2)A/M} \end{aligned} \quad (32)$$

where $0 \leq \tau'_1 \leq \dots \leq \tau'_k \leq \tau$ are sorted times from τ_1, \dots, τ_k , and $0 \leq m'_1 \leq \dots \leq m'_k \leq M-1$ are sorted integers from m_1, \dots, m_k . The resulting expression is only an approximation of the ideal evolution, but the truncation and discretization error can be made arbitrarily small by choosing K and M sufficiently large. These parameters give a contribution to the complexity, which scales as a polynomial in $\log(1/\epsilon)$, which is omitted when giving complexity using \tilde{O} .

As discussed in Sec. III, the bottleneck requirements for realizing this Dyson series are that one implements e^{-iT} and also block encodes $U + V$ a number of times scaling as $(\lambda_U + \lambda_V)/\epsilon$. For the most part, one can reuse parts of the circuits in Sec. II to perform the block encoding of $U + V$ with complexity $\tilde{O}(\eta)$. Furthermore, it is relatively straightforward to provide circuits that fast forward the operator e^{-iT} with gate complexity $\tilde{O}(\eta)$. Thus, this suggests that the overall asymptotic complexity will scale as

$$\tilde{O}\left(\frac{(\lambda_U + \lambda_V)\eta}{\epsilon}\right) = \tilde{O}\left(\frac{\eta^3 N^{1/3}}{\epsilon \Omega^{1/3}}\right) = \tilde{O}\left(\frac{\eta^{8/3} N^{1/3}}{\epsilon}\right). \quad (33)$$

However, the most challenging part of the interaction-picture algorithm is to efficiently tie together these primitives within the Dyson-series simulation framework and propagate errors from that simulation through the phase-estimation circuits. While past work has discussed quantum simulation of the Dyson series [5,71] in asymptotic terms, no prior work has provided concrete circuits for the task and worked out the constant factors in the scaling. One of the primary contributions of Sec. III is to work out those circuits in the context of chemistry, optimize their implementations, and bound the overall Toffoli cost, as reported in Theorem 5. In doing so, we make several technical contributions, including a novel qubitization of the pre-amplified Dyson series (Sec. III A), an efficient strategy to update the kinetic energy (Sec. III B), and a higher-order discretization of the time integrals (Appendix H), which greatly improves over the naive approaches. We note that we also work out the constant factor scalings for the interaction-picture simulation in a more general context (not related to chemistry), which we report in Appendix I.

Finally, in Sec. IV our paper concludes with numerical comparisons between these two algorithms assessing

the resources required to deploy them to real molecular systems. In addition to comparing the two algorithms, we also compare to prior work analyzing the fault-tolerant viability of second quantized Gaussian orbital approaches. Although the comparison is somewhat nuanced, the first quantized approaches developed here appear to give lower Toffoli complexities than any past work for realistic simulations of both material Hamiltonians and nonperiodic molecules. We note that throughout this work we focus on estimating the constant factor Toffoli complexity and number of ancilla required. We focus on Toffoli complexity (as opposed to the complexity of other gates) because these are the operations that would be most challenging to implement within most practical error-correcting codes such as the surface code [72]. Furthermore, the algorithms described here have the same asymptotic Clifford complexities as the Toffoli complexities that we report. This makes it more likely that the implementation would actually be bottlenecked by Toffoli distillation. This stands in contrast to other recent chemistry algorithms compiled to fault tolerance with extensive use of quantum read-only memory (QROM) [20,73]. Approaches relying heavily on QROM such as Refs. [20,22,26,27] tend to have a worse asymptotic Clifford complexity than Toffoli complexity, thus necessitating a careful layout of how the algorithm might be realized using lattice surgery or other surface-code fault-tolerant gates, in order to assess whether state distillation is actually the dominant bottleneck.

There are a number of new techniques that we introduce, which can be used in many other applications. One is to note that in some cases amplitude amplification can be eliminated from the state-preparation procedure for the linear combination of unitaries. The reason for this is that the Hamiltonian can be composed of a number of parts, where that state preparation is only needed for one part of the Hamiltonian. When the state preparation is regarded as having “failed,” we can simply flag that as a part of the state, which will control the select for another part of the Hamiltonian. It is useful to eliminate the amplitude amplification, which would otherwise triple the cost (or more if further amplification steps are needed).

Another new technique is to eliminate the arithmetic for some parts of the Hamiltonian. If a component of the Hamiltonian gives a multiplication of the state by a function of the basis state, then it can be broken up into a sum of individual bit products. The bit product requires only a single Toffoli, and the sum is implemented via the linear combination of unitaries. In our case the component of the Hamiltonian is the kinetic energy, and we break up the kinetic energy into a sum of products of bits of the momenta. This dramatically reduces the Toffoli complexity.

In the case of the interaction picture we can greatly reduce the complexity needed for dealing with the kinetic energy (though not completely eliminate arithmetic like we

can for qubitization). The idea is to use a kinetic energy register. This register can be updated with relatively small complexity and used for the phasing. Another approach we use for the interaction picture is to use a novel qubitization of the Dyson series. This again eliminates the need for an amplitude amplification, which would otherwise triple the complexity. This method works generally for applications where the goal is to measure the eigenvalue, rather than to generate the time evolution under the Hamiltonian.

Our implementation of these algorithms allows us to report the first constant factors for the scaling of any first quantized chemistry algorithm (or any Dyson-series-based simulation), and also optimizes those constant factors more than a thousandfold. This allows us to compare the viability of these approaches to other methods based on second quantization, and to each other. Ultimately, we find that the qubitization algorithm is usually more practical and that first quantized methods often require much less surface-code space-time volume for simulating millions of plane waves than the best second quantized algorithms require for simulating hundreds of Gaussian orbitals.

II. THE QUBITIZATION-BASED ALGORITHM

We now give an overview of the qubitization algorithm and its circuit implementation for simulating the first quantized quantum chemistry. For this purpose, we seek to represent the target Hamiltonian as a linear combination of unitaries $H = \sum_{\ell=1}^L \alpha_{\ell} H_{\ell}$, where H_{ℓ} are unitary operators and α_{ℓ} are positive coefficients. Recall from Sec. I B that we can construct a state preparation subroutine PREP_H and a selection subroutine SEL_H , such that the block encoding $\langle 0 | \text{PREP}_H^{\dagger} \cdot \text{SEL}_H \cdot \text{PREP}_H | 0 \rangle = H/\lambda$ holds with normalization factor $\lambda = \sum_{\ell} \alpha_{\ell}$. Then, the operator $Q = (2 | 0 \rangle \langle 0 | - I) \cdot \text{PREP}_H^{\dagger} \cdot \text{SEL}_H \cdot \text{PREP}_H$ has eigenvalues $e^{\pm i \arccos(E_k/\lambda)}$ corresponding to the energy E_k of H . We can thus estimate the spectrum of H by performing phase estimation on Q with classical postprocessing.

For the chemistry Hamiltonian $H = T + U + V$, we see from Sec. I B that the terms T , U , and V can indeed be expressed as linear combinations of unitaries. This means that we have PREP_T , PREP_{U+V} and SEL_T , SEL_{U+V} , such that the block encodings

$$\begin{aligned} \langle 0 | \text{PREP}_T^{\dagger} \cdot \text{SEL}_T \cdot \text{PREP}_T | 0 \rangle &= \frac{T}{\lambda_T}, \\ \langle 0 | \text{PREP}_{U+V}^{\dagger} \cdot \text{SEL}_{U+V} \cdot \text{PREP}_{U+V} | 0 \rangle &= \frac{U+V}{\lambda_U + \lambda_V} \end{aligned} \quad (34)$$

hold with normalization factors λ_T and $\lambda_U + \lambda_V$, respectively. As a result, we find that the state-preparation

subroutine

$$\left(\sqrt{\frac{\lambda_T}{\lambda_T + \lambda_U + \lambda_V}} |0\rangle + \sqrt{\frac{\lambda_U + \lambda_V}{\lambda_T + \lambda_U + \lambda_V}} |1\rangle \right) \otimes \text{PREP}_T |0\rangle \otimes \text{PREP}_{U+V} |0\rangle \quad (35)$$

and the selection subroutine

$$|0\rangle\langle 0| \otimes \text{SEL}_T \otimes I + |1\rangle\langle 1| \otimes I \otimes \text{SEL}_{U+V} \quad (36)$$

block encode the target Hamiltonian H with normalization factor $\lambda_T + \lambda_U + \lambda_V$, giving a circuit implementation of the qubitization algorithm.

There is a subtle issue with the above naive analysis. While the kinetic operator T can be encoded with a normalization factor close to λ_T , in the most efficient approach for preparation the potential operators $U + V$ are only encoded with a normalization factor close to $4(\lambda_U + \lambda_V)$. To be more specific, we can construct a state preparation

$$\widetilde{\text{PREP}}_{U+V} |0\rangle \otimes I \approx \frac{1}{2} |0\rangle \otimes \text{PREP}_{U+V} + \frac{\sqrt{3}}{2} |1\rangle \otimes \text{PREP}_{U+V}^\perp \quad (37)$$

that succeeds only with a probability close to $1/4$, where $|0\rangle$ flags that the preparation is successful and $|1\rangle$ indicates a failure. If we were to use such a probabilistic subroutine in our block encoding, we would implement qubitization with a normalization factor close to $\lambda_T + 4(\lambda_U + \lambda_V)$. Alternatively, we could perform a single step of oblivious amplitude amplification to boost the success probability close to unity, but this triples the cost of the state preparation and can introduce overhead to the overall gate complexity.

Our new approach improves the above implementations by proceeding even when the state preparation of $U + V$ fails. To elaborate, first consider the special case where $\lambda_T = 3(\lambda_U + \lambda_V)$. Then, we propose to perform the selection subroutine of T following the failure of the preparation of $U + V$, i.e.,

$$I \otimes |0\rangle\langle 0| \otimes \text{SEL}_{U+V} + \text{SEL}_T \otimes |1\rangle\langle 1| \otimes I. \quad (38)$$

This new selection operator, combined with the state preparation

$$\text{PREP}_T |0\rangle \otimes \widetilde{\text{PREP}}_{U+V} |0, 0\rangle, \quad (39)$$

gives the block encoding

$$\begin{aligned} \frac{3T}{4\lambda_T} + \frac{U+V}{4(\lambda_U + \lambda_V)} &= \frac{T}{4(\lambda_U + \lambda_V)} + \frac{U+V}{4(\lambda_U + \lambda_V)} \\ &= \frac{T+U+V}{\lambda_T + \lambda_U + \lambda_V} \end{aligned} \quad (40)$$

with the desired normalization factor $\lambda_T + \lambda_U + \lambda_V$.

The general case can be handled by using an ancilla qubit to adjust the relative amplitudes between T and $U + V$. If $\lambda_T < 3(\lambda_U + \lambda_V)$, then we introduce an ancilla qubit and perform only the selection of T conditioned on both the failure of the preparation of $U + V$ AND the ancilla qubit is in state $|0\rangle$, i.e.,

$$\begin{aligned} I \otimes I \otimes |0\rangle\langle 0| \otimes \text{SEL}_{U+V} + |0\rangle\langle 0| \otimes \text{SEL}_T \otimes |1\rangle\langle 1| \\ \otimes I + |1\rangle\langle 1| \otimes I \otimes |1\rangle\langle 1| \otimes I. \end{aligned} \quad (41)$$

This selection operator, combined with the state preparation

$$(\cos \theta |0\rangle + \sin \theta |1\rangle) \otimes \text{PREP}_T |0\rangle \otimes \widetilde{\text{PREP}}_{U+V} |0, 0\rangle, \quad (42)$$

gives the block encoding

$$\frac{U+V}{4(\lambda_U + \lambda_V)} + \frac{3T \cos^2 \theta}{4\lambda_T} \quad (43)$$

up to a shifting by multiples of the identity operator. For this to be an accurate encoding of the Hamiltonian, this expression must be proportional to $U + V + T$. That means we need $1/[4(\lambda_U + \lambda_V)] = (3 \cos^2 \theta)/(4\lambda_T)$; solving for θ , we find that $\theta = \arccos \sqrt{\lambda_T/3(\lambda_U + \lambda_V)}$, which gives the target Hamiltonian encoded with normalization factor $4(\lambda_U + \lambda_V)$.

In the case where $\lambda_T > 3(\lambda_U + \lambda_V)$, we would still introduce an ancilla qubit to adjust the relative amplitudes. However, we now perform the selection of T conditioned on failure of the preparation subroutine OR the ancilla qubit being in state $|0\rangle$ (and perform SEL_T in the remaining situations), i.e.,

$$\begin{aligned} |1\rangle\langle 1| \otimes I \otimes |0\rangle\langle 0| \otimes \text{SEL}_{U+V} + |1\rangle\langle 1| \otimes \text{SEL}_T \otimes |1\rangle\langle 1| \\ \otimes I + |0\rangle\langle 0| \otimes \text{SEL}_T \otimes I \otimes I. \end{aligned} \quad (44)$$

Combined with the state preparation

$$(\cos \theta |0\rangle + \sin \theta |1\rangle) \otimes \text{PREP}_T |0\rangle \otimes \widetilde{\text{PREP}}_{U+V} |0, 0\rangle, \quad (45)$$

this gives the block encoding

$$\frac{(U+V) \sin^2 \theta}{4(\lambda_U + \lambda_V)} + \frac{3T \sin^2 \theta}{4\lambda_T} + \frac{T \cos^2 \theta}{\lambda_T}. \quad (46)$$

For this to be proportional to $U + V + T$ we need

$$\frac{\sin^2 \theta}{4(\lambda_U + \lambda_V)} = \frac{3 \sin^2 \theta}{4\lambda_T} + \frac{\cos^2 \theta}{\lambda_T}. \quad (47)$$

Solving for θ , we find that $\theta = \arcsin [2\sqrt{(\lambda_U + \lambda_V)/(\lambda_T + \lambda_U + \lambda_V)}]$, which gives the target

Hamiltonian encoded with the normalization factor $\lambda_T + \lambda_U + \lambda_V$.

In the following, we sketch the high-level ideas behind the circuit implementation of the state preparation and selection subroutines, and leave a discussion of the details to subsequent sections, including a more detailed description of the implementation in Sec. II A. For simplicity, we

focus on the case where $\lambda_T < 3(\lambda_U + \lambda_V)$ although similar analysis applies to the other cases as well. Note that λ_T is not significantly larger than $\lambda_U + \lambda_V$ in this parameter regime and this is when the qubitization algorithm is expected to be advantageous.

For state preparation, we aim to prepare the quantum state

$$\begin{aligned} & (\cos \theta |0\rangle + \sin \theta |1\rangle)_a |+\rangle_b \frac{1}{\sqrt{\eta}} \left(\sqrt{\eta-1} |0\rangle_c \sum_{i \neq j=1}^{\eta} |i\rangle_d |j\rangle_e + |1\rangle_c \sum_{j=1}^{\eta} |j\rangle_d |j\rangle_e \right) \left(\frac{1}{\sqrt{3}} \sum_{w=0}^2 |w\rangle_f \right) \\ & \otimes \left(\frac{1}{2^{n_p-1}-1} \sum_{r,s=0}^{n_p-2} 2^{(r+s)/2} |r\rangle_g |s\rangle_h \right) \left(\sqrt{\frac{\lambda_U}{\lambda_U + \lambda_V}} |0\rangle + \sqrt{\frac{\lambda_V}{\lambda_U + \lambda_V}} |1\rangle \right)_i \\ & \otimes \left(\sqrt{\frac{p_v}{\lambda_v}} |0\rangle_j \sum_{v \in G_0} \frac{1}{\|v\|} |v\rangle_k + \sqrt{1-p_v} |1\rangle_j |v^\perp\rangle_k \right) \left(\frac{1}{\sqrt{\sum_{\ell} \zeta_{\ell}}} \sum_{\ell=1}^L \sqrt{\zeta_{\ell}} |\ell\rangle \right), \end{aligned} \quad (48)$$

with p_v the probability of successfully preparing the momentum state, where the first ancilla register is used to adjust the relative amplitudes between T and $U+V$ (as explained above), and the i register (holding state with relative amplitudes $\sqrt{\lambda_U}$ and $\sqrt{\lambda_V}$) is used to select between U and V . Here, the preparation subroutine PREP_T acts on the registers b, f, g, h and outputs the corresponding states, whereas PREP_{U+V} acts on registers c, d, e, j, k, l . Note that as some of these registers are used for multiple terms in the Hamiltonian, the assignment of common registers such as $|\cdot\rangle_b$ to any particular preparation circuit is arbitrary. Quantum circuits for these preparation subroutines are detailed in Secs. II B and II C, respectively. The normalization factor λ_v is defined in Eq. (25), and $|v^\perp\rangle$ is a state obtained in the failure of initializing the momentum register. We see that the probability of successfully preparing the momentum state $(1/\sqrt{\lambda_v}) \sum_{v \in G_0} (1/\|v\|) |v\rangle$ is close to $1/4$, which is consistent with our discussion above.

Note that several important details, including the flagging of $i \neq j$ in register c and the precision in preparing the ancilla state $\cos \theta |0\rangle + \sin \theta |1\rangle$, are omitted in our analysis for exposition purposes. We provide full details on how the encodings between T, U , and V are performed and what precision the ancilla state is prepared with in Appendix D.

We now describe the key steps in the implementation of the selection operations. For SEL_T , we can compute the product $p_{w,r} p_{w,s}$ in an ancilla register and use that to perform a controlled Z gate on $|+\rangle$, followed by a final uncomputation. The net effect of this transformation is

$$\begin{aligned} \text{SEL}_T : & |b\rangle_b |j\rangle_e |w\rangle_f |r\rangle_g |s\rangle_h |p_j\rangle \\ \mapsto & (-1)^{b(p_{w,r} p_{w,s} \oplus 1)} |b\rangle_b |j\rangle_e |w\rangle_f |r\rangle_g |s\rangle_h |p_j\rangle, \end{aligned} \quad (49)$$

where $|b\rangle$ is either the component $|0\rangle$ or $|1\rangle$ of the state $|+\rangle$ and $|p_j\rangle$ is the j th particle's momentum. To implement SEL_{U+V} , we first add v into the i th momentum register. Since this should only be done for V , we perform the addition controlled by the register selecting between U and V . Next, we subtract v from the j th momentum register for both U and V , and the subtraction does not need to be controlled by the corresponding selection register. We now check if $(p-v) \notin G$ (for U) or $[(p+v) \notin G] \vee [(q-v) \notin G]$ (for V) and use the result to perform controlled Z gate on $|+\rangle$. For the case of U , we also perform the phase rotation $-e^{-ik_v \cdot R_{\ell}}$. Overall, this implements

$$\begin{aligned} \text{SEL}_U : & |b\rangle_b |j\rangle_e |0\rangle_i |v\rangle_k |\ell\rangle_l |q_j\rangle \\ \mapsto & |b\rangle_b |j\rangle_e |0\rangle_i |v\rangle_k |\ell\rangle_l |q_j - v\rangle \\ \mapsto & (-1)^{b[(p-v) \notin G]} |b\rangle_b |j\rangle_e |0\rangle_i |v\rangle_k |\ell\rangle_l |q_j - v\rangle \\ \mapsto & -e^{-ik_v \cdot R_{\ell}} (-1)^{b[(p-v) \notin G]} |b\rangle_b |j\rangle_e |0\rangle_i |v\rangle_k \\ & \otimes |\ell\rangle_l |q_j - v\rangle \end{aligned} \quad (50)$$

and

$$\begin{aligned} \text{SEL}_V : & |b\rangle_b |i\rangle_d |j\rangle_e |1\rangle_i |v\rangle_k |p_i\rangle |q_j\rangle \\ \mapsto & |b\rangle_b |i\rangle_d |j\rangle_e |1\rangle_i |v\rangle_k |p_i + v\rangle |q_j - v\rangle \\ \mapsto & (-1)^{b[(p+v) \notin G] \vee [(q-v) \notin G]} |b\rangle_b |i\rangle_d |j\rangle_e |1\rangle_i \\ & \otimes |v\rangle_k |p_i + v\rangle |q_j - v\rangle \end{aligned} \quad (51)$$

for U and V , respectively, as controlled by the register selecting them. See Secs. II B and II D for a detailed description of the selection operations and their circuit implementations.

We bound the error of our implementation and provide a complete costing of the qubitization algorithm in Sec. II E, with our main result summarized in Theorem 4.

A. Circuit implementation of qubitization

Our aim in this section is to provide a high-level overview of how qubitization can be realized to sample from the energy eigenvalues of a first quantized Hamiltonian. The general strategy is to combine the circuits used for state preparation and selection between T , U , and V in order to minimize the overall cost. In particular, we prepare superpositions over i and j , which are used independent of which of T , U , and V is being block encoded. Then we need to perform operations on the momentum given in registers i and j , so we use i and j to control the swap of the momenta into temporary ancillae in order to perform the needed operations. For the case of T we are just computing a phase factor ± 1 dependent on the value of p . In contrast, for U we are performing arithmetic on the momentum from register j , and for V we are performing arithmetic on the momenta from both registers i and j . Therefore, the operations $p + \nu$ and $p - \nu$ will be controlled on registers that select between T , U , and V .

Our protocol for preparing the state, Eq. (48), consists of the following steps.

1. Rotate the qubit for selecting between T and $U + V$.
2. Prepare an equal superposition state over η values of i and j , and check that they are not equal.
3. Prepare superpositions over w , r , and s that will be used for the T part of the select operation. These preparations do not need to be controlled, because it is only the SEL part that is controlled on the qubit for T . The costing for this is shown to be $2n_p + 2b_r - 7$ in the subsection below for T .
4. Prepare an appropriate superposition state for selecting between U and V . We show that the complexity is approximately $4 \lceil \log(\eta + \lambda_\zeta) \rceil + 2b_r - 9$ Toffoli gates.
5. Prepare a superposition over ν with weights proportional to $1/\|\nu\|$. The costing for this is described below for the U and V parts of the Hamiltonian.
6. Prepare a state with amplitudes proportional to $\sqrt{\zeta_\ell}$, which may be done with complexity $\lambda_\zeta = \sum_\ell \zeta_\ell$.

Because we write the state to be prepared as a tensor product of the individual states, the overall cost of the state preparation is just the cost of the preparation of the states on the individual subsystems.

In order to perform the SEL operation, the key parts are as follows.

1. Add ν into the momentum register at location i , controlled on the first and sixth registers, which select V .

2. Subtract ν from the momentum register at location j , controlled on the first qubit that is $|1\rangle$ for U or V .
3. Perform a Z gate on the $|+\rangle$ state if $(q - \nu) \notin G$ (for U) or $[(p + \nu) \notin G] \vee [(q - \nu) \notin G]$ (for V).
4. Perform a Z gate on $|+\rangle$ state controlled on $p_{w,r}p_{w,s}$ for the case of T .
5. For the case of U perform the phase rotation $-e^{-ik_\nu \cdot R_\ell}$.

In addition, for our phase measurement procedure the SEL operation should be self-inverse. It is easily seen that the component of the SEL for T is self-inverse. To account for the SEL for U and V we can introduce an ancilla qubit that selects whether we add or subtract ν . Controlling addition versus subtraction is done with no additional Toffoli complexity. This ancilla qubit is flipped each step, to produce a self-inverse operation in the same way as in Eq. (8) of [21].

The following discussion provides a statement of the block-encoding procedure for the Born-Oppenheimer Hamiltonian.

Lemma 1. *The operator $(\text{PREP}^\dagger \otimes I)\text{SEL}(\text{PREP} \otimes I)$ forms, for $N_a = 10 + 2n_\eta + 3n_p + 2\lceil \log(n_p - 1) \rceil + \lceil \log(L) \rceil$, an $(\lambda, N_a, 0)$ -block encoding of H_{BO} , meaning that*

$$\begin{aligned} & \langle\langle 0|^{N_a} \otimes I \rangle\rangle (\text{PREP}^\dagger \otimes I) \text{SEL}(\text{PREP} \otimes I) \langle\langle 0|^{N_a} \otimes I \rangle\rangle \\ &= \frac{|0\rangle\langle 0|^{\otimes N_a} \otimes H_{BO}}{\lambda}. \end{aligned}$$

Proof. The proof of this lemma follows by noting that the number of qubits needed to encode the state yielded by PREP as given in Eq. (48). That coincides with the ancillary space for block encoding using the LCU lemma as in Refs. [6,68]. Further, the LCU lemma specifically shows that if $\text{PREP}'|0\rangle^{N_a} = \sum_\ell \alpha_\ell |j\rangle$ and $\text{SEL}'|\ell\rangle|\psi\rangle = |\ell\rangle U_\ell |\psi\rangle$ for $\alpha_\ell \geq 0$ then

$$\begin{aligned} & \langle\langle 0| \rangle\langle 0|^{\otimes N_a} \otimes I \rangle\rangle (\text{PREP}'^\dagger \text{SEL}' \text{PREP}') |0\rangle^{N_a} |\psi\rangle \\ &= \frac{|0\rangle^{N_a} \sum_\ell \alpha_\ell H_\ell |\psi\rangle}{\sum_\ell \alpha_\ell}. \end{aligned} \quad (52)$$

The unitary operations U_ℓ are given implicitly in the above list for SEL. In this section we do not consider the complexity of implementing the SEL operator and so use only the expression as a definition of the action of the operator on input states.

Specifically, recall that up to a constant energy offset due to the nuclear positions the Hamiltonian is $H = T + U + V$ where

$$\begin{aligned}
T &= \frac{\pi^2}{\Omega^{2/3}} \sum_{j=1}^{\eta} \sum_{w \in \{x,y,z\}} \sum_{r=0}^{n_p-2} \sum_{s=0}^{n_p-2} 2^{r+s} \sum_{b \in \{0,1\}} \left(\sum_{p \in G} (-1)^{b(p_w, r p_w, s \oplus 1)} |p\rangle \langle p|_j \right), \\
U &= \sum_{v \in G_0} \sum_{\ell=1}^L \frac{2\pi \zeta_{\ell}}{\Omega \|k_v\|^2} \sum_{j=1}^{\eta} \sum_{b \in \{0,1\}} \left(-e^{-ik_v \cdot R_{\ell}} \sum_{q \in G} (-1)^{b[(q-v) \notin G]} |q-v\rangle \langle q|_j \right), \\
V &= \sum_{v \in G_0} \frac{\pi}{\Omega \|k_v\|^2} \sum_{i \neq j=1}^{\eta} \sum_{b \in \{0,1\}} \left(\sum_{p,q \in G} (-1)^{b\{[(p+v) \notin G] \vee [(q-v) \notin G]\}} |p+v\rangle \langle p|_i \otimes |q-v\rangle \langle q|_j \right),
\end{aligned} \tag{53}$$

here $p_{w,r}$, for example, denotes the r th bit of the w th component of the momentum.

In this form, it is clear that the Hamiltonian operators in the LCU decomposition satisfy for each ℓ , $H_{\ell} \in \{S_T, S_U, S_V\}$ where each of these unitary operators are defined in Eqs. (49), (50), and (51). Thus we need to demonstrate two facts. First, that the operations in SEL selectively apply these transformations and that the coefficients of the expansion match the square roots of those in the preceding expression.

Consider steps 1 and 2 in the definition of SEL. These steps can be implemented using a controlled unitary adder. A unitary subtractor subtracting the vector v from a momentum register $|q\rangle_j$ has the action

$$|v\rangle |q\rangle_j \mapsto |v\rangle |q-v\rangle_j, \tag{54}$$

This therefore has the correct action for implementing the subtraction needed to implement U and V .

There are two cases we need to consider. First, assume $q-v \in G$. In this case the operator describing the transformation has the same action as $|v\rangle \langle v| \otimes |p-v\rangle \langle p|_j$. Now let us assume that $p-v \notin G$. In this case we have an overflow. However, since $\sum_{b \in \{0,1\}} (-1)^{b[(q-v) \notin G]} = 0$ in Eq. (21) the action of the unitary adder need not be defined in this case and hence an additional qubit is not needed in the register to address potential overflow problems. (We use that later to simplify the implementation.) With this fact in mind, it is then straightforward to show that steps 1 and 2 block encode the V and U terms in the Hamiltonian.

To explain how this gives the required linear combination of unitaries, note that the general procedure is to implement an operator of the form $\sum_{\ell} \alpha_{\ell} H_{\ell}$, prepare an ancilla with amplitudes proportional to $\sqrt{\alpha_{\ell}}$, then perform a controlled H_{ℓ} operation. The first step in SEL gives the addition $p+v$ in momentum register i required for V . Step 2 gives the subtraction $q-v$ required for V , the together with step 3 the complete operation needed for V is performed as described in Eq. (51). Step 2 also gives $q-v$ as required for U , step 3 also gives the sign flip for U , and step 5 gives the phase factor needed for U .

Together they give the required operations for U as shown in Eq. (50). The correct amplitudes are obtained by the amplitudes $1/\|v\|$ in Eq. (48), which gives the correct weights $1/\|v\|^2$ in the Hamiltonian. The SEL operation for T is obtained entirely by step 4, which gives the sign flip. The appropriate weights in the sum are obtained by using the superpositions over j, w, r, s in Eq. (48). With these definitions in place we can formally state the following.

The block encoding of T is a little less obvious, but it ultimately follows similarly from the definitions of the involved terms. Recalling from Eq. (25) that $\lambda_T = 6\eta\pi^2(2^{n_p-1}-1)^2/\Omega^{2/3}$ and noting the definition of T given in Eq. (24) we see immediately that the theorem holds after substituting the terms in the LCU decomposition in Eq. (24) into the U_j in Eq. (52). ■

For the rotation of the qubit to select between T and $U+V$, we find that error in the rotation angle of $\Delta\theta$ results in error in the energy no more than $\lambda\Delta\theta$ (see Appendix D). As a result, if we take the number of bits for this rotation to be n_T , then the error from this source is bounded as

$$\frac{\pi\lambda}{2^{n_T}} \leq \epsilon_T. \tag{55}$$

The Toffoli complexity of performing this rotation is $n_T - 3$. In some applications it may be possible to reduce the complexity by selecting Ω appropriately. That is, the ratio $\lambda_T/(\lambda_U + \lambda_V)$ will depend on Ω , so one can choose Ω to that the rotation may be performed exactly with a relatively small number of bits.

In order to quantify the complexity of performing the state preparation for selecting between U and V , we can take account of the relative amplitudes between the different parts. We find that

$$\frac{\lambda_U}{\lambda_V} = \frac{2\lambda_{\zeta}}{\eta-1} \quad \lambda_{\zeta} = \sum_{\ell=1}^L \zeta_{\ell}, \tag{56}$$

where we remind the reader that $\lambda_{\zeta} = \eta$ for charge neutral systems. Even if that is not the case, we see that λ_U/λ_V is simply a ratio of integers. Rather than performing a

rotation, one can instead achieve an exact state preparation as follows. First, prepare an equal preparation over $\eta - 1 + 2\lambda_\zeta$ numbers. Then, perform an inequality test between this register and η to give the correct relative weightings $\eta - 1$ and $2\lambda_\zeta$. There is a subtlety in that we initially prepare superpositions over all numbers i and j including the case $i = j$. This effectively increases the value of λ_V , so the ratio we want is $2\lambda_\zeta/\eta$. Then we prepare an equal superposition over $\eta + 2\lambda_\zeta$ numbers (rather than $\eta - 1 + 2\lambda_\zeta$).

After preparing the superposition state and applying the inequality test, use a Toffoli between the result of the inequality test and the first qubit selecting between T and $U + V$, to give a qubit flagging that we are applying V (so will perform the addition $p + v$). Writing the number of qubits needed to store $\eta + 2\lambda_\zeta$ as $n_{\eta\zeta}$, the complexity of preparing the equal superposition state is [27] (see Appendix A.2, p. 39 of version 2)

$$3n_{\eta\zeta} + 2b_r - 9, \quad (57)$$

and the cost of the inequality test is $n_{\eta\zeta} - 1$. Here b_r is a number of bits of precision used for rotation on an ancilla. Together with the extra Toffoli, the cost is

$$4n_{\eta\zeta} + 2b_r - 9. \quad (58)$$

The convenience of this is that it introduces little error due to the state preparation.

The probability of success of the state preparation is usually above 0.999 for $b_r = 7$ (see Fig. 10 of [74]) given in Eq. (205) of Ref. [74], and is

$$P_s(n, b_r) = \frac{n}{2^{\lceil \log n \rceil}} \left\{ \left[1 + \left(2 - \frac{4n}{2^{\lceil \log n \rceil}} \right) \sin^2 \theta(n, b_r) \right]^2 + \sin^2 [2\theta(n, b_r)] \right\}, \quad (59)$$

where n is the number of basis states to create an equal superposition over, and $\theta(u, b_r)$ is the angle of rotation. Since we are using b_r bits for the rotation angle, the angle is of the form $2\pi n/2^{b_r}$ for some n , and

$$\theta(n, b_r) = \frac{2\pi}{2^{b_r}} \text{round} \left[\frac{2^{b_r}}{2\pi} \arcsin \left(\sqrt{\frac{2^{\lceil \log n \rceil}}{4n}} \right) \right]. \quad (60)$$

More generally, we prove that the probability of error is no more than $(3/2)^2(\pi/2^{b_r})^2$ with a slight adjustment to the formula for θ (see Appendix J). A simpler case is that where $\lambda_\zeta = \eta$, which occurs when the total charge is zero. In that case, the ratio we want is just 2. Instead of needing to prepare an equal superposition over $\eta + 2\lambda_\zeta$ values, we need only a superposition over 3. Including $n_T - 3$ Toffolis

to perform the rotation on the first qubit, the total complexity for the preparation on the qubits selecting T , U , and V would be

$$n_T + 4n_{\eta\zeta} + 2b_r - 12. \quad (61)$$

Next we consider the cost of the preparation of the equal superposition of i and j in unary. The steps needed are as follows.

1. Prepare an equal superposition in binary, using the approach in Appendix A.2 of Ref. [27], with cost $3n_\eta + 2b_r - 9$ for each, where $n_\eta = \lceil \log \eta \rceil$ and b_r is the angle of rotation on an ancilla qubit.
2. Test if i and j are equal in binary, flagging the result in a qubit. This qubit will need to be combined with another success qubit for applying the potential, so the cost is n_η .
3. The testing $i \neq j$ can be inverted with n_η Toffolis.
4. The inversion of the equal superposition of i and j can be inverted with cost $3n_\eta + 2b_r - 9$ for each.

Adding all these costs together gives a total complexity

$$14n_\eta + 8b_r - 36. \quad (62)$$

For the failure case where $i = j$ and we would otherwise be selecting V (rather than U), we can simply combine this with the failure of the state preparation for $|v\rangle$, and perform SEL for T in that case.

For convenience of the rest of the procedure, we produce three qubits for selecting T , U , and V . These are as follows.

1. A qubit that is 0 if we are to perform T .
2. A qubit that is 0 if we are performing the subtraction $q - v$. This will be the case for U or V .
3. A qubit that is 0 for V only, which controls the addition $p + v$.

First, there are four preparations of equal superposition states, and checking that all of those succeeded takes 3 Toffolis. There are also another four qubits needed to select T, U, V . These are the rotated qubit [register 1 in Eq. (48)], the qubit resulting from the inequality test to select between U and V , the qubit flagging $i \neq j$, and the qubit flagging failure of the state preparation for v . For simplicity of the explanation, let us call these logic values A, B, C, D , respectively, with 0 corresponding to True.

To check that the state preparation for $U + V$ succeeded, we can use $(B \vee C) \wedge D$, which takes two Toffolis (one for each logic operation). In one scenario (OR), T is selected as $A \vee \neg[(B \vee C) \wedge D]$, and in the other (AND) T is selected as $A \wedge \neg[(B \vee C) \wedge D]$. Either takes one more Toffoli, but an extra Toffoli is needed to ensure that qubit is set only for

success of the preparation of the equal superposition states. That is 7 Toffolis so far.

Now, in the OR case, we set the second qubit (selecting U or V) if $\neg\{A \vee \neg[(B \vee C) \wedge D]\}$ and there is success of the equal superposition preparations. That takes one more Toffoli. In the AND case we set that second qubit according to $(B \vee C) \wedge D$ and success of the preparations, which is again one Toffoli. The third qubit can then be set by taking the AND of $\neg B$ and the result in the second qubit. That is one more Toffoli for a total of 9.

For the preparations of equal superposition states, the four needed are the equal superposition over three basis states for w , one with $\eta + 2\lambda_\zeta$ basis states, and two with η . For the superposition over three basis states required for w , it is highly efficient to take $b_r = 8$, so we use that rather than allowing arbitrary b_r there. As a result, the block-encoded state is multiplied by the probability

$$P_{\text{eq}} := P_s(3, 8) P_s(\eta + 2\lambda_\zeta, b_r) P_s(\eta, b_r)^2. \quad (63)$$

This corresponds to giving an effective λ divided by P_{eq} .

B. Block encoding the kinetic operator

Recall that the usual form of the kinetic operator is given by

$$T = \sum_{j=1}^{\eta} \sum_{p \in G} \frac{\|k_p\|^2}{2} |p\rangle \langle p|_j. \quad (64)$$

We rewrite this expression to express the square explicitly as a sum of products of bits, as

$$T = \frac{2\pi^2}{\Omega^{2/3}} \sum_{j=1}^{\eta} \sum_{p \in G} \sum_{w \in \{x, y, z\}} \sum_{r=0}^{n_p-2} \sum_{s=0}^{n_p-2} 2^{r+s} p_{w,r} p_{w,s} |p\rangle \langle p|_j, \quad (65)$$

where $p_{w,r}$ is indicating bit r of the w component of p . To block encode this, we need to prepare the superposition state proportional to

$$|+\rangle \sum_{j=1}^{\eta} |j\rangle \sum_{w=0}^2 |w\rangle \sum_{r=0}^{n_p-2} 2^{r/2} |r\rangle \sum_{s=0}^{n_p-2} 2^{s/2} |s\rangle. \quad (66)$$

The preparation of the equal superposition over j is addressed above. For the equal superposition over w , we take $b_r = 8$ and $n = 2$, so the formula $3n + 2b_r - 9$ for the complexity gives a cost of only 13 Toffolis.

The superposition over r and s can be prepared with high chance of success by using the following approach. First,

we aim to prepare the following state in unary:

$$2^{-n_p/2} |n_p - 1\rangle + \sum_{r=0}^{n_p-2} 2^{-(r+1)/2} |r\rangle. \quad (67)$$

Half the amplitude is on $|0\rangle$, then of the remaining states half the amplitude is on $|1\rangle$, and so forth. This means that to prepare the state in unary, we perform a Hadamard on the first qubit (labeled 0), then controlled on that we perform a Hadamard on the second qubit, and so forth. There are in total $n_p - 2$ controlled Hadamards, and each controlled Hadamard can be performed with a single Toffoli and a catalytically used $|T\rangle$ state as in Fig. 17 of Ref. [27]. The qubit for $n_p - 1$ is being used to flag success of the state preparation (it flags success with $|0\rangle$). Omitting the state $|n_p - 1\rangle$, the state flagged for success is (up to normalization)

$$\sum_{r=0}^{n_p-2} 2^{-(r+1)/2} |r\rangle. \quad (68)$$

We then flip all these remaining bits so that the state becomes

$$\sum_{r=0}^{n_p-2} 2^{-(r+1)/2} |n_p - 2 - r\rangle = 2^{-(n_p-1)/2} \sum_{r=0}^{n_p-2} 2^{r/2} |r\rangle. \quad (69)$$

We want the unary to be one hot, but that conversion may be performed with CNOTs so the complexity is omitted in the Toffoli count. Note that to account for the case $r = 0$, one can add one qubit in the state $|1\rangle$, so $r = 0$ is encoded as $|1000\dots\rangle$, $r = 1$ is encoded as $|1100\dots\rangle$, and so forth. Then, the sequence of CNOTs gives the desired one-hot unary. We therefore get a complexity of $n_p - 2$ for this state preparation. Together with the 13 Toffolis used to prepare the superposition over w , the overall complexity for the preparation over the w , r , and s registers is

$$2n_p + 9. \quad (70)$$

The effect of the preparation having a small probability of failure is that the effective value of λ is increased to

$$\lambda'_T = \frac{6\eta\pi^2}{\Omega^{2/3}} 2^{2(n_p-1)}, \quad (71)$$

where we replace $2^{n_p-1} - 1$ with 2^{n_p-1} .

In order to perform the SEL operation, we need to use the value of p in momentum register j . Because we need to perform arithmetic on the momentum in register j for U and V as well, the most efficient procedure is to simply control a swap of the momentum in register j into an ancilla register. Similarly, the momentum in register i will

be swapped into an ancilla register. The operations we perform on the momenta in these ancilla registers will then be controlled by the first two qubits selecting between T, U, V . That is more efficient than performing the operations in a controlled way on every momentum register. The complexity of controlling the swaps of the momenta out of and back into momentum registers i and j is $12\eta n_p$. For the control, we perform unary iteration on i and j . That requires $\eta - 2$ Toffolis for each of the four times (swapping the register out and back for i , then again for j), for a total of $4\eta - 8$. Together with $12\eta n_p$ that gives a total of

$$12\eta n_p + 4\eta - 8. \quad (72)$$

We consider that cost separately from the costs for implementing SEL specifically for T, U, V . In the case of T , the other operations needed to perform SEL are as follows.

1. Use w to control copying of component w of p into another ancilla. This has cost $3(n_p - 1)$.
2. Copy bit r of component w of p to an ancilla qubit, with cost $n_p - 1$.
3. Copy bit s of component w of p to an ancilla qubit, with cost $n_p - 1$.
4. Controlled on those two bits, perform a phase flip on the $|+\rangle$ state unless both bits are one. This gives an orthogonal $|-\rangle$ state unless both bits are one. The orthogonal state gives no contribution to the block encoding. Only one Toffoli is needed here.
5. Erase the two qubits containing $p_{w,r}$ and $p_{w,s}$, then the register containing component w of p . These erasures can be done with measurements and Clifford gates.

The overall Toffoli complexity is only about as much as is needed to run through the qubits representing p . It is $5(n_p - 1) + 1$. There is also some complexity due to the fact that we also have a control qubit selecting T , but control on one extra qubit needs only one more Toffoli. That gives a total SEL cost for T

$$5(n_p - 1) + 2, \quad (73)$$

in addition to the SEL cost $12\eta n_p$ that is common between T, U, V .

C. State preparation for U and V

The U and V potential operators both have a $1/\|k_v\|^2$ weighting, which means that in order to do the state preparation for the block encoding we need to prepare a state proportional to

$$\sum_{v \in G_0} \frac{1}{\|v\|} |v\rangle, \quad (74)$$

where $v \in G_0 = [- (N^{1/3} - 1), N^{1/3} - 1]^3 \setminus \{(0, 0, 0)\}$ is a three-component integer vector. In our implementation, we use eight registers $|\mu\rangle |v_x\rangle |v_y\rangle |v_z\rangle |m\rangle |0\rangle$ to hold this state. These subsystems are used as follows.

- (i) $|\mu\rangle$ flags which box is used in the series of nested boxes for the state preparation.
- (ii) $|v_x\rangle |v_y\rangle |v_z\rangle$ are the components of v given as signed integers.
- (iii) $|m\rangle$ is an ancilla in an equal superposition used to give the correct amplitude via an inequality test.
- (iv) $|0\rangle$ flags that the state preparation is successful.

There are four parts of the preparation where there is some amplitude for failure. First, there is the preparation of μ . Second, there is the negative zero in the signed integer representation of v , which is not allowed. Third, there is testing whether the position is inside the inner box in the state preparation. Fourth, there is the success of the inequality test used in the state preparation. Success for all four is indicated by $|0\rangle$ on the flag qubit above, which can then be used in amplitude amplification.

The state that we prepare has the form

$$\frac{1}{\sqrt{\mathcal{M}2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \sum_{m=0}^{\lceil \mathcal{M}(2^{\mu-2}/\|v\|)^2 \rceil - 1} \times \frac{1}{2^\mu} |\mu\rangle |v_x\rangle |v_y\rangle |v_z\rangle |m\rangle |0\rangle + |\Psi^\perp\rangle, \quad (75)$$

where B_μ is a subset of v to be defined later and $(I \otimes \langle 111|) |\Psi^\perp\rangle = 0$. Conditioned on the qubit flagging success with $|0\rangle$, the amplitude for each v is then

$$\sqrt{\frac{\lceil \mathcal{M}(2^{\mu-2}/\|v\|)^2 \rceil}{\mathcal{M}2^{2\mu}2^{n_p+2}}} \approx \frac{1}{8\sqrt{2^{n_p}}} \frac{1}{\|v\|}. \quad (76)$$

The probability of the flag qubit in state $|0\rangle$ is around $1/4$, and a single step of amplitude amplification boosts the success probability to close to 1 [42, Fig. 1]. Note that the probability is slightly smaller here than in Ref. [42], because we use a preparation on μ with nonzero amplitude for failure.

The approach is to use a hierarchy of nested boxes in $[-(N^{1/3} - 1), N^{1/3} - 1]$ indexed by μ and to prepare a set of v belonging to each box. We initially prepare a unary-encoded superposition state

$$\frac{1}{\sqrt{2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sqrt{2^\mu} |\mu\rangle = \frac{1}{\sqrt{2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sqrt{2^\mu} |0 \dots 0 \underbrace{1 \dots 1}_\mu\rangle. \quad (77)$$

This state can be created using Hadamard and controlled Hadamard gates in a similar way as we perform the state

preparation for r and s in order to implement the squaring of p . Here we omit the failure part of that preparation, so this state is subnormalized, with the norm indicating the probability of success. There is just one more bit than in the case where we are preparing r and s , so the cost is $n_p - 1$ Toffolis.

In the next stage, we prepare

$$\frac{1}{\sqrt{2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sum_{v_x, v_y, v_z = -(2^{\mu-1}-1), -0}^{2^{\mu-1}-1} \frac{1}{2^\mu} |\mu\rangle |v_x\rangle |v_y\rangle |v_z\rangle, \quad (78)$$

where the basis state is explicitly represented as

$$\begin{aligned} & |0 \cdots 0 \underbrace{1 \cdots 1}_\mu \rangle |v_{x,\text{sign}} 0 \cdots 0 v_{x,\mu-2} \cdots v_{x,0}\rangle \\ & \otimes |v_{y,\text{sign}} 0 \cdots 0 v_{y,\mu-2} \cdots v_{y,0}\rangle \\ & \otimes |v_{z,\text{sign}} 0 \cdots 0 v_{z,\mu-2} \cdots v_{z,0}\rangle. \end{aligned} \quad (79)$$

Note that in this signed integer representation, the number zero appears twice, with both a plus sign and a minus sign. The sum with “ -0 ” indicates that the negative zero is included at this point. This state preparation can be implemented using six Hadamards and $3(n_p - 1)$ controlled Hadamards. Specifically, we create the uniform superposition state in registers $|v_{x,\text{sign}}, v_{x,0}\rangle |v_{y,\text{sign}}, v_{y,0}\rangle |v_{z,\text{sign}}, v_{z,0}\rangle$ using six Hadamards. We always have an equal superposition on $v_{x,0}$, $v_{y,0}$ and $v_{z,0}$ because μ has a minimum value of 2. Then we generate uniform superpositions in $|v_{x,j}\rangle$, $|v_{y,j}\rangle$, and $|v_{z,j}\rangle$ controlled on the $(j + 1)$ th qubit in $|\mu\rangle$ from the right for $j = 2, \dots, n_p$. As before, the controlled Hadamards can be implemented with a Toffoli and a catalytic state, so the Toffoli cost is $3(n_p - 1)$.

Now, to remove the minus zero in this representation, we flag it as a failure. This can be done by checking each $|v_x\rangle$, $|v_y\rangle$, and $|v_z\rangle$ whether the sign bit is one and the remaining bits are zero, which requires controlled Toffolis each having $n_p + 1$ controls. We then use another two Toffoli gates to flag if any of these checkings returned true. Altogether, the Toffoli cost is

$$3n_p + 2. \quad (80)$$

In the next stage, we test whether all of v_x , v_y , and v_z are smaller in absolute value than $2^{\mu-2}$. First, we convert the μ register to one-hot unary using CNOTs. Then, conditioned on the μ th qubit being one, we flag $|f_{\text{box}}\rangle$ depending on whether the $(\mu - 1)$ th qubit $|v_{x,\mu-2}\rangle$, $|v_{y,\mu-2}\rangle$, and $|v_{z,\mu-2}\rangle$ are all zero, using a controlled Toffoli gate with four controls, which has a cost of three Toffolis. Because this must

be done n_p times, this gives a cost of

$$3n_p. \quad (81)$$

The costs so far are then $4(n_p - 1)$ controlled Hadamards and $6n_p + 2$ Toffolis. Although the complexity of the controlled Hadamard is one Toffoli, we need to pay this cost again when inverting the preparation, but the Toffolis can be erased with Clifford gates. The state excluding the failures then becomes

$$\frac{1}{\sqrt{2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \frac{1}{2^\mu} |\mu\rangle |v_x\rangle |v_y\rangle |v_z\rangle, \quad (82)$$

where B_μ is the set of v such that $|v_x|$, $|v_y|$, $|v_z|$ are less than $2^{\mu-1}$, excluding the case that they are all less than $2^{\mu-2}$.

We now prepare an ancilla state in an equal superposition of $|m\rangle$ for $m = 0$ to $\mathcal{M} - 1$. The parameter \mathcal{M} is a power of 2 and is chosen to be sufficiently large to obtain an accurate block encoding. This state can be prepared entirely using Hadamards, so no non-Clifford gates are required. The resulting state excluding the failures is now

$$\frac{1}{\sqrt{\mathcal{M}2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \sum_{m=0}^{\mathcal{M}-1} \frac{1}{2^\mu} |\mu\rangle |v_x\rangle |v_y\rangle |v_z\rangle |m\rangle. \quad (83)$$

In the final stage, we test the inequality

$$(2^{\mu-2})^2 \mathcal{M} > m(v_x^2 + v_y^2 + v_z^2). \quad (84)$$

We then need to use the four flag qubits for success to produce an overall success flag. The final state is then

$$\begin{aligned} & \frac{1}{\sqrt{\mathcal{M}2^{n_p+2}}} \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \sum_{m=0}^{\lceil \mathcal{M}(2^{\mu-2}/\|v\|)^2 \rceil - 1} \\ & \times \frac{1}{2^\mu} |\mu\rangle |v_x\rangle |v_y\rangle |v_z\rangle |m\rangle |0\rangle + |\Psi^\perp\rangle \end{aligned} \quad (85)$$

as desired. To achieve this, we compute $(2^{\mu-2})^2 \mathcal{M}$ and $m(v_x^2 + v_y^2 + v_z^2)$ separately in two ancilla registers and compare the results. The unsigned number $m(v_x^2 + v_y^2 + v_z^2)$ is bounded by $3\mathcal{M}2^{2n_\mu-2}$, so we need at most $2n_\mu + n_{\mathcal{M}}$ qubits to represent it, where we use

$$n_{\mathcal{M}} = \lceil \log(\mathcal{M}) \rceil. \quad (86)$$

The value $(2^{\mu-2})^2 \mathcal{M}$ is upper bounded by $2^{2n_\mu-4} \mathcal{M}$, so we could represent it with $2n_\mu - 4 + n_{\mathcal{M}}$ qubits.

Because \mathcal{M} is a classically chosen value, and μ is given in unary, $(2^{\mu-2})^2 \mathcal{M}$ can be computed in place with n_0 gates, and simply a relabeling of qubits. We can place

zeroed qubits in between the qubits of the (one-hot) unary representation of μ , to give 2μ . For example, $\mu = 1$ is changed from 000001 to 000000000010, $\mu = 2$ is changed from 000010 to 000000001000, and so forth, where the underlined zeros are those added. Then this unary representation of 2μ is also a binary representation of $2^{2\mu+1}$. Lastly, to multiply by $2^{-5}\mathcal{M}$, we simply pad with $\log \mathcal{M} - 5$ zeros. In practice these zeroed ancillae do not even need to be included, and we simply relabel the qubits for μ .

To compute $m(v_x^2 + v_y^2 + v_z^2)$, we first use the result Lemma 8 given in Appendix G to cost the sum of three squares, which is $3n^2 - n - 1$ for n -bit numbers. We can regard v_x , v_y , and v_z as unsigned numbers of n_p bits, so squaring them costs $3n_p^2 - n_p - 1$ Toffolis. Finally, we multiply two binary numbers of length $\log(\mathcal{M})$ and $2n_p + 2$ using a similar approach (see Lemma 10) with

$$2n_{\mathcal{M}}(2n_p + 2) - n_{\mathcal{M}} \quad (87)$$

Toffoli gates. To test the inequality $(2^{\mu-2})^2 \mathcal{M} > m(v_x^2 + v_y^2 + v_z^2)$, we use a $(2n_p + n_{\mathcal{M}} + 2)$ -bit comparator, which has cost [75]

$$2n_p + n_{\mathcal{M}} + 2. \quad (88)$$

The net cost of the squaring, multiplication, and inequality test is then

$$3n_p^2 + n_p + 1 + 4n_{\mathcal{M}}(n_p + 1) \quad (89)$$

Toffoli gates. There are a further 3 Toffolis needed to produce the overall success flag qubit.

Now for the overall block encoding, the state preparation needs to be inverted. Rather than paying this cost again, we can keep the ancillae used to perform the arithmetic and inequality test. As a result, the inequality test and arithmetic can be inverted with Clifford gates. The 3 Toffolis used for the overall success flag qubit can be inverted with Clifford gates too. Similarly, the $6n_p + 2$ Toffolis used for the testing on v can be inverted with zero non-Clifford cost. The only operations in this state preparation that have a non-Clifford cost in the inverse preparation are the $4(n_p - 1)$ controlled Hadamards used for the preparation of the μ register and superpositions on the v register. As a result, the total Toffoli cost for the preparation and inverse preparation is

$$3n_p^2 + 15n_p - 7 + 4n_{\mathcal{M}}(n_p + 1). \quad (90)$$

In the case where amplitude amplification is performed, the cost will be 3 times this, due to the need to invert the inequality test and calculation, then perform it again.

The other key parts of the state preparation for U and V are to prepare the superpositions over i , j , and ℓ . The preparation over i and j are just simple preparations

of equal superposition states, which we have addressed already. The final preparation we need to consider here is therefore the preparation over ℓ with weightings $\sqrt{\zeta_\ell}$. In the preparation over the second qubit selecting between U and V , we have already prepared a superposition, which will be an equal superposition over $2\lambda_\zeta$ values for the case of U . That is the case where we need to prepare this superposition state, so we can just prepare it in this case, rather than preparing it for all cases as described above.

For the preparation, we can just use a QROM on this register, outputting the value of ℓ in an output register. When there are ζ_ℓ values in the input register that give the same value of ℓ in the output register, the result is obtaining an amplitude exactly proportional to $\sqrt{\zeta_\ell}$. Moreover, we need to compute $k_v \cdot R_\ell$ in performing the SEL operation. The value of R_ℓ can be output in another ancilla register at the same time. The complexity of the QROM, using the simple form of QROM, is just $\lambda_\zeta - 1$. The QROM should be controlled on the qubit, which $U + V$, so R_ℓ are not output there. The QROM automatically does not output R_ℓ for V , because this input register is outside the range for V .

Although we have aimed to prepare a state with amplitudes $\sqrt{\zeta_\ell}$, note that the only way that ℓ is used is in outputting R_ℓ . That means, if we are using QROM to output both ℓ and R_ℓ at once, then the ℓ register is not actually used. That means we can simply omit the ℓ register, and output R_ℓ . That saves ancilla qubits, though does not change the Toffoli count.

In inverting this preparation, erasing the output of the QROM has a Toffoli complexity of $2^k + \lceil 2^{-k}\lambda_\zeta \rceil$ where k is taken to be a power of 2. The resulting complexity is approximately the square root of the initial QROM. For convenience of describing the complexities, we define

$$\text{Er}(x) := \min_{k \in \mathbb{N}} (2^k + \lceil 2^{-k}x \rceil), \quad (91)$$

so this cost is $\text{Er}(\lambda_\zeta)$. In terms of this function, the cost of outputting the values of R_ℓ via QROM and erasing them again can be given as

$$\lambda_\zeta + \text{Er}(\lambda_\zeta). \quad (92)$$

D. The SEL operations for U and V

We have described the block encoding for T , and the state preparation for U and V , and now give the method to perform the SEL operations for U and V . The general principle is as follows.

1. Use the i and j registers to control a swap of the momentum registers p and q into ancillae.
2. Controlled on the qubit that is $|0\rangle$ only for V , add v (in place) into the ancilla for p .
3. Controlled on the qubit that is $|0\rangle$ for U or V , subtract v into the ancilla for q .

4. Check if either $p + \nu$ or $p - \nu$ is outside the box G , and if so perform a phase gate on the $|+\rangle$ state.
5. Control swap the ancillae back into the momentum registers.
6. Apply the phase $e^{-ik_\nu \cdot R_\ell}$ in the case of U only.

The cost of the controlled swaps was discussed above, and is $12\eta n_p$. The key cost to discuss now is the addition and subtraction of ν to momenta. The difficulty is that the momentum and ν are encoded as signed integers, whereas the usual methods for addition and subtraction are for integers encoded in two's complement.

In order to achieve this, just one of the numbers needs to be converted into two's complement. For example, since $p + \nu$ is computed in place, we can convert p to two's complement, then use the sign bit of ν to control adding or subtracting $|\nu|$ into p . This control of addition versus subtraction can be done with no more complexity than regular addition and subtraction. This is done for each component, so the overall complexity of converting between signed and two's complement is approximately 12 times n_p for the conversions. The addition and subtraction also need to be controlled by the qubits selecting between T, U, V , so have complexity twice uncontrolled additions and subtractions. That gives an overall complexity about

$$24n_p. \quad (93)$$

More specifically, to convert an n digit number from signed to two's complement, we take the sign qubit and use it as a control on CNOTs with the remaining qubits. Then we add it to the remaining $n - 1$ bits with complexity $n - 2$ Toffolis. We can just invert this process for converting in the opposite direction. Because p and q have n_p bits in each of their three components, the complexity of converting them to two's complement is $6(n_p - 2)$.

For the cost of the controlled addition (or subtraction), there is cost $n_p + 1$ to make each component controlled, which can be achieved by (for example) controlling copying out of ν to an ancilla. (That can be erased without Toffolis by measurements and phase fixups.) Then the addition has cost $n_p + 1$, giving a total cost of $2(n_p + 1)$ for the controlled addition or subtraction of each component. The total cost is $12(n_p + 1)$ for these controlled additions and subtractions.

Lastly we need to convert $p + \nu$ and $q - \nu$ back to signed form. Adding or subtracting ν expands these numbers by 2 bits, so the cost is now $6n_p$. The total cost is then

$$6(n_p - 2) + 12(n_p + 1) + 6n_p = 24n_p \quad (94)$$

Toffolis. Thus we get exactly the same cost as approximately estimated above.

Another major part of the SEL operations is to check whether $p + \nu$ or $q - \nu$ is outside the bounding box G .

Because the complexity depends only on N through the number of bits used in the arithmetic n_p , it is best to take N as large as possible given n_p , which means that the bounds of the box $\pm(N^{1/3} - 1)/2$ are taken to be of the form $2^{n_p-1} - 1$. That means that, if $p + \nu$ or $q - \nu$ is outside the bounding box, then at least one of the extra qubits introduced in the arithmetic is set to $|1\rangle$. These qubits are additional, so are not swapped from the ancilla back to the main momentum registers.

If we perform the checking as presented above, then we use these qubits to control a Z on the $|+\rangle$ state, in order to remove parts of the Hamiltonian where $p + \nu$ or $q - \nu$ are outside the bounding box, which are excluded. However, note that in the block encoding we would select on $|0\rangle$ for these ancillae, which means that these parts of the Hamiltonian are automatically eliminated. It is therefore unnecessary to perform this controlled phase in the SEL operation.

The last thing we need to do for implementing SEL is to apply the phase factor $-e^{-ik_\nu \cdot R_\ell}$. To achieve this we need to compute $k_\nu \cdot R_\ell$ then add it into a phase gradient register to apply the phase. We do not need to make the addition into the phase gradient register controlled on the qubits that select U , because the register with R_ℓ is nonzero only for U . The minus sign in this phase does need to be performed in a controlled way, but that is just a controlled phase on two qubits, which is a Clifford gate.

In order to perform the phase shift, a subtlety is that it is given in terms of k_ν rather than ν . However, the conversion factor between k_ν and ν can simply be absorbed into R_ℓ , rather than explicitly multiplied. This multiplication is most conveniently performed with signed integers. The product of each pair of components will then have a sign that is the product of the signs of the individual components. That product of signs can be computed with a CNOT. However, if we were to then add the components to compute $k_\nu \cdot R_\ell$, we would need to deal with the problem of adding signed integers again. Instead, it is convenient to add the products of each pair of components into the phase gradient state separately. They can be added or subtracted into the phase gradient state controlled on the sign with no more complexity than a regular addition.

The maximum value of a component of k_ν is $2\pi/\Omega^{1/3}$ times $N^{1/3} - 1$, which is the maximum value of a component of ν , and the maximum value of a component of R_ℓ is $\Omega^{1/3}/2$, since $\Omega^{1/3}$ is the width of the spatial region. Therefore, a component of k_ν times R_ℓ can be rewritten as a component of ν times $2\pi R_\ell/\Omega^{1/3}$, which has a maximum value of π . There is one initial CNOT used to find the product of the signs of the component of ν and R_ℓ . That will be used to control addition or subtraction of numbers into the phase gradient state. For the multiplication of the least significant bit of ν times $2\pi R_\ell/\Omega^{1/3}$, we have cost $n_R - 1$ to compute the bit products, because we now exclude the sign bit. The most significant bit of that product being added

into the phase gradient state corresponds to a phase shift of $\pi/2$, and the least significant bit corresponds to a phase shift of $2\pi/2^{n_R}$. As a result, the cost of adding these bits into the phase gradient state is $n_R - 2$. For the next most significant bit of v , the most significant bit of the product would give phases of π , so we still need to compute the $n_R - 1$ bit products, but since all phases are multiplied by 2, the cost of adding into the phase gradient state is reduced to $n_R - 3$. After that, there is cost $n_R - 2$ for the bit products, and $n_R - 4$ for addition into the phase gradient state. For the k 'th bit of v , we have cost $n_R - k + 1$ for the bit products and $n_R - k - 1$ for the addition for a total of $2(n_R - k)$. There is a subtlety in that if $n_R \leq n_p$ then $n_R - k - 1$ can be negative. Ignoring that case for the moment, because we would typically expect $n_R > n_p$, the Toffoli cost is

$$-1 + \sum_{k=1}^{n_p} 2(n_R - k) = 2n_p n_R - n_p(n_p + 1) - 1. \quad (95)$$

In the case where $n_R = n_p$, then we get one more Toffoli, because we simply have $n_R - k$ instead of $n_R - k - 1$ for the maximum value of k . Then the formula for the complexity can be given as $2n_p n_R - n_p(n_p + 1) = n_R(n_R - 1)$. In the case where $n_R < n_p$, then at $k = n_R + 1$ we have both costs equal to zero, and so the sum can be taken to $k = n_R$, to give

$$\sum_{k=1}^{n_R} 2(n_R - k) = n_R(n_R - 1). \quad (96)$$

Here we remove the -1 at the beginning because the expression in the sum overcounts the Toffolis by 1 at $k = n_R$. This cost is multiplied by 3, because it is done

once for each component. Therefore the Toffoli cost may be given as

$$\begin{cases} 3[2n_p n_R - n_p(n_p + 1) - 1], & n_R > n_p \\ 3n_R(n_R - 1), & n_R \leq n_p. \end{cases} \quad (97)$$

For convenience of the final expression, we can give the Toffoli cost as just $6n_p n_R$, but the cost can be computed more accurately with Eq. (97). The maximum number of ancilla qubits used is $n_R - 1$ for the bit products, plus $n_R - 3$ for the addition into the phase gradient register, for a total of $2n_R - 4$.

The total Toffoli costs for the block encoding of the Hamiltonian are now as in Table II. Here b_r is a number of bits used for ancilla qubit rotations in preparations of equal superposition states, and can be taken to be 7 for a probability of success no less than 0.999.

The complete step needed for the qubitization, to give the step that one would perform phase estimation on, needs a reflection on the control ancilla. The cost of this reflection corresponds to the numbers of qubits used in the state preparation. The total qubit cost, and the qubits needed to be reflected upon, are given in Appendix C. There are a total of

$$n_{\eta\zeta} + 2n_{\eta} + 6n_p + n_{\mathcal{M}} + 16 \quad (98)$$

qubits to reflect on. The corresponding Toffoli cost is two less. However, this reflection needs to be controlled on the qubits used as control for the phase measurement. Also, there is a single Toffoli cost for iterating those control qubits. We bundle those extra two Toffolis together with this cost, giving the cost as in Eq. (98).

TABLE II. The costs involved in the block encoding of the Hamiltonian for qubitization.

Procedure	Toffoli cost
Preparing the superposition on the registers selecting between T, U, V , where $n_{\eta\zeta}$ is the number of bits needed to represent $\eta + \lambda_{\zeta}$; see Eq. (61).	$2(n_T + 4n_{\eta\zeta} + 2b_r - 12)$
Preparing equal superpositions over η values of i and j in unary; see Eq. (62).	$14n_{\eta} + 8b_r - 36$
The state-preparation cost for the w, r, s registers used for T , including the factor of 2 for inversion; see (70).	$2(2n_p + 9)$
Controlled swaps of the p and q registers into and out of ancillae (which is used for T, U , and V); see Eq. (72).	$12\eta n_p + 4\eta - 8$
The SEL cost for T ; see Eq. (73).	$5(n_p - 1) + 2$
The preparation of the state with amplitude $1/\ v\ $ and inversion; see Eq. (90).	$3n_p^2 + 15n_p - 7 + 4n_{\mathcal{M}}(n_p + 1)$
The QROM for R_{ℓ} , combined with the state preparation with amplitudes $\sqrt{\zeta_{\ell}}$ (which may be implicit); see Eq. (92).	$\lambda_{\zeta} + \text{Er}(\lambda_{\zeta})$
Additions and subtractions of v into the momentum registers; see Eq. (93).	$24n_p$
Phasing by $-e^{-ik_v \cdot R_{\ell}}$; see Eq. (97).	$6n_p n_R$.

E. Total cost of constructing the qubitization operator

This is a complete accounting for the cost of a single step of the block encoding, except that we need to estimate the error due to the finite values of \mathcal{M} and n_R in order to determine the appropriate values to use for these quantities. The finite value of n_R corresponds to replacing R_ℓ with an n_R -digit approximation \tilde{R}_ℓ , so that the potential term U that the algorithm actually implements is close to

$$U = -\frac{4\pi}{\Omega} \sum_{\ell=1}^L \sum_{j=1}^{\eta} \sum_{\substack{p,q \in G \\ p \neq q}} \left(\zeta_\ell \frac{e^{ik_{q-p} \cdot \tilde{R}_\ell}}{\|k_{p-q}\|^2} \right) |p\rangle \langle q|_j. \quad (99)$$

By choosing n_R sufficiently large, this error can be at most an appropriate fraction of the overall error budget ϵ . To bound the error due to this source and choose the value of n_R , we use the following Lemma.

Lemma 2. *Given $\|\tilde{R}_\ell - R_\ell\| \leq \delta_R$ for all ℓ , if λ_j is the j^{th} eigenvalue of the true Hamiltonian and Eq. (99) is used to approximate the nuclear potential, then there exists an eigenvalue of the approximated Hamiltonian $\tilde{\lambda}_j$ such that*

$$|\lambda_j - \tilde{\lambda}_j| \leq \frac{2\delta_R \eta \lambda_\zeta}{\Omega^{2/3}} \sum_{v \in G_0} \frac{1}{\|v\|}. \quad (100)$$

Proof. First, note from Ref. [76, Corollary 6.3.4] that for any normal matrices M and M' that $\|M - M'\| \leq \epsilon$ implies that for every eigenvalue λ_j of M that there exists an eigenvalue of M' within distance ϵ . Thus it suffices to bound the spectral norm of the difference between the true potential operator and the approximated potential. Our claim then follows from the following bound on the difference between the two Hamiltonians:

$$\|U - \tilde{U}\| = \left\| \sum_{v \in G_0} \sum_{\ell=1}^L \frac{4\pi \zeta_\ell}{\Omega k_v^2} \sum_{j=1}^{\eta} (e^{-ik_v \cdot \tilde{R}_\ell} - e^{-ik_v \cdot R_\ell}) \left(\sum_{p \in G} |p - v\rangle \langle p|_j \right) \right\|, \quad (101)$$

$$\leq \eta \sum_{v \in G_0} \sum_{\ell=1}^L \frac{4\pi \zeta_\ell}{\Omega \|k_v\|^2} |e^{-ik_v \cdot \tilde{R}_\ell} - e^{-ik_v \cdot R_\ell}| \leq \eta \sum_{v \in G_0} \sum_{\ell=1}^L \frac{4\pi \zeta_\ell}{\Omega \|k_v\|^2} |k_v \cdot (\tilde{R}_\ell - R_\ell)|, \quad (102)$$

$$\leq \eta \sum_{v \in G_0} \sum_{\ell=1}^L \frac{4\pi \zeta_\ell}{\Omega \|k_v\|^2} \|k_v\| \|\tilde{R}_\ell - R_\ell\| \leq \frac{2\delta_R \eta \lambda_\zeta}{\Omega^{2/3}} \sum_{v \in G_0} \frac{1}{\|v\|}. \quad (103)$$

Note that $\sum_{v \in G_0} (1/\|v\|)$ can be further upper bounded by a function of N . However, we use the current bound as it is tighter and more versatile for numerical implementation. ■

Next we consider the error due to the finite value of \mathcal{M} . As can be seen from (85), the amplitude for each v is proportional to

$$\sqrt{\frac{\lceil \mathcal{M}(2^{\mu-2}/\|v\|)^2 \rceil}{\mathcal{M}2^{2\mu}2^{n_p+2}}}, \quad (104)$$

where μ is chosen such that $v \in B_\mu$. From the definition of B_μ , this means that μ is the minimum integer such that $|v_x|$, $|v_y|$, and $|v_z|$ are all less than $2^{\mu-1}$. That is,

$$\mu = \lceil \log \max(|v_x|, |v_y|, |v_z|) \rceil + 2. \quad (105)$$

As a result, the success probability of this state preparation is given by

$$p_v = \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \frac{\lceil \mathcal{M}(2^{\mu-2}/\|v\|)^2 \rceil}{\mathcal{M}2^{2\mu}2^{n_p+1}}. \quad (106)$$

As discussed in Ref. [42, Sec. II D], this is a constant around 1/4 provided the ceiling function is not used. In the case where the ceiling function is used, then the probability of success is only increased.

To bound the error due to the ceiling function, we can write each weight that is prepared as $1/\|v'\|^2$, for a modified v' . Now it would appear that all coefficients are larger, because of the ceiling function. However, when accounting for the normalization of the state, that will be larger, so each coefficient is multiplied by a constant that is smaller than 1, resulting in approximate coefficients that may be larger or smaller than the exact coefficients. In fact, because we are choosing the amplitude on the potential terms, we have freedom to adjust it to minimize the error. Let us therefore call the constant factor α . The squared amplitudes in the state we have prepared correspond to approximately

$$\frac{1}{2^{n_p+6}} \frac{1}{\|v\|^2}. \quad (107)$$

Therefore, let us put the values of v' as

$$\frac{1}{2^{n_p+6}} \frac{1}{\|v'\|^2} = \alpha \frac{\lceil \mathcal{M}2^{\mu-2}/\|v\| \rceil^2}{\mathcal{M}2^{2\mu}2^{n_p+2}}, \quad (108)$$

so

$$\frac{1}{\|v'\|^2} = 16\alpha \frac{\lceil \mathcal{M}2^{2\mu}/(16\|v\|^2) \rceil}{\mathcal{M}2^{2\mu}}. \quad (109)$$

Using this approach the error in the eigenvalues due to the approximation of the state preparation may be bounded as follows.

Lemma 3. *Given that the state preparation for v is performed as described above with a given \mathcal{M} , if λ_j is the j^{th} eigenvalue of the true Hamiltonian then there exists an eigenvalue of the approximated Hamiltonian $\tilde{\lambda}_j$ such that*

$$\begin{aligned} |\lambda_j - \tilde{\lambda}_j| &\leq \frac{2\eta}{\pi \mathcal{M} \Omega^{1/3}} (\eta - 1 + 2\lambda_\zeta) (7 \times 2^{n_p+1} - 9n_p - 11 - 3 \times 2^{-n_p}). \\ &= \frac{2^{n_p} (28\eta) (\eta - 1 + 2\lambda_\zeta)}{\pi \mathcal{M} \Omega^{1/3}} [1 + \mathcal{O}(2^{-n_p})]. \end{aligned}$$

Proof. Again we use the fact that the norm of the difference in the operators can be used to upper bound the difference in the eigenvalues. The error in the approximation of U can be bounded as

$$\begin{aligned} \|U - \tilde{U}\| &\leq \left\| \sum_{v \in G_0} \sum_{\ell=1}^L \frac{\zeta_\ell}{\pi \Omega^{1/3}} \left(\frac{1}{v^2} - \frac{1}{v'^2} \right) \sum_{j=1}^{\eta} e^{-ik_v \cdot R_\ell} \left(\sum_{p \in G} |p - v\rangle \langle p|_j \right) \right\|, \\ &\leq \sum_{v \in G_0} \frac{\eta \lambda_\zeta}{\pi \Omega^{1/3}} \left| \frac{1}{\|v'\|^2} - \frac{1}{\|v\|^2} \right|, \\ &= \frac{\eta \lambda_\zeta}{\pi \Omega^{1/3}} \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \left| \frac{1}{\|v'\|^2} - \frac{1}{\|v\|^2} \right|. \end{aligned} \quad (110)$$

Similarly, the error in the approximation of V can be bounded as

$$\begin{aligned} \|V - \tilde{V}\| &\leq \left\| \sum_{v \in G_0} \frac{1}{2\pi \Omega^{1/3}} \left(\frac{1}{v^2} - \frac{1}{v'^2} \right) \sum_{\substack{i,j=1 \\ i \neq j}}^{\eta} \left(\sum_{p,q \in G} |p + v\rangle \langle p|_i \otimes |q - v\rangle \langle q|_j \right) \right\|, \\ &\leq \sum_{v \in G_0} \frac{\eta(\eta - 1)}{2\pi \Omega^{1/3}} \left| \frac{1}{\|v\|^2} - \frac{1}{\|v'\|^2} \right|, \\ &= \frac{\eta(\eta - 1)}{2\pi \Omega^{1/3}} \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \left| \frac{1}{\|v'\|^2} - \frac{1}{\|v\|^2} \right|. \end{aligned} \quad (111)$$

Thus the total error in the potential component of the Hamiltonian can be upper bounded as

$$\|U + V - (\tilde{U} + \tilde{V})\| \leq \frac{\eta}{2\pi \Omega^{1/3}} (\eta - 1 + 2\lambda_\zeta) \sum_{\mu=2}^{n_\mu} \sum_{v \in B_\mu} \left| \frac{1}{\|v'\|^2} - \frac{1}{\|v\|^2} \right|. \quad (112)$$

For given parameters, it is possible to bound the error by numerically performing this sum. To give an analytic result, we take $\alpha = 1$ and place an upper bound on the sum. We obtain

$$\begin{aligned}
& \sum_{\mu=2}^{n_\mu} \sum_{v \in B_\mu} \left| \frac{1}{\|v'\|^2} - \frac{1}{\|v\|^2} \right| \\
& \leq \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \frac{16}{\mathcal{M}2^{2\mu}}, \\
& = \sum_{\mu=2}^{n_p+1} [(2^\mu - 1)^3 - (2^{\mu-1} - 1)^3] \frac{16}{\mathcal{M}2^{2\mu}}, \\
& = \frac{4}{\mathcal{M}} (7 \times 2^{n_p+1} - 9n_p - 11 - 3 \times 2^{-n_p}). \quad (113)
\end{aligned}$$

Thus the total error in the approximation of the Hamiltonian due to finite \mathcal{M} can be upper bounded by

$$\begin{aligned}
& \|U + V - (\tilde{U} + \tilde{V})\| \\
& \leq \frac{2\eta}{\pi \mathcal{M} \Omega^{1/3}} \left(\eta - 1 + 2 \sum_{\ell=1}^L \zeta_\ell \right) \\
& \quad \times (7 \times 2^{n_p+1} - 9n_p - 11 - 3 \times 2^{-n_p}). \quad (114)
\end{aligned}$$

This expression then provides the upper bound on the difference in eigenvalues as required. ■

As a result, if we are allowing an error due to the state preparation of $\epsilon_{\mathcal{M}}$, then the value of \mathcal{M} that should be chosen is then about

$$\mathcal{M} \approx \frac{28\eta 2^{n_p}}{\epsilon_{\mathcal{M}} \pi \Omega^{1/3}} (\eta + 2\lambda_\zeta). \quad (115)$$

Note that \mathcal{M} is required to be a power of 2, so $2^{n_{\mathcal{M}}}$. In numerical testing, we find that computing the actual sum gives a result about 50% of the upper bound. This is what would be expected, because the ceiling will add 1/2 on average, which is 50% of the upper bound of 1. If $\alpha \neq 1$ is used, then the error can be reduced further to about 30% of the upper bound, using α in the range $1 - 3/2\mathcal{M}$ to $1 - 1/\mathcal{M}$.

The errors due to approximation of v and approximation of R_ℓ can be upper bounded by the sum of the two sources of error. Calling the U with both approximations \bar{U} , and that with just R_ℓ approximated \tilde{U} , we have $\|\bar{U} - U\| \leq \|\bar{U} - \tilde{U}\| + \|\tilde{U} - U\|$. Here $\|\tilde{U} - U\|$ is upper bounded as in Lemma 2. For $\|\bar{U} - \tilde{U}\|$ exactly the same reasoning as in Eq. (110) holds, except with R_ℓ replaced with \tilde{R}_ℓ . This means that the same error bound holds and so the bounds on the error due to the two sources can be added.

The other main source of error is the error in the phase estimation. With m steps of phase estimation, the rms error in the estimated energy due to the phase estimation is $\pi\lambda/(2m)$. It is possible to use m that is not a power of 2 simply by using unary iteration on the control register

for the phase estimation. The effective value of λ will vary dependent on the relative sizes of the T , U , and V operators.

As discussed above, if $p_v \lambda_T = (1 - p_v)(\lambda_U + \lambda_V)$, then in the case of the failure of the inequality test for the preparation of v , we apply the SEL operation for T , and the net value of λ is $\lambda_T + \lambda_U + \lambda_V$. If $p_v \lambda_T > (1 - p_v)(\lambda_U + \lambda_V)$, we apply SEL for T in the case of failure of the inequality test, *or* $|0\rangle$ on an ancilla qubit, and again $\lambda = \lambda_T + \lambda_U + \lambda_V$. In the case where $p_v \lambda_T < (1 - p_v)(\lambda_U + \lambda_V)$, we apply SEL for T in the case of failure of the inequality test, *and* $|0\rangle$ on an ancilla qubit. In that case, the value of λ is $(\lambda_U + \lambda_V)/p_v$. To account for the three cases, we have

$$\lambda = \max[\lambda_T + \lambda_U + \lambda_V, (\lambda_U + \lambda_V)/p_v]. \quad (116)$$

We can alternatively perform amplitude amplification in the state preparation for $|v\rangle$, which triples that cost, but boosts the probability of success to

$$p_v^{\text{amp}} = \sin^2[3 \arcsin(\sqrt{p_v})], \quad (117)$$

which is close to 1. As a result the value of λ can be given as

$$\lambda = \max[\lambda_T + \lambda_U + \lambda_V, (\lambda_U + \lambda_V)/p_v^{\text{amp}}]. \quad (118)$$

A further subtlety is that we are testing $i \neq j$ in preparing the equal superposition over i and j for V . In that case, if we apply T in the case where this test fails, then the effective λ becomes

$$\lambda = \max\{\lambda_T + \lambda_U + \lambda_V, [\lambda_U + \lambda_V/(1 - 1/\eta)]/p_v\}. \quad (119)$$

The reason for this expression is that the preparation over all values of i and j gives an effective λ_V , which is proportional to η^2 rather than $\eta(\eta - 1)$. This means the effective λ_V is $\lambda_V \eta^2 / [\eta(\eta - 1)] = \lambda_V / (1 - 1/\eta)$. Then, in the case of amplitude amplification on $|v\rangle$, the effective λ would be

$$\lambda = \max\{\lambda_T + \lambda_U + \lambda_V, [\lambda_U + \lambda_V/(1 - 1/\eta)]/p_v^{\text{amp}}\}. \quad (120)$$

See Appendix D for a more in-depth derivation of this expression.

Another adjustment to the value of λ is that needed to account for the probabilities of failure in the preparation of the equal superposition states for w , i , and j . These equal superpositions are required for both parts of the Hamiltonian (kinetic and potential), and so the value of λ is divided by $P_{\text{eq}} = P_s(3, 8) P_s(\eta + 2\lambda_\zeta, b_r) P_s(\eta, b_r)^2$. That gives the resulting values of λ without amplitude

amplification

$$\lambda = \max\{\lambda_T + \lambda_U + \lambda_V, [\lambda_U + \lambda_V/(1 - 1/\eta)]/p_v\}/P_{\text{eq}}, \quad (121)$$

and with amplitude amplification

$$\lambda = \max\{\lambda_T + \lambda_U + \lambda_V, [\lambda_U + \lambda_V/(1 - 1/\eta)]/p_v^{\text{amp}}\}/P_{\text{eq}}. \quad (122)$$

Lastly we note that the λ values are modified by our block-encoding method. We use λ'_T as given by Eq. (71), because of the state preparation used for T . Moreover, because we use an imperfect state preparation for v , the values of

λ_U, λ_V are modified. We define the modified λ_v value

$$\lambda_v^\alpha := \alpha \sum_{v \in G_0} \frac{[\mathcal{M}(2^{\mu-2}/\|v\|)^2]}{\mathcal{M}2^{2\mu-4}}. \quad (123)$$

Then the λ values for U and V are scaled in this way to give

$$\lambda_U^\alpha := \lambda_v^\alpha \lambda_U / \lambda_v, \quad \lambda_V^\alpha := \lambda_v^\alpha \lambda_V / \lambda_v. \quad (124)$$

As a result of these approaches, we have the following theorem for the complexity.

Theorem 4. *The eigenvalue of the Hamiltonian given by Eq. (8) can be estimated to within rms error ϵ using a number of Toffoli gates*

$$\left\lceil \frac{\pi \lambda}{2\epsilon_{pha}} \right\rceil \left\{ 2(n_T + 4n_{\eta\zeta} + 2b_r - 12) + 14n_\eta + 8b_r - 36 + a[3n_p^2 + 15n_p - 7 + 4n_{\mathcal{M}}(n_p + 1)] \right. \\ \left. + \lambda_\zeta + Er(\lambda_\zeta) + 2(2n_p + 2b_r - 7) + 12\eta n_p + 5(n_p - 1) + 2 + 24n_p + 6n_p n_R + 18 \right. \\ \left. + n_{\eta\zeta} + 2n_\eta + 6n_p + n_{\mathcal{M}} + 16 + \mathcal{O}[\log(1/\epsilon)] \right\} \quad (125)$$

where “ a ” is equal to 3 if we perform amplitude amplification and 1 otherwise. For this complexity we require that a phase gradient state and a $|T\rangle$ state are given as a resource. The value of λ is given by

$$\lambda = \max\{\lambda'_T + \lambda_U^1 + \lambda_V^1, [\lambda_U^1 + \lambda_V^1/(1 - 1/\eta)]/p_v^{\text{amp}}\}/P_{\text{eq}} \quad (126)$$

if we perform amplitude amplification, and

$$\lambda = \max\{\lambda'_T + \lambda_U^1 + \lambda_V^1, [\lambda_U^1 + \lambda_V^1/(1 - 1/\eta)]/p_v\}/P_{\text{eq}} \quad (127)$$

otherwise, where

$$p_v = \sum_{\mu=2}^{n_p+1} \sum_{v \in B_\mu} \frac{[\mathcal{M}(2^{\mu-2}/\|v\|)^2]}{\mathcal{M}2^{2\mu}2^{n_\mu+1}}, \quad (128)$$

$$p_v^{\text{amp}} = \sin^2 [3 \arcsin(\sqrt{p_v})], \quad (129)$$

$$P_{\text{eq}} = P_s(3, 8) P_s(\eta + 2\lambda_\zeta, b_r) P_s(\eta, b_r)^2. \quad (130)$$

The value of ϵ_{pha} is the allowable rms error in phase estimation. We require that $\epsilon_{pha} > 0$, $n_R, n_{\mathcal{M}}, n_T \in \mathbb{N}$ are

chosen such that

$$\epsilon^2 \geq \epsilon_{pha}^2 + (\epsilon_{\mathcal{M}} + \epsilon_R + \epsilon_T)^2, \quad (131)$$

$$\epsilon_{\mathcal{M}} = \frac{2\eta}{2^{n_{\mathcal{M}}}\pi\Omega^{1/3}} (\eta - 1 + 2\lambda_\zeta) \\ \times (7 \times 2^{n_p+1} - 9n_p - 11 - 3 \times 2^{-n_p}), \quad (132)$$

$$\epsilon_R = \frac{\eta\lambda_\zeta}{2^{n_R}\Omega^{1/3}} \sum_{v \in G_0} \frac{1}{\|v\|}, \quad (133)$$

$$\epsilon_T = \frac{\pi\lambda}{2^{n_T}}. \quad (134)$$

Proof. The complexity given is the sum of the costs for the block encoding given in the list in Table II, plus the number of qubits that need to be reflected upon given in Eq. (98). The term 18 is to account for the Toffolis used to produce the qubits selecting between T, U, V , and the one flagging success of the state preparation. That takes 9 Toffolis both in the preparation and the inverse preparation. The term $\mathcal{O}[\log(1/\epsilon)]$ is to account for complexity of preparing the initial state to use as control for the phase measurement, and the processing of this control at the end to obtain the phase estimate.

The appropriate values of λ in the cases with and without amplitude amplification are given above in Eqs. (122) and (121). In those expressions we replace $\lambda_T, \lambda_U, \lambda_V$ with

$\lambda'_T, \lambda'_U, \lambda'_V$ to account for our block encoding, where we take $\alpha = 1$ in order to use Lemma 3. The error due to the phase estimation is independent from all other sources of error, so we can upper bound the rms error due to all sources as the sum of squares of the error due to the phase estimation and the error from the other sources. The bound $\epsilon_{\mathcal{M}}$ on the error due to finite \mathcal{M} is obtained from Eq. (115), which is based on Lemma 3. The bound ϵ_R on the error due to the nuclear positions is obtained from Lemma 2. In that expression we substitute $\delta_R = \Omega^{1/3}/2^{n_R+1}$, which comes from the cell length of a side being $\Omega^{1/3}$, and the maximum error in the position being half of $1/2^{n_R}$ times the length of the region. The bound ϵ_T on the error due to the rotation of the qubit selecting between T and $U + V$ is given in Eq. (55).

The phase gradient state and $|T\rangle$ state are used catalytically throughout the procedure, with the phase gradient state being used to implement phase rotations and the $|T\rangle$ state used to implement controlled Hadamards. These states can be prepared at the beginning of the algorithm with negligible complexity compared to the rest of the algorithm. ■

In order to choose parameters for the algorithm, we first take some maximum allowable values for $\epsilon_{\mathcal{M}}$, ϵ_R , and ϵ_T as, for example, $\epsilon/10$. Then we substitute those values into Eqs. (132), (133), and (55) to find minimum values of $n_{\mathcal{M}}$, n_R and n_T . We can then use those values on the rhs of Eq. (132) and Eq. (133) to compute updated (smaller) values of $\epsilon_{\mathcal{M}}$ and ϵ_R . Those values can then be substituted into Eq. (131) to solve for the maximum ϵ_{pha} that satisfies the inequality. These values of the parameters can then be substituted into Eq. (125) to determine the total Toffoli cost. In order to optimize the parameters, instead of using just the values of $n_{\mathcal{M}}$, n_R , and n_T as given by this method, we can

compute the cost for an array of values of $n_{\mathcal{M}}$, n_R , and n_T , for example testing 4 below and 4 above the numbers of bits given by this approach. We can then select from this array the parameters that give the minimum Toffoli complexity.

Another improvement is to compute the exact bound on the state-preparation error, rather than just using the bound given in Lemma 3. That is, we use

$$\epsilon_{\mathcal{M}} = \frac{\eta}{2\pi\Omega^{1/3}} (\eta - 1 + 2\lambda_{\xi}) \sum_{\mu=2}^{n_{\mu}} \sum_{v \in B_{\mu}} \left| \frac{1}{\|v'\|^2} - \frac{1}{\|v\|^2} \right|, \quad (135)$$

with

$$\frac{1}{\|v'\|^2} = \frac{16\alpha}{\mathcal{M}2^{2\mu}} \left[\frac{\mathcal{M}2^{2\mu}}{16\|v\|^2} \right]. \quad (136)$$

Choosing a finite number of bits for n_T corresponds to giving a value of $\alpha \neq 1$. Since the lowest $\epsilon_{\mathcal{M}}$ is obtained in the region $\alpha = 1 - 3/2\mathcal{M}$ to $1 - 1/\mathcal{M}$, we can choose n_T in order to give α in that range and recompute $\epsilon_{\mathcal{M}}$. The value of n_T needed to obtain α in that range is numerically found to be typically around $n_{\mathcal{M}}$ plus 4 to 8. With this choice of n_T and α , $\epsilon_{\mathcal{M}}$ is about 30% of the bound proven in Lemma 3, and there is no additional error due to finite n_T .

III. THE INTERACTION-PICTURE-BASED ALGORITHM

In the interaction-picture approach, one has a time-independent Hamiltonian $H = A + B$, where the norm of B is much smaller than that of A . The time evolution under $A + B$ for time τ can then be given as (with \mathcal{T} the time-ordering operator)

$$\begin{aligned} e^{-i(A+B)\tau} &= e^{-i\tau A} \mathcal{T} \exp \left(-i \int_0^{\tau} ds e^{isA} B e^{-isA} \right), \\ &= \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} \int_0^{\tau} d\tau_1 \int_0^{\tau} d\tau_2 \cdots \int_0^{\tau} d\tau_k e^{-i(\tau-\tau'_k)A} B e^{-i(\tau'_k-\tau'_{k-1})A} B \cdots B e^{-i(\tau'_2-\tau'_1)A} B e^{-i\tau_1 A}, \\ &= \lim_{\substack{K \rightarrow \infty \\ M \rightarrow \infty}} \sum_{k=0}^K \frac{(-i\tau)^k}{M^k k!} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \cdots \sum_{m_k=0}^{M-1} e^{-i\tau(M-1/2-m'_k)A/M} B e^{-i\tau(m'_k-m'_{k-1})A/M} B \cdots \\ &\quad \cdots B e^{-i\tau(m'_2-m'_1)A/M} B e^{-i\tau(m'_1+1/2)A/M}, \end{aligned} \quad (137)$$

where τ'_1, \dots, τ'_k are sorted times from τ_1, \dots, τ_k , and m'_1, \dots, m'_k are sorted integers from m_1, \dots, m_k . The approximation comes from both truncating the sum at K

and discretizing the integrals. The discretization of the integral here is taken by using $\tau_j = (m_j + 1/2)/M$, so the midpoints of the intervals are used. This does not affect the

final expression, except that there are shifts by 1/2 on the beginning and ending rotations.

Note that we do not immediately get a linear combination of unitaries from the above Dyson-series expansion, as the operator B from the Hamiltonian is not unitary in general. Instead, we now describe the Hamiltonian input model for the interaction-picture algorithm in which B is specified. We assume that B is given by a block encoding as

$$\langle 0 | \text{PREP}_B^\dagger \cdot \text{SEL}_B \cdot \text{PREP}_B | 0 \rangle = \frac{B}{\lambda_B}. \quad (138)$$

Here, PREP_B can be viewed as some state-preparation procedure acting nontrivially only on an ancilla register initialized in state $|0\rangle$, and SEL_B is a selection unitary acting jointly on the ancilla register and the register of the target system. By preparing the ancilla state, performing

the selection operation, and unpreparing the state, we can implement operator B on the target system with normalization factor $\lambda_B \geq \|B\|$. We describe a quantum circuit in Sec. C that block encodes B for the first quantized quantum chemistry simulation.

For the operator A , we assume that the matrix exponential e^{-isA} can be efficiently implemented for an arbitrary real number s . This is the case for instance when A is already diagonalized and contains pairwise commuting terms, each of which can be directly exponentiated. In particular, this holds for the kinetic operator T when the first quantized Hamiltonian is represented in the plane-wave basis. See Sec. III B for a detailed discussion of the circuit implementation.

Using the block encoding of B , we now rewrite the Dyson-series expansion as

$$\begin{aligned} e^{-i(A+B)\tau} &\approx \sum_{k=0}^K \frac{(-i\lambda_B\tau)^k}{M^k k!} \langle 0 |^{\otimes k} \text{PREP}_B^{\dagger \otimes k} \sum_{m_1, \dots, m_k=0}^{M-1} \left(e^{-i\tau(M-1/2-m'_k)A/M} \text{SEL}_B e^{-i\tau(m'_k-m'_{k-1})A/M} \text{SEL}_B \dots \right. \\ &\quad \left. \dots \text{SEL}_B e^{-i\tau(m'_2-m'_1)A/M} \text{SEL}_B e^{-i\tau(m'_1+1/2)A/M} \right) \text{PREP}_B^{\otimes k} | 0 \rangle^{\otimes k}, \\ &= \left(\langle 0 | \text{PREP}_B^\dagger \right)^{\otimes K} \sum_{k=0}^K \frac{(-i\lambda_B\tau)^k}{M^k k!} \sum_{m_1, \dots, m_k=0}^{M-1} \left(e^{-i\tau(M-1/2-m'_k)A/M} \text{SEL}_B e^{-i\tau(m'_k-m'_{k-1})A/M} \text{SEL}_B \dots \right. \\ &\quad \left. \dots \text{SEL}_B e^{-i\tau(m'_2-m'_1)A/M} \text{SEL}_B e^{-i\tau(m'_1+1/2)A/M} \right) (\text{PREP}_B | 0 \rangle)^{\otimes K}. \end{aligned} \quad (139)$$

Here for each fixed value of k , the k selection operators SEL_B act on k different ancilla registers and the same target register; the last equality follows from the fact that the remaining $K - k$ ancilla states $\text{PREP}_B | 0 \rangle$ cancel with $\langle 0 | \text{PREP}_B^\dagger$ in pairs and so the number of their occurrences can be made independent of k . Note that the resulting expression is expressed as a linear combination of unitaries and can thus in principle be implemented using quantum circuits.

Although we have written this as performing all the preparations at the beginning and inverting them at the end, in practice we perform the preparations and inverse preparations for each SEL_B before the next preparation. That avoids needing to use all qubits for all preparations at once, which would have a much higher qubit cost. A further simplification is that we can separate out $e^{-i\tau A/2M}$ at the beginning and the end. Then $e^{-i\tau(M-1/2-m'_k)A/M}$ is replaced with $e^{-i\tau(M-1-m'_k)A/M}$. The advantage here is that $M - 1 - m'_k$ can be computed simply by performing NOT gates on all qubits, with zero Toffoli cost.

To implement this linear combination of unitaries, we need to prepare quantum states encoding the coefficients from the combination. For readability, we describe a naive circuit implementation here, and leave a detailed discussion of the subtleties and improvements in subsequent subsections. We first prepare

$$\frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(-i\lambda_B\tau)^k}{k!}} |k\rangle, \quad (140)$$

where β is the normalization factor and $|k\rangle$ is encoded in unary as $|1^k 0^{K-k}\rangle$. This state may be prepared using the naive approach with one rotation gate and K controlled rotation gates. However, we propose an improved approach based on the inequality test, which can significantly reduce the implementation cost; see Sec. III D for details.

Next, we prepare the correct time indices used in the interaction-picture simulation. One possible method is to prepare an equal superposition over k time indices

m_1, \dots, m_k and quantumly sort them [71], giving

$$\frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(-i\lambda_B\tau)^k}{M^k k!}} |k\rangle \sum_{m_1, \dots, m_k=0}^{M-1} |m'_1, \dots, m'_k\rangle, \quad (141)$$

where again $0 \leq m'_1 \leq \dots \leq m'_k \leq M-1$ are sorted integers from m_1, \dots, m_k . We then compute the time differences as

$$\frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(-i\lambda_B\tau)^k}{M^k k!}} |k\rangle \sum_{m_1, \dots, m_k=0}^{M-1} |m'_1, m'_2 - m'_1, \dots, m'_k - m'_{k-1}, M-1 - m'_k\rangle. \quad (142)$$

These difference values are further used to control the matrix exponentials in the Dyson-series expansion. Specifically, we have the decomposition

$$\sum_{m=0}^{M-1} |m\rangle \langle m| \otimes e^{-i\tau mA/M} = \prod_{j=0}^{\log(M)} \left(|0\rangle \langle 0|_j \otimes I + |1\rangle \langle 1|_j \otimes e^{-i\tau 2^j A/M} \right), \quad (143)$$

where each controlled exponential of A can be directly implemented. In practice, in the implementation it is convenient to apply the controlled exponential jointly controlled by all bits of m , rather than controlling on the bits of m independently. On the other hand, we apply the selection unitaries SEL_B controlled by the unary-encoded k values. These operations, together with the preparation and unpreparation procedures, implement the Dyson-series expansion up to the normalization factor β . Beyond the sorting method, previous work also suggested alternative methods based on a ‘‘compressed encoding’’ [71] and ‘‘inequality tests’’ [5]. We compare these methods and justify our choice of implementation in Sec. III A.

For K sufficiently large, the normalization factor β is approximately given by

$$\beta = \sum_{k=0}^K \frac{(\lambda_B\tau)^k}{k!} \approx \exp(\lambda_B\tau). \quad (144)$$

In the original interaction-picture algorithm, we choose $\tau \approx \ln(2)/\lambda_B$ so that $\beta \approx 1/2$. Then, a single step of oblivious amplitude amplification boosts the success amplitude

close to 1. The ideal evolution can be simulated for time t by repeating this $\lambda_B t / \ln(2)$ times. However, for performing phase estimation, we develop a new approach by applying qubitization on the preamplified interaction-picture subroutine, which significantly reduces the implementation cost. We discuss our approach in detail in Sec. III A.

A. Implementation choices for the interaction-picture algorithm

There are two main choices for how to provide the ordered integers m'_j . One is to generate a sequence of differences according to a negative exponential weighting. Then one directly has the differences $m'_{j+1} - m'_j$, but it is necessary to add these differences together to determine which difference gives a total over M . That is used to determine k , rather than generating a register for k directly. This method is called the ‘‘compressed encoding’’ in Ref. [71]. The preparation of the state with the negative exponential weighting can be performed as described in Ref. [77] with a sequence of controlled rotations. That results in state-preparation complexity corresponding to a factor of $\log(1/\epsilon)$ times the number of qubits in the time registers.

The other choice for preparing the state with the ordered integers m'_j is to prepare a set of unordered integers, then sort them. The type of sort needed is a quantum version of a sorting network, as described in Refs. [78,79]. There are algorithms for generating sorting networks, but in practice for small numbers (as we have here) it is best to choose optimized sorting networks. The number of comparator operations required for the best-known sorting networks for up to 16 items are as shown in Table III [80]. Each comparator needed corresponds to an inequality test and a controlled swap. The inequality test uses approximately a Toffoli for each qubit of the registers, and the controlled swap uses a Toffoli for each qubit of the registers. The number of Toffolis therefore corresponds to twice the number of comparators multiplied by the qubits used for each time register. For the case of $n = 16$, that is 7.5 times the number of qubits in the time registers, as compared to $\log(1/\epsilon)$ for the compressed encoding approach. This will be a smaller multiplying factor, which is why we use the sorting approach instead of the compressed encoding.

For the implementation of $e^{-i(A+B)\tau}$ as a linear combination of unitaries, the general method is to prepare a superposition state with weights that are the square roots of the amplitudes in the linear combination of unitaries, then apply controlled unitary operations. Obvious

TABLE III. The number of comparators needed for a quantum sort of K items. Each comparator involves an inequality test and a controlled swap. The number of comparators for a given K is denoted $\text{St}(K)$.

K	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Comparators	1	3	5	9	12	16	19	25	29	35	39	45	51	56	60

amplitude amplification is used in order to implement the step deterministically, and in order to need only a single step of oblivious amplitude amplification it is necessary to choose the parameters so the amplitude of success is $1/2$.

In the case where the sorting is performed, the sum of the amplitudes needed for the approximation of $e^{-i(A+B)\tau}$ as a linear combination of unitaries is

$$\sum_{k=0}^K \frac{\tau^k}{k!} \lambda_B^k \leq \exp(\tau \lambda_B). \quad (145)$$

An inequality is written here due to the truncation at finite K , but the two sides of this equation will be very close (with error corresponding to error in the evolution), so it is sufficient to assume that they are equal for the purpose of costing the algorithm. The amplitude for success in the linear combination of unitaries is the inverse of this value, so in order to obtain amplitude $1/2$, we need $\exp(\tau \lambda_B) \approx 2$, which corresponds to choosing $\tau \approx \ln 2 / \lambda_B$. In order to perform an accurate phase estimation we need long time evolution, so greater efficiency is obtained by the time being broken into intervals τ that are as large as possible.

An alternative approach is given by Ref. [5], where instead of performing a sort, inequality tests are performed in the time registers. When performing the inequality test, the factor of $k!$ is removed in the denominator, because that corresponds to the number of permutations of times. As a result, the sum of amplitudes in the method of Ref. [5] is

$$\sum_{k=0}^K \tau^k \lambda_B^k \leq \frac{1}{1 - \tau \lambda_B}. \quad (146)$$

Again the two sides of the expression are approximately equal. For the amplitude amplification, we need this amplitude to be 2 (because the amplitude for success is the inverse of the sum of amplitudes). That means that in the approach of Ref. [5] one needs $\tau \approx 1/(2\lambda_B)$. This value of τ is about 72% of the value of τ in the method of Ref. [71], which means that the overall complexity is about 39% larger. The method of Ref. [5] saves the complexity of the sort, but that complexity is trivial as compared to other complexities in the algorithm.

So far we assume that phase estimation is performed on the time evolution, which requires performing amplitude amplification. That triples the cost, because even a single step of oblivious amplitude amplification requires a forward step, a reverse step, then a forward step again. An alternative method can be obtained by simply performing phase estimation on the single step of the evolution, as per the discussion in Refs. [8,21]. The difference here is that we are block encoding the evolution $e^{-i(A+B)\tau}$ rather than the Hamiltonian, and the evolution is not Hermitian so the analysis does not hold. However, we can easily block encode $\sin[(A+B)\tau] = [e^{i(A+B)\tau} - e^{-i(A+B)\tau}]/(2i)$.

All we need to do is use a single qubit that chooses between the block encoding of $e^{i(A+B)\tau}$ and $e^{-i(A+B)\tau}$. All that qubit needs to do is control the sign of the exponential in terms like $e^{-i\tau(m'_2 - m'_1)\lambda/M}$, which can be achieved with negligible complexity.

If the operation $\sin[(A+B)\tau]$ has eigenvalues $\mu_j = \sin(E_j \tau)$, where E_j are eigenvalues of $A+B$, then as discussed in Ref. [21] the qubitization procedure gives an operator with eigenvalues $\pm e^{\pm i \arcsin(\mu_j/\lambda)}$, where λ is the sum of the amplitudes in the block encoding of $\sin[(A+B)\tau]$. (In Ref. [21] H is equivalent to $\sin[(A+B)\tau]$ here, and E_k there is equivalent to μ_j here.) Now the value of λ in the block encoding of $\sin[(A+B)\tau]$ is the same as that of $e^{-i(A+B)\tau}$, and so with the method of Ref. [71] is $\exp(\tau \lambda_B)$. The magnitude of the eigenvalues in the exponential are therefore $\arcsin[\sin(E_j \tau) / \exp(\tau \lambda_B)]$. In the case of small E_j (as we expect here), the linearization is approximately $E_j \tau / \exp(\tau \lambda_B)$, so this is equivalent to using an effective evolution time $\tau_{\text{eff}} = \tau / \exp(\tau \lambda_B)$. If we were to choose $\tau \approx \ln 2 / \lambda_B$, then we would have $\tau_{\text{eff}} \approx \ln 2 / (2\lambda_B)$. That is, the evolution time is halved from τ , which corresponds to doubling the cost for the phase estimation. On the other hand, the oblivious amplitude amplification triples the cost, so this qubitization method gives $2/3$ the cost for performing phase estimation on the time evolution. However, we can further reduce the cost by increasing τ to $1/\lambda_B$. That gives an effective time $\tau_{\text{eff}} \approx 1/(e\lambda_B)$ where e is Euler's number. That means the complexity is then 63% of the method via time evolution from Ref. [71], or 45% of the complexity using the method of Ref. [5]. Therefore, we see that even though these individual improvements are only minor, there is total improvement by more than a factor of 2 over the method of Ref. [5].

For this effective time, we are assuming that E_j is near zero, so $\arcsin[\sin(E_j \tau)/e]$ (where we take $\tau = 1/\lambda_B$) is approximately $E_j \tau / e$. In cases where E_j is not near zero, then the uncertainty in estimating $\arcsin[\sin(E_j \tau)/e]$ will correspond to uncertainty in estimating E_j with a different constant factor. However, we have freedom in adding a phase shift to the operator we are implementing. Instead of implementing $e^{\pm i(A+B)\tau}$, we can implement $e^{\pm i(A+B-E_0)\tau}$, simply by performing an additional phase shift. That means we are then estimating eigenvalues $\arcsin\{\sin[(E_j - E_0)\tau]/e\}$. Provided we have an initial estimate of E_j , we can ensure that $E_j - E_0$ is near zero so we are in the region where the sine and arcsine can be linearized.

To be more specific, the slope at $E_j - E_0$ is increased to

$$\frac{\tau}{e} \left(1 + \frac{(e^2 - 1)(E_j - E_0)^2 \tau^2}{2e^2} + \mathcal{O}[(E_j - E_0)^4 \tau^4] \right). \quad (147)$$

This means that there is a relative increase in the number of steps needed to obtain a given accuracy of

$\mathcal{O}[(E_j - E_0)^2 \tau^2]$. There is also a slight increase in the estimation error due to the nonlinearity of this function. This relative increase turns out to be on the order of the standard deviation of the phase estimation, so scales as $\mathcal{O}(1/\mathcal{N})$ for phase estimation on \mathcal{N} steps. This higher-order increase in the error can be accounted for by increasing \mathcal{N} by $\mathcal{O}(1)$. We do not need to take account of these factors when considering phase estimation for the qubitization, because there the effect of the nonlinearity is to reduce the error.

To summarize, there are two main choices for the algorithm. First, there is the choice of how to prepare the time registers, for which there are three main alternatives.

1. The ‘‘compressed encoding’’ method from Ref. [71], which requires controlled rotations by arbitrary angles on the qubits for the time registers, and therefore has a complexity scaling like the logarithm of the inverse error for each qubit of the time registers.
2. The sorting method from Ref. [71], which requires a number of Toffolis for each qubit that can be determined from Table III, and is expected to be around 8 (so less than for compressed encoding).
3. The method of Ref. [5], which involves postselecting on success of inequality tests, rather than a sort. That reduces the complexity of the state preparation for the time registers, but means the number of steps is increased by a factor of $2 \ln 2 \approx 1.38$ over the methods of Ref. [71]. This is expected to result in an increased overall cost, because the state preparation for the time registers is a small part of the overall cost in each step.

The other choice is between the following two alternatives.

1. One can simulate the time evolution under the Hamiltonian, and perform phase estimation on the time evolution. Then the number of steps scales as $\lambda_B T / \ln 2$ for the methods of Ref. [71], or $2\lambda_B T$ for the method of Ref. [5].
2. One can instead perform a qubitization of $\sin(H\tau)$, and perform phase estimation on that. That reduces the complexity because it removes the need for amplitude amplification for time evolution (and the corresponding error). The number of steps would be $e\lambda_B T$ for the methods of Ref. [71] or $2e \ln 2 \lambda_B T$ for the method of Ref. [5]. The number of steps is increased, but the complexity of each step is reduced by a factor of 3 as compared to the time-evolution approach.

We focus on the case where the time registers are prepared using a sort, and we perform qubitization of $\sin(H\tau)$, as this gives the best performance. As in the case of the qubitization approach, we need to ensure that our SEL

procedure is self-inverse. We use an ancilla to control $e^{-i(A+B)\tau}$ versus $e^{i(A+B)\tau}$. To ensure that it is self-inverse, it will be controlled on the ancilla qubit as

$$|0\rangle \langle 1| \otimes e^{-i(A+B)\tau} + |1\rangle \langle 0| \otimes e^{i(A+B)\tau}. \quad (148)$$

This ancilla qubit is prepared as $|+\rangle$, ensuring we obtain the sum of $e^{-i(A+B)\tau}$ and $e^{i(A+B)\tau}$ in the block encoding. However, we do need to ensure that the implementation of $e^{-i(A+B)\tau}$ is the inverse of $e^{i(A+B)\tau}$. To achieve that, we apply the following method.

- (i) Use the qubit as control for a phase gate on each qubit in the unary encoding of k to control between $(-i)^k$ and $(i)^k$ in the Dyson series. There is no extra Toffoli cost.
- (ii) Use the qubit to control whether the sort of the time registers sorts them in ascending or descending order. One can simply use the qubit as a control with the qubits flagging the result of the inequality test as target. There is again no Toffoli cost.
- (iii) Use the qubit to control taking the absolute values of the time differences (see Appendix E). This is the only part with a Toffoli cost.
- (iv) Use this qubit to control between adding and subtracting v , which has no extra Toffoli cost.
- (v) Use the qubit to control positive versus negative phasing by the kinetic energy, which can be performed without Toffolis by controlling between adding and subtracting the phase into the phase gradient state.

B. Exponentiating the kinetic operator

It is estimated in Ref. [42] that the cost of simulating the first quantized electronic structure Hamiltonian is dominated by the exponentiation $e^{-i\tau T}$. The reason for that is the kinetic energy needs to be computed in the computational basis, which requires squaring 3η numbers (for the three components of the momenta of each electron) and adding them together. Squaring an n_p -bit number can be performed with approximately n_p^2 Toffolis, giving an overall complexity of approximately $3\eta n_p^2$.

However, it is possible to greatly reduce the complexity, by keeping the sum of squares in an ancilla, instead of recomputing it every time. Then when we do the block encoding of the $U + V$, we can update this sum of squares. The cost of updating the sum of squares is only logarithmically dependent on η . To be more specific, for V we have p replaced with $p + v$ and q replaced with $q - v$. For the first the change in the sum of squares is $\|p + v\|^2 - \|p\|^2 = 2p \cdot v + \|v\|^2$, and for the second the difference in squares is $\|q - v\|^2 - \|q\|^2 = -2q \cdot v + \|v\|^2$. The quantity $\|v\|^2$ has already been computed in preparing the state with amplitudes $1/\|v\|$, so does not need to be recomputed.

This method also enables us to take account of the energy offset E_0 we are using. We aim to subtract all energies by E_0 to put us in the linear region of the function we are estimating. We can simply subtract E_0 from the kinetic energy register at the beginning of the procedure. That means the energy of every step will be shifted by E_0 with no more gate complexity (other than the initial subtraction).

Therefore the majority of the cost is in the computation of $p \cdot v$ and $q \cdot v$. The cost for computing the product of the *absolute value* of each component is $2n_p(n_p - 1) - n_p = 2n_p^2 - 3n_p$, and the sign of the product can be computed with a CNOT. Rather than adding all three components then adding into the energy register, it is convenient to add them one at a time. This is because we have signed integers, and adding the three signed integers would incur an additional cost for converting to two's complement. When adding the signed integer into the energy register, we can instead use the sign bit to control addition or subtraction into the energy register, which incurs no additional cost.

The number of bits needed to represent the sum of momenta is $n_\eta + 2n_p$. We are guaranteed that the net result of all the arithmetic will result in no overflow qubit, but because we are adding terms in succession there may be overflow in intermediate steps. Therefore the cost of addition of $\|v\|$ will be $n_\eta + 2n_p$. On the other hand, because we have twice $p \cdot v$ and $q \cdot v$, the least significant bit is 0, and the cost of the addition of each component is reduced by 1 to $n_\eta + 2n_p - 1$.

We have qubits that control whether U and V are implemented, and these will need to control whether each of the additions are performed. For $\|v\|^2$ we have a cost of the control each time of $2n_p$, for a total of $4n_p$. To eliminate the cost of controlling the addition of $p \cdot v$ and $q \cdot v$, we can instead make the swaps of these momenta into the ancilla controlled on the qubits, which control whether U and V are implemented. That control on the unary iteration on i and j takes only one extra Toffoli each time, for a total cost of 4 Toffolis there. However, if the swap was not performed, then $p \cdot v$ or $q \cdot v$ was computed with a momentum of zero, giving zero. Therefore these additions do not need to be made controlled.

There is an important nuance in that the phase needed will not be exactly a multiple of $2\pi/2^{b_{\text{grad}}}$ for some integer b_{grad} . The kinetic energy is given by $2\pi^2/\Omega^{2/3}$ times the sum of squares of the integers stored in the momentum registers, and this energy is multiplied by $\tau = 1/(\lambda_U + \lambda_V)$ to give the phase required. The ideal situation is that $2\tau\pi^2/\Omega^{2/3}$ (the smallest phase shift) is exactly $2\pi/2^{b_{\text{grad}}}$. Then one can simply take the sum of squares of momenta and add it into the phase gradient register.

On the other hand, say that for example $2\tau\pi^2/\Omega^{2/3}$ is instead $3 \times 2\pi/2^{b_{\text{grad}}}$. One can then instead use the

modified phase gradient state

$$\frac{1}{\sqrt{2^{b_{\text{grad}}}}} \sum_{\ell=0}^{2^{b_{\text{grad}}}-1} e^{-6\pi i \ell / 2^{b_{\text{grad}}}} |\ell\rangle. \quad (149)$$

Then it is easy to see that adding an integer to the register will give a phase shift that is a multiple of $3 \times 2\pi/2^{b_{\text{grad}}}$. Typically, $2\tau\pi^2/\Omega^{2/3}$ would not be of the form of an integer multiple of $2\pi/2^{b_{\text{grad}}}$, but we have freedom to adjust the time τ . Say for example that we are using a multiple with 10 bits. We use b_T for the number of bits in the multiple. Then the worst case is when $2\tau\pi^2/\Omega^{2/3}$ with $\tau = 1/(\lambda_U + \lambda_V)$ is about 512.5 times $2\pi/2^{b_{\text{grad}}}$. Then rounding to 513 gives the largest relative error, whereas rounding numbers like, for example, 599.5 to 600 would result in smaller relative error. (Numbers below 512 could be represented with fewer bits.) Adjusting the value of τ to make the multiple exactly 513, means that the value of $\tau/\exp(\tau\lambda_B)$ (which governs the number of steps needed in the phase estimation) would be changed by less than 1 part in a million. In general, as discussed in Appendix F, one can take

$$b_{\text{grad}} = b_T - \left\lceil \log \left(\frac{\pi}{(\lambda_U + \lambda_V)\Omega^{2/3}} \right) \right\rceil, \quad (150)$$

and the increase in the λ will be no more than $1 + 1/2^{2b_T+1}$.

To summarize, the costs associated with the phasing by T are as follows.

1. We need to compute time differences from two times of n_t qubits with no carry bit, which has cost $n_t - 1$.
2. There is cost $2n_t(n_\eta + 2n_p) - n_t$ for the multiplication of the sum of momenta by the difference of times.
3. The addition into the phase gradient register has cost $b_{\text{grad}} - 2$ Toffolis.
4. The cost of computing all the products of components of p or q with v is $6 \times (2n_p^2 - 3n_p) = 12n_p^2 - 18n_p$ altogether.
5. The cost of the addition of all these components into the energy register is $6(n_\eta + 2n_p - 1)$.
6. The cost of two additions of $\|v\|^2$ into the energy is $2(n_\eta + 2n_p)$.
7. The cost of making each addition of $\|v\|^2$ controlled is $4n_p$.
8. For the kinetic energy phasing for time $\tau/2M$ at the beginning and the end we simply need to add the (appropriately bit-shifted) kinetic energy into the phase gradient state, with cost $b_{\text{grad}} - 2$ Toffolis each time.

These costs are needed different numbers of times. For the time difference, there is only Toffoli cost where we compute differences of times given in quantum registers,

which is done $K - 1$ times. This is because $M - 1 - m'_k$ can be computed with NOT gates (see Appendix E for more detail). The multiplication in step 2 and the phasing in step 3 are both needed $K + 1$ times for all of the phasings. The remaining costs are updating the energy register, which needs only to be done between phasings, so we have K of these costs.

The arithmetic for the time differences and multiplications can be performed in place with ancillae, so the arithmetic can be inverted without Toffolis. This gives a complexity for each time difference $n_t - 1$. There is also cost $n_t - 1$ to take the absolute value of the time, as controlled by the qubit selecting between forward and reverse time evolution (see Appendix E). With $K - 1$ differences, the total cost is

$$2(K - 1)(n_t - 1). \quad (151)$$

For the phase shift the total Toffoli cost is

$$(K + 1)[2n_t(n_\eta + 2n_p) - n_t + b_{\text{grad}} - 2] + 2(b_{\text{grad}} - 2), \quad (152)$$

where we include $2(b_{\text{grad}} - 2)$ for the initial and final phasings. For updating the sum of momenta

$$\begin{aligned} & K\{[12n_p^2 - 18n_p] + [6(n_\eta + 2n_p - 1)] \\ & + [2(n_\eta + 2n_p)] + [4n_p]\} = K(12n_p^2 + 2n_p + 8n_\eta). \end{aligned} \quad (153)$$

Note that the complexity is now only logarithmically dependent on η . There is still linear dependence on η for selecting momentum registers in block encoding $U + V$.

TABLE IV. The costs involved in the block encoding of the Hamiltonian for the interaction picture, following the same entries as in Table II.

Procedure	Toffoli cost
Preparing the superposition on the qubit selecting between U and V , but not T ; see Eq. (58).	$2(4n_{\eta\zeta} + 2b_r - 9)$
Preparing equal superpositions over η values of i and j in unary; see Eq. (62).	$14n_\eta + 8b_r - 36$
The state preparation cost for the w, r, s registers used for T is eliminated.	0
Controlled swaps of the p and q registers into and out of ancillae (which is used for U and V), with another 4 Toffolis to give overall control; see Eq. (72).	$12\eta n_p + 4\eta - 4$
The SEL cost for T is eliminated.	0
The preparation of the state with amplitude $1/\ \nu\ $ and inversion; see Eq. (90).	$3n_p^2 + 15n_p - 7 + 4n_{\mathcal{M}}(n_p + 1)$
The QROM for R_ℓ , combined with the state preparation with amplitudes $\sqrt{\xi_\ell}$ (which may be implicit); see Eq. (92).	$\lambda_\zeta + \text{Er}(\lambda_\zeta)$
The additions and subtractions of ν into the momentum registers; see Eq. (93).	$24n_p$
Phasing by $-e^{-ik_\nu \cdot R_\ell}$; see Eq. (97).	$6n_p n_R$
Updating the sum of momenta; see Eq. (153).	$12n_p^2 + 2n_p + 8n_\eta$

C. Implementing SEL and PREP

In this section, we discuss the costing of the block encoding of $U + V$ for the interaction picture. As part of this block encoding, we update the sum of squares of momenta, as described above. We have previously described the complexity for block encoding $T + U + V$. The complexity of block encoding $U + V$ is very similar, except the parts needed to block encode T are omitted, and we need to update the register storing the sum of momenta. Therefore our previous costing of the block encoding from Table II can be modified as in Table IV. We make the controlled swaps of momenta into the ancillas controlled on qubits flagging whether U and V are to be performed. That overall control takes one extra Toffoli for each of the controlled swaps, for a total of 4. The final item in this table is the new cost for updating the sum of momenta.

For the total complexity, in the Dyson series the most efficient method is to qubitize a single step of the Dyson series, and perform phase estimation on it, rather than to use amplitude amplification to obtain the explicit time evolution. As a result, the effective time for a single step is approximately $1/[e\lambda_B]$. The rms error in the energy due to phase estimation is then $\pi e\lambda_B/(2\mathcal{N})$, where \mathcal{N} is the number of steps, so one should take

$$\mathcal{N} = \left\lceil \frac{\pi e\lambda_B}{2\epsilon_{\text{pha}}} \right\rceil, \quad (154)$$

where ϵ_{pha} is the allowable error due to phase estimation. For the application to the electronic structure problem, we take $\lambda_B = \lambda_U + \lambda_V$. As described above, we are not using exactly the optimal size of the time step, but the increase

in the cost is upper bounded by $1 + 1/2^{2b_T+1}$, giving

$$\mathcal{N} = \left[(1 + 1/2^{2b_T+1}) \frac{\pi e \lambda_B}{2\epsilon_{\text{pha}}} \right]. \quad (155)$$

D. Total cost of the interaction-picture simulation

In the block encoding of the Dyson series, there are a number of parts that need to be costed.

1. Preparing the superposition over k in unary.
2. The superpositions on the time registers will be prepared with controlled Hadamards.
3. The K registers can be sorted with a number of comparators as given in Table III.

The contributions to the error are as follows.

1. The finite cutoff on the Dyson series, K .
2. The finite angles of the rotations for preparing the superposition over k .
3. The finite number of bits used to represent the times, n_t .

For the time steps we are using for the qubitized Dyson-series approach, the time interval is $\tau = 1/\lambda_B$, which means that the size of term k in the Dyson series is simply $1/k!$, and the error is approximately $1/(K+1)!$. More precisely, the error is

$$e - \sum_{k=0}^K \frac{1}{k!}. \quad (156)$$

This is error in block encoding $\sin[(A+B)\tau] = \sin[(A+B)/\lambda_B]$, and therefore corresponds to error in the estimation of the energy of

$$\lambda_B \left(e - \sum_{k=0}^K \frac{1}{k!} \right) \approx \frac{\lambda_B}{(K+1)!}. \quad (157)$$

Note that here we are assuming that the eigenvalue is near zero, so the sine function can be linearized in estimating the error. In the case where E_j (the energy eigenvalue to be measured) is not near zero, it is possible to just apply a phase shift so we are block encoding $\sin[(A+B-E_0)\tau]$ with E_0 near E_j .

For the error in preparing the superposition over k , we note that because the amplitudes we need are precisely $1/\sqrt{k!}$, the state can be written as proportional to

$$\sum_{k=0}^K \sqrt{\frac{K!}{k!}} |k\rangle, \quad (158)$$

where the amplitudes are the square roots of integers. The $-i$ factors in the Dyson series can be implemented just

with S gates on the qubits with the unary encoding of k , so do not require non-Clifford gates. We also control between forward and reverse time evolution, which means that we control between $\pm i$. That can be achieved simply with a controlled phase gate, and again there is no non-Clifford cost. We can therefore prepare an equal superposition over a number of integers $\Sigma(0) \approx K!e$, where we define

$$\Sigma(k) = \sum_{\ell=k}^K \frac{K!}{\ell!}, \quad (159)$$

and use inequality testing to exactly obtain the required state. The equal superposition over the integers is not performed exactly, but the failure, which has small probability, is flagged so it corresponds only to slightly reducing the size of the block-encoded operation, rather than introducing any error into that block-encoded operation.

For the complexity, the contributions are as follows.

1. First there is the preparation of an equal superposition over a number of integers that is upper bounded by $K!e$. The number of bits can be taken to be $n_k = \lceil \log[\Sigma(0)] \rceil$, so the cost of preparing the equal superposition is (according to the usual formula)

$$3n_k + 2b_r - 9. \quad (160)$$

2. To select $k=0$ or $k=1$, we can perform an inequality test between the equal superposition register and $\Sigma(2)$. The Toffoli cost of this inequality test is $n_k - 1$. To distinguish $k=0$ and $k=1$, we simply use a Hadamard on an ancilla qubit and use a Toffoli to set the qubit.
3. For $1 < k < K$, we perform an inequality test with the number $\Sigma(k+1)$. To save complexity, we can perform only an inequality test with a number of bits necessary to store $\Sigma(k)$, conditioned on already passing the inequality test with $\Sigma(k)$ in the previous step. The Toffoli complexity may then be given as $\lceil \log[\Sigma(k)] \rceil - 1$ for the inequality test, and another Toffoli is needed for the conditioning.

The first contribution here is a complexity that must be paid again in the inverse preparation, giving a total cost

$$2(3n_k + 2b_r - 9). \quad (161)$$

For contributions 2 and 3, the Toffolis and inequality tests may be inverted with Clifford gates, giving a Toffoli cost

$$n_k + \sum_{k=2}^{K-1} \lceil \log[\Sigma(k)] \rceil. \quad (162)$$

In order to ensure that these inequality tests and Toffolis can be inverted with Cliffords, the same number of qubits

are kept throughout the calculation. For a more detailed explanation of the costing, see Appendix E.

The other contributor to the error listed above is the finite number of bits used to represent the time, n_t . In prior work, this error was estimated based on a bound on the derivative of the Hamiltonian. Here we use a tighter bound on the error based on the second derivative. This is a lengthy derivation, so is relegated to Appendix H. The result is that, for time interval τ , the error in the time-ordered exponential can be bounded by

$$\left(\frac{2\|A\|}{\|B\|} + \frac{\|A\|^2}{\|B\|^2} \right) \times \left[2^{n_t} \left(e^{\|B\|\tau/2^{n_t}} - 1 \right) - \|B\|\tau(1 + \|B\|\tau/2^{n_t+1}) \right]. \quad (163)$$

Here we take $\tau = 1/(\lambda_U + \lambda_V)$, and replace $\|B\|$ with $\lambda_U + \lambda_V$ and $\|A\|$ with λ_T . That gives the bound

$$\left(\frac{2\lambda_T}{\lambda_U + \lambda_V} + \frac{\lambda_T^2}{(\lambda_U + \lambda_V)^2} \right) \times \left[2^{n_t} \left(e^{2^{-n_t}} - 1 \right) - (1 + 2^{-(n_t+1)}) \right]. \quad (164)$$

Again, this is error in block encoding of $\sin[(A+B)/\lambda_B]$, and corresponds to error in the estimation of the energy that is multiplied by $\lambda_B = \lambda_U + \lambda_V$, so the error in the energy is

$$\left(2\lambda_T + \frac{\lambda_T^2}{\lambda_U + \lambda_V} \right) \left[2^{n_t} \left(e^{2^{-n_t}} - 1 \right) - (1 + 2^{-(n_t+1)}) \right]. \quad (165)$$

The complexity of the controlled Hadamards to generate the superpositions over times is therefore Kn_t . The same complexity is needed in inverting the state preparation, giving total cost

$$2Kn_t. \quad (166)$$

Sorting the times with n_t bits gives complexity $2n_t$ times the numbers of comparators given in Table III. These numbers of comparators are denoted $\text{St}(K)$, and the same cost is needed to invert the sort, giving a total sorting cost

$$4n_t \text{St}(K). \quad (167)$$

For the times, we need to compute the differences. There is a subtlety here, in that we need to compute the difference with the maximum time for the final time difference for a given k . That can be achieved by taking all times where we have *not* performed the Hadamards to be set to the maximum time (all ones). One can simply perform X gates after the controlled Hadamards. The total cost of the

differences and absolute values was then given before as $2(K-1)(n_t-1)$.

That accounts for the need to make the exponentiation of the kinetic operator controlled on the k register, but we also need to do this for the block encoding of $U+V$. The block encoding is a combination of the PREP operation, a SEL operation, and an inverse preparation PREP^\dagger . Because the preparation and inverse preparation cancel, only the SEL operation need be made controlled. The cost of this control may be quantified as follows.

First, we need to control on success of the three equal superposition state preparations and the preparation of v , as well as the qubit in the k register. That takes 4 Toffolis to produce a qubit that is the AND of all five bits. We apply another 4 Toffolis in erasing this qubit (though we could also keep the ancillas to erase it without Toffolis).

For the phasing by $-e^{-ik_v \cdot R_\ell}$, the R_ℓ can be made zero in the case where we do not want to perform SEL by simply controlling the QROM for outputting of R_ℓ on that qubit. That QROM is already controlled on the qubit selecting between U and V , and controlling on another qubit merely needs one more Toffoli. There also needs to be control of the minus sign on that qubit as well. However, in controlling the output of the QROM, we are generating an ancilla that can be used to control that sign flip, and no additional Toffolis are required.

The addition $p+v$ is already controlled by the qubit selecting between U and V . Again, making it controlled on one more qubit only takes one more Toffoli. We cost the subtraction $q-v$ as being controlled, but for the interaction picture that subtraction only needs to be controlled on this one qubit. We have already given the cost for this subtraction to be controlled, so there is *no* additional Toffoli cost over what we have already presented. As a result, the total additional complexity to make the block encoding of $U+V$ controlled on a qubit is just *two* Toffolis, which is completely trivial compared to the other costs in the circuit.

In the block encoding, we also need to reflect on all qubits used that are not guaranteed to be returned to zero. The detailed analysis is given in Appendix C, and yields a Toffoli cost of

$$K(n_{\eta\zeta} + 2n_\eta + 4n_p + n_{\mathcal{M}} + n_t + 12) + n_k + 3. \quad (168)$$

In this costing, we use intermediate checks between the block encodings of $U+V$ to ensure that the control qubits are rezeroed, and we can reuse them. Lastly we provide the total cost of a single step of the interaction picture including inversion in Table V. There, 10 Toffolis are added to the costs from Table IV to account for the complexity of making the operations for U and V controlled.

To determine the effective λ , we need to also account for the probabilities of success of the amplitude amplification for the various state preparations we have performed. For the block encoding of $U+V$, these are as follows.

TABLE V. The costs involved in the block encoding of the interaction-picture representation of the Hamiltonian.

Procedure	Toffoli cost
The preparation of the equal superposition over $\Sigma(0)$ numbers; see Eq. (161).	$2(3n_k + 2b_r - 9)$
Inequality tests used for the preparation of the superposition over k ; see Eq. (162).	$n_k + \sum_{k=2}^{K-1} \lceil \log(\Sigma(k)) \rceil$
Controlled Hadamards for preparing the superpositions over times; see Eq. (166).	$2Kn_t$
The sorting of the time registers; see Eq. (167).	$4n_t \text{St}(K)$
The differences of the times; see Eq. (151).	$2(K-1)(n_t - 1)$
The kinetic operator needs to be exponentiated $K+1$ times; see Eq. (152).	$(K+1)[2n_t(n_\eta + 2n_p) - n_t + b_{\text{grad}} - 2] + 2(b_{\text{grad}} - 2)$
Perform the controlled block encoding of the potential operator K times.	$K(10 + \text{costs in Table IV})$
The reflection cost for qubitization.	Eq. (168)

1. The preparation of the state with amplitudes $1/\|v\|$, which has probability of success denoted p_v^{amp} .
2. The preparation of the equal superposition state for selecting between U and V , which has probability $P_s(\eta + 2\lambda_\xi, b_r)$.
3. Preparing the equal superpositions over i and j have probabilities of success $P_s(\eta, b_r)$ each.
4. The probability of obtaining $i \neq j$ is $1 - 1/\eta$.

The value of $\lambda_U + \lambda_V$ needs to be divided by all the probabilities in parts 1 to 3, and just λ_V needs to be divided by $1 - 1/\eta$. It is also possible to perform a joint state preparation over i and j , eliminating the need to divide by $1 - 1/\eta$. We do not explain that here as the procedure is already quite complicated.

In performing the block encoding of $\sin[(U + V)\tau]$ there is a further preparation of an equal superposition state in order to prepare the superposition over k . That has probability of success $P_s[\Sigma(0), b_r]$. Because we are considering a region where the sine can be linearized, the effect is the same as multiplying $U + V$ by this probability, and we should therefore multiply all these probabilities. Also taking into account the factor $(1 + 1/2^{2b_T+1})$ due to adjusting the length of time, we get

$$P_{\text{eq}} = p_v^{\text{amp}} P_s[\Sigma(0), b_r] P_s(\eta + 2\lambda_\xi, b_r) P_s(\eta, b_r)^2 / (1 + 1/2^{2b_T+1}). \quad (169)$$

The value of $\lambda_U + \lambda_V$ then needs to be divided by P_{eq} to obtain the effective λ . Therefore, to get the total cost, the costs for a single step of the interaction picture need to be multiplied by a number of steps

$$\mathcal{N} = \left\lceil \frac{\pi e[\lambda_U + \lambda_V/(1 - 1/\eta)]}{2\epsilon_{\text{pha}} P_{\text{eq}}} \right\rceil. \quad (170)$$

Now we take account of the nonlinearity in using an operator with eigenvalue $\arcsin\{\sin[(E_j - E_0)\tau]/e\}$. As discussed above, the error of the estimation is increased by a relative amount $\mathcal{O}[(\lambda_U + \lambda_V)^2 \Delta E^2]$, so we need to introduce this factor to the number of steps used. Also, the nonlinearity gives an $\mathcal{O}(1)$ increase to the number of steps needed. If we have an initial estimate of the energy with precision ΔE , we can choose E_0 such that the number of steps needed is

$$\mathcal{N} = \frac{\pi e[\lambda_U + \lambda_V/(1 - 1/\eta)]}{2\epsilon_{\text{pha}} P_{\text{eq}}} \times \{1 + \mathcal{O}[(\lambda_U + \lambda_V)^2 \Delta E^2]\} + \mathcal{O}(1). \quad (171)$$

In this expression, we should also use the values $\lambda_U^\alpha, \lambda_V^\alpha$ given by the block encoding to be precise, though in practice the difference should be negligible.

In computing these costs, we need to ensure that the total error is no larger than ϵ , where the sources of error are as follows.

1. The error due to the phase estimation ϵ_{pha} .
2. The error due to the truncation of the Dyson series, $\epsilon_K \approx (\lambda_U + \lambda_V)/(K+1)!$.
3. The error ϵ_R due to the finite number of bits used to represent the nuclear positions, which is upper bounded as in Lemma 2.
4. The error $\epsilon_{\mathcal{M}}$ due to the preparation of the state with amplitudes $1/\|v\|$, which is upper bounded as in Lemma 3.
5. The error due to the finite number of time steps, ϵ_t , which is upper bounded as in Eq. (165).

Because the error ϵ_{pha} is a rms value, we combine it with the other errors as a square root of a sum of squares. We therefore require that

$$\epsilon_{\text{pha}}^2 + (\epsilon_K + \epsilon_R + \epsilon_{\mathcal{M}} + \epsilon_t)^2 \leq \epsilon^2. \quad (172)$$

To summarize, the result can be given as in the following theorem.

Theorem 5. *Using the interaction picture, assuming an initial estimate with accuracy ΔE , the eigenvalue of the Hamiltonian given by Eq. (8) can be estimated to within rms error ϵ , using a number of Toffoli gates*

$$\begin{aligned} \mathcal{N} & \left(2(3n_k + 2b_r - 9) + n_k + \sum_{k=2}^{K-1} [\log[\Sigma(k)]] + 2Kn_t + 4n_t \text{St}(K) \right. \\ & + 2(K-1)(n_t - 1) + (K+1)[2n_t(n_\eta + 2n_p) - n_t + b_{\text{grad}} - 2] + 2(b_{\text{grad}} - 2) \\ & + K \left\{ 10 + 2(4n_{\eta\zeta} + 2b_r - 9) + (14n_\eta + 8b_r - 36) + 3[3n_p^2 + 15n_p - 7 + 4n_{\mathcal{M}}(n_p + 1)] \right. \\ & + \lambda_\zeta + E_r(\lambda_\zeta) + (12\eta n_p + 4\eta - 4) + 24n_p + 6n_p n_R + (12n_p^2 + 2n_p + 8n_\eta) \left. \right\} \\ & \left. + K(n_{\eta\zeta} + 2n_\eta + 2 + 4n_p + n_{\mathcal{M}} + n_t + 12) + n_k + 3 + \mathcal{O}[\log(1/\epsilon)] \right), \end{aligned} \quad (173)$$

where

$$\mathcal{N} = \frac{\pi e[\lambda_U^1 + \lambda_V^1/(1-1/\eta)]}{2\epsilon_{\text{pha}} P_{\text{eq}}} \{1 + \mathcal{O}[(\lambda_U + \lambda_V)^2 \Delta E^2]\} + \mathcal{O}(1), \quad (174)$$

$$P_{\text{eq}} = p_v^{\text{amp}} P_s[\Sigma(0), b_r] P_s(\eta + 2\lambda_\zeta, b_r) P_s(\eta, b_r)^2 / (1 + 1/2^{2b_T+1}), \quad (175)$$

$$b_{\text{grad}} = b_T - \left\lceil \log \left(\frac{\pi}{(\lambda_U + \lambda_V)\Omega^{2/3}} \right) \right\rceil, \quad (176)$$

with $b_T \in \mathbb{N}$. For this complexity we require appropriately chosen phase gradient states to be given as a resource. The quantities $\epsilon_{\text{pha}} > 0$ and $K, n_R, n_{\mathcal{M}}, n_T \in \mathbb{N}$ are chosen such that

$$\epsilon^2 \geq \epsilon_{\text{pha}}^2 + (\epsilon_K + \epsilon_R + \epsilon_{\mathcal{M}} + \epsilon_t)^2, \quad (177)$$

$$\epsilon_K = (\lambda_U + \lambda_V) \left(e - \sum_{k=0}^K \frac{1}{k!} \right), \quad (178)$$

$$\epsilon_{\mathcal{M}} = \frac{2\eta}{2^{n_{\mathcal{M}}}\pi\Omega^{1/3}} \left(\eta + 2 \sum_{\ell=1}^L \zeta_\ell \right) (7 \times 2^{n_p+1} - 9n_p - 11 - 3 \times 2^{-n_p}) \quad (179)$$

$$\epsilon_R = \frac{\eta\lambda_\zeta}{2^{n_R}\Omega^{2/3}} \sum_{v \in G_0} \frac{1}{\|v\|}, \quad (180)$$

$$\epsilon_t = \left(2\lambda_T + \frac{\lambda_T^2}{\lambda_U + \lambda_V} \right) \left[2^{n_t} (e^{2^{-n_t}} - 1) - (1 + 2^{-(n_t+1)}) \right]. \quad (181)$$

Proof. The Toffoli costs given in this theorem are simply the sum of the costs found above in Table V. We take ϵ_{pha} , ϵ_K , ϵ_R , $\epsilon_{\mathcal{M}}$, and ϵ_t to be the allowable errors from phase estimation, truncation of the Dyson series at order K , the finite approximation of the positions of the nuclei, preparation of the state with amplitudes $1/\|v\|$, and finite number of bits for the time, respectively. Because the error in the

phase estimate is independent of all other errors, the sum of the squares of this error and all other errors needs to be smaller than ϵ^2 . The expression for ϵ_K is from Eq. (157), the expression for ϵ_R is from Lemma 2, the expression for $\epsilon_{\mathcal{M}}$ is from Lemma 3, and the expression for ϵ_t is obtained from the upper bound in Eq. (165). The probability of success of the state preparations is given by Eq. (169), and the

number of repetitions needed for the state preparation is given by Eq. (171). We use $\lambda_U^\alpha, \lambda_V^\alpha$ with $\alpha = 1$ so we can use Lemma 3. The term $\mathcal{O}[\log(1/\epsilon)]$ is to account for complexity of preparing the initial state to use as control for the phase measurement, and the processing of this control at the end to obtain the phase estimate. ■

As discussed previously for the case of qubitization, the general way to use this theorem is to choose some allowable maxima for the errors, compute the appropriate values of the parameters required to obtain those errors, then compute the actual error to obtain $\epsilon_K, \epsilon_M, \epsilon_R$, and ϵ_t . Those are then used in Eq. (177) to obtain ϵ_{pha} , which is then used to obtain the total complexity. The errors ϵ_M and ϵ_R were discussed above for Theorem 4. For the case of ϵ_K , we can choose K proportional to $\ln[(\lambda_U + \lambda_V)/\epsilon_V]/\ln[(\lambda_U + \lambda_V)/\epsilon_V]$. For ϵ_t we can use the approximation that the expression in the square brackets is about $2^{-2n_t}/6$ to estimate the appropriate value of n_t . The value of b_T is chosen to be an integer, usually around 8. We can then adjust the values used for n_M, n_R, K, n_t , and b_T to minimize the complexity.

IV. RESULTS AND DISCUSSION

A. The resources required to simulate real systems

In this section we numerically compute and compare the costs of the central algorithms of the preceding two sections to each other, and to other methods in the literature. The constant factor Toffoli complexity is given for our qubitization approach by Theorem 4 and given for our interaction-picture approach by Theorem 5. While those expressions are rather complex, the simpler asymptotic complexities are given in Eqs. (26) and (33), respectively. We see that we can characterize the complexity of these approaches in terms of η, N, Ω , and ϵ . The number of nuclear charges L also appears in the scaling but only inside of subdominant terms that do not significantly impact the complexity. Thus, our algorithm has no leading-order dependence on the location, magnitude, or number of nuclear charges. This is a stark contrast to algorithms based on Gaussian orbitals where the basis itself, and hence the λ values as well as the cost of the quantum walk steps, are greatly influenced by the particular details of the system being simulated. Because our costs do not depend strongly on such properties, we are able to analyze the scaling of our algorithms much more broadly than is possible with the molecular orbital based algorithms.

While the logarithmic factors in the complexity sometimes depend on N without Ω (a result of the $\log N$ sized ancilla and momenta registers), in the leading-order complexity we see that N and Ω always appear together as a ratio. For example, assuming fixed precision (in these numerics we target chemical accuracy, $\epsilon = 0.0016$ Hartree) we can express the leading-order algorithmic gate

complexities as

$$\begin{aligned} & \text{gates (qubitization)} \\ &= \tilde{\mathcal{O}} \left(\frac{\eta^2}{\Delta^2} + \frac{\eta^3}{\Delta} \right) \quad \text{gates (interaction pic)} \\ &= \tilde{\mathcal{O}} \left(\frac{\eta^3}{\Delta} \right) \quad \Delta = \left(\frac{\Omega}{N} \right)^{1/3}, \end{aligned} \quad (182)$$

where we define Δ as the spacing of the plane-wave reciprocal lattice. Intuitively, Δ can be thought of as the length scale of resolution within our basis (e.g., wavefunction features smaller than Δ cannot be resolved but those larger than Δ can be resolved). It is easy to see that the $1/\Delta$ dependence arises from the maximum energy of the Coulomb operator between grid points separated by Δ whereas the $1/\Delta^2$ dependence arises due to the maximum energy of the kinetic operator, which is proportional to the square of velocities on such a grid. If one desires to model a nonperiodic system using plane waves then the typical strategy is to choose a sufficiently large box size Ω such that the periodic images of the Coulomb operator do not interact [81]. At first glance, the larger one makes the box, the cheaper our algorithms due to the inverse dependence on Ω . However, as Ω gets larger the resolution Δ decreases, and thus to maintain reasonable precision one must also increase N . By expressing the complexity as in Eq. (182) we see that if the goal is to perform a calculation with fixed spatial resolution Δ , the exact size of the box will not substantially impact the complexity.

In Fig. 1 we compare the complexities of the two algorithms of this paper by plotting the total Toffolis required against η for several different values of Δ . While the interaction-picture algorithm has slightly better scaling in terms of Δ , the constant factors are meaningfully worse. To make the comparison between the complexities more directly, in Fig. 2 we show both the ratio of the Toffoli complexities and the parameter regimes where each approach is more efficient. It would appear that the qubitization algorithm has lower cost in almost all regimes of interest. The exception would be for *extremely* high-resolution simulations where both η and Δ are on the smaller side of what we would imagine requiring (there is little dependence on N). For example, the interaction-picture algorithm is cheaper when $\eta = 20$ and $\Delta = 10^{-3}a_0$. While such precisions might be appropriate for performing dynamics, we do not anticipate that those precisions will typically be required for preparing and sampling observables from molecular eigenstates. In terms of spatial resolution, $\Delta = 10^{-3}a_0$ is quite high since it corresponds to a reciprocal lattice grid with a billion lattice points per Bohr radius cubed. Of course, at higher values of η or Δ the term with η^3/Δ scaling dominates in the qubitization cost, which gives a scaling that is similar to that of the interaction-picture algorithm but with lower constant factors. *A priori*, it was

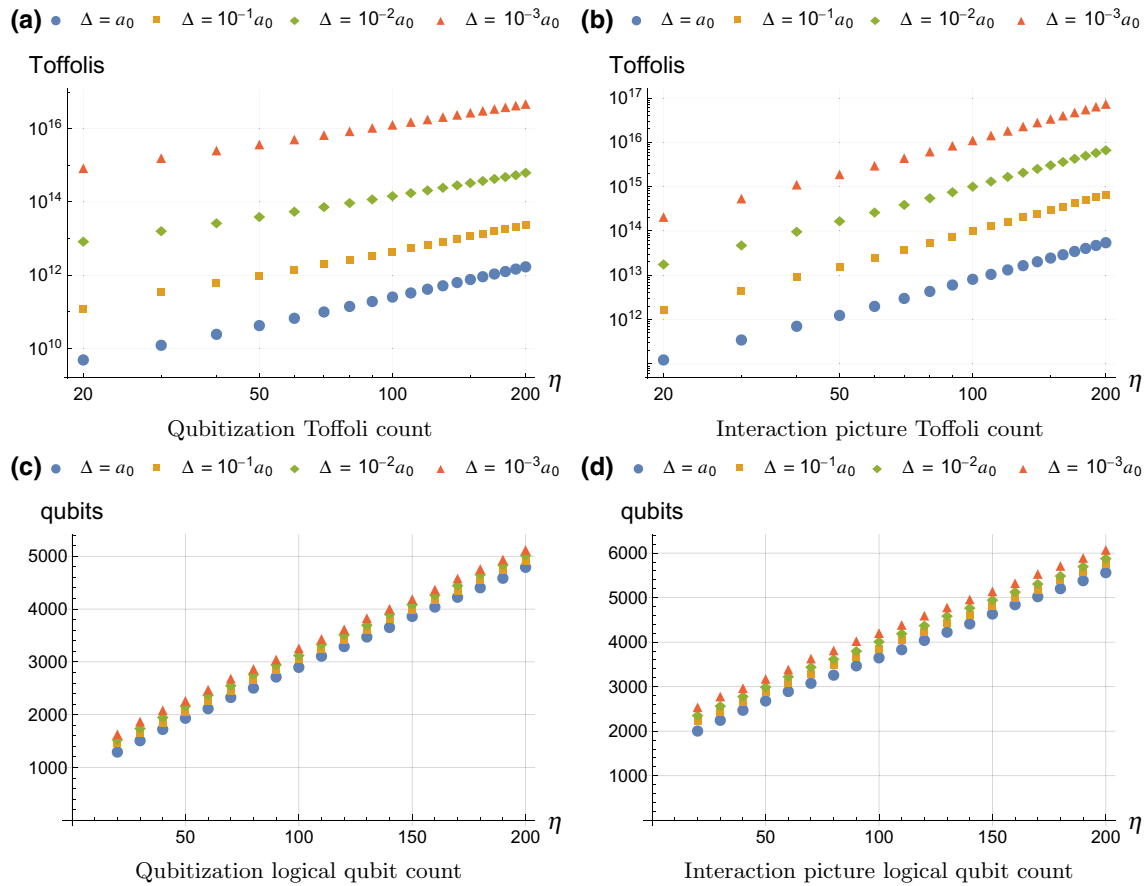


FIG. 1. Toffoli (top) and logical qubit (bottom) counts of the qubitization algorithm (left) and interaction-picture algorithm (right) as a function of η . Here we take $\epsilon = 0.0016$ Hartree (corresponding to “chemical accuracy”) and $N = 2^{18} \approx 2.6 \times 10^5$ but we note that outside of the dependence on Δ , algorithm costs depend only logarithmically on N . For simplicity these plots correspond to the uniform electron gas (i.e., $U = 0$) but block encoding the external potential adds only marginal overhead so we expect these costs to be broadly representative of the resources required to simulate arbitrary molecular systems as a function of η and Δ . We emphasize that $\Delta = 10^{-3}a_0$ is an excessively accurate discretization that would correspond to a grid of one billion points per each Bohr radius cubed of system volume. In most cases, we expect that a resolution of $\Delta = 10^{-2}a_0$ (which is one million grid points per Bohr radius cubed) would correspond to more accurate energies than large correlation consistent Gaussian basis sets. In the bottom panels we see that the qubit counts of the two algorithms are similar.

not obvious to us that the qubitization algorithm would be more efficient than the interaction-picture algorithm for modest values of η . Thus, the fact that qubitization is usually more practical for chemistry problems, is an important finding of our work.

Next, we turn to the topic of how our algorithm scales for the important application of simulating periodic systems such as crystals, metals, surfaces and polymers. Such systems are ubiquitous in nature and are a major focus of electronic structure methods deployed to materials science. While it is important to understand how our algorithm scales as one approaches the continuum basis set limit where $\Delta \rightarrow 0$ (as discussed in the context of Fig. 1) for condensed phase applications it is also important to understand how our algorithm scales towards the thermodynamic limit. That is, the computational cell size should be larger than the correlation length in order for the

simulation to faithfully capture the behavior of the infinite system. The computational cell will be composed of multiple unit cells (the simplest repeating unit in the crystalline lattice). For such simulations, one can imagine adding more unit cells to the computational cell but keeping the overall density of particles (i.e., η/Ω) fixed. Often, behavior in the thermodynamic limit is extrapolated from a series of progressively larger calculations. In order to discuss simulations with fixed η/Ω in more intuitive terms, we analyze scaling in terms of the Wigner-Seitz radius, r_s (a common measure of a system’s average radius between electrons), defined through the relation

$$\frac{4}{3}\pi r_s^3 = \frac{\Omega}{\eta}, \quad r_s = \left(\frac{3\Omega}{4\pi\eta}\right)^{1/3}. \quad (183)$$

We report Wigner-Seitz radii for several common materials in Table VI and note that they are mostly between

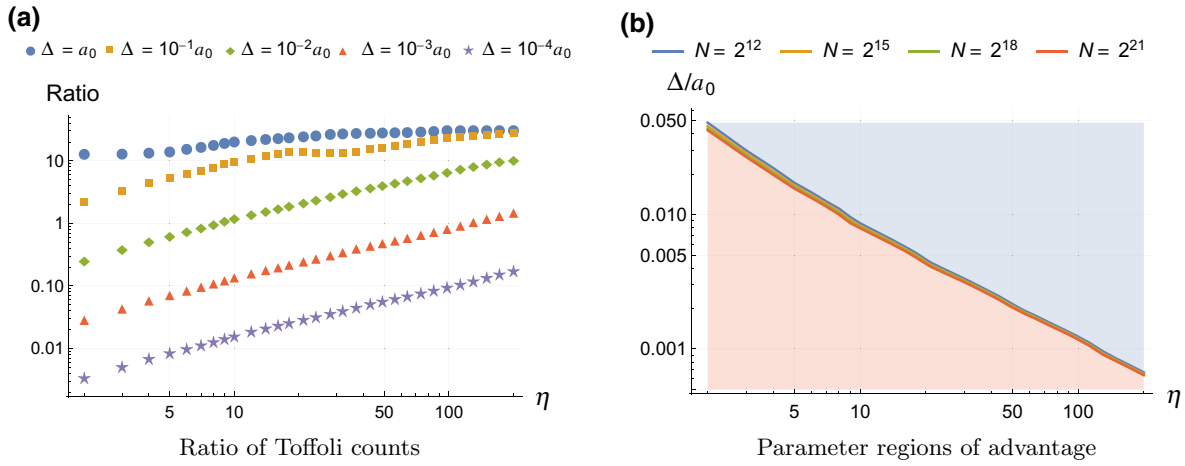


FIG. 2. (a) The ratio of the number of Toffolis for the interaction-picture approach to the number of Toffolis for the qubitization approach. We take $\epsilon = 0.0016$ Hartree (corresponding to “chemical accuracy”) and $N = 2^{18} \approx 2.6 \times 10^5$. We again note that outside of the dependence on Δ , algorithm costs depend only logarithmically on N . For simplicity these plots correspond to the uniform electron gas (i.e., $U = 0$) but block encoding the external potential adds only marginal overhead so we expect these costs to be broadly representative of the resources required to simulate arbitrary molecular systems as a function of η and Δ . (b) The parameter region of Δ versus η where there is an advantage for the Toffoli complexity for the interaction-picture approach versus qubitization. The region where the interaction picture has the advantage is shown in red, and the region where qubitization has the advantage is shown in blue. That is, the lower left of (b) is showing the values of Δ for which the cost ratio plotted in (a) is less than one, as a function of η . In this figure we show the crossover for four different values of N , demonstrating that it barely changes as a function of N . In both figures we see that the interaction-picture approach only has an advantage for low particle numbers and/or when simulations target unusually high precision.

one and a few Bohr radii. In terms of Wigner-Seitz radii we can express the overall Toffoli complexities as

$$\begin{aligned}
 & \text{Toffolis (qubitization)} \\
 &= \tilde{\mathcal{O}}\left(\frac{\eta^{4/3}N^{2/3}}{r_s^2} + \frac{\eta^{8/3}N^{1/3}}{r_s}\right) \text{ Toffolis (interaction pic)} \\
 &= \tilde{\mathcal{O}}\left(\frac{\eta^{8/3}N^{1/3}}{r_s}\right). \tag{184}
 \end{aligned}$$

TABLE VI. Values of η , Ω , and r_s for a variety of common materials. Above, a_0 is the Bohr radius of the hydrogen atom and thus r_s values are in units of Bohr and Ω values are in units of Bohr³. Often one would want to perform a simulation with multiple unit cells in order to resolve correlations in between the units. Doing this would not change r_s but it would increase η proportional to the number of unit cells included. The valence η and r_s values correspond to simulations with core electrons removed. Doing this would require the use of pseudopotentials, which is beyond the scope of the current paper but likely an important next step. Note further that the unit-cell volumes tabulated here correspond to conventional unit cells, rather than primitive unit cells. Using primitive unit cells would necessitate adapting our approach to reciprocal spaces beyond the cubic Bravais lattices considered here.

System	Conventional cell type	Ω/a_0^3	η (cell total)	r_s/a_0 (total)	η (cell valence)	r_s/a_0 (valence)
Metallic lithium	Body-centered-cubic	284.94	6	2.25	2	3.24
Metallic potassium	Body-centered-cubic	961.67	38	1.82	2	4.86
Diamond	Diamond cubic	307.04	48	1.15	32	1.32
Crystalline silicon	Diamond cubic	1080.43	112	1.32	32	2.01
Iron (II) oxide	NaCl structure	539.84	136	0.98	52	1.35
Cobalt oxide	NaCl structure	522.81	140	0.96	60	1.28
Aluminium arsenide	Zinblende	1197.86	184	1.16	32	2.08
Indium phosphide	Zinblende	1364.93	256	1.08	32	2.17

In Figs. 3 and 4 we plot the estimated resources required to perform condensed phase simulations for the qubitization and interaction-picture algorithms, respectively. We make these plots between $\eta = 20$ and $\eta = 200$, between $N = 2^{12}$ and $N = 2^{21}$ and between $r_s = 0.1$ and $r_s = 100$. We observe that the qubitization algorithm is most practical due to lower constant factors across all regimes plotted. As before, in the limit of large η , the qubitization algorithm also recovers the better scaling of the interaction-picture algorithm (but with better constant factors). By comparing

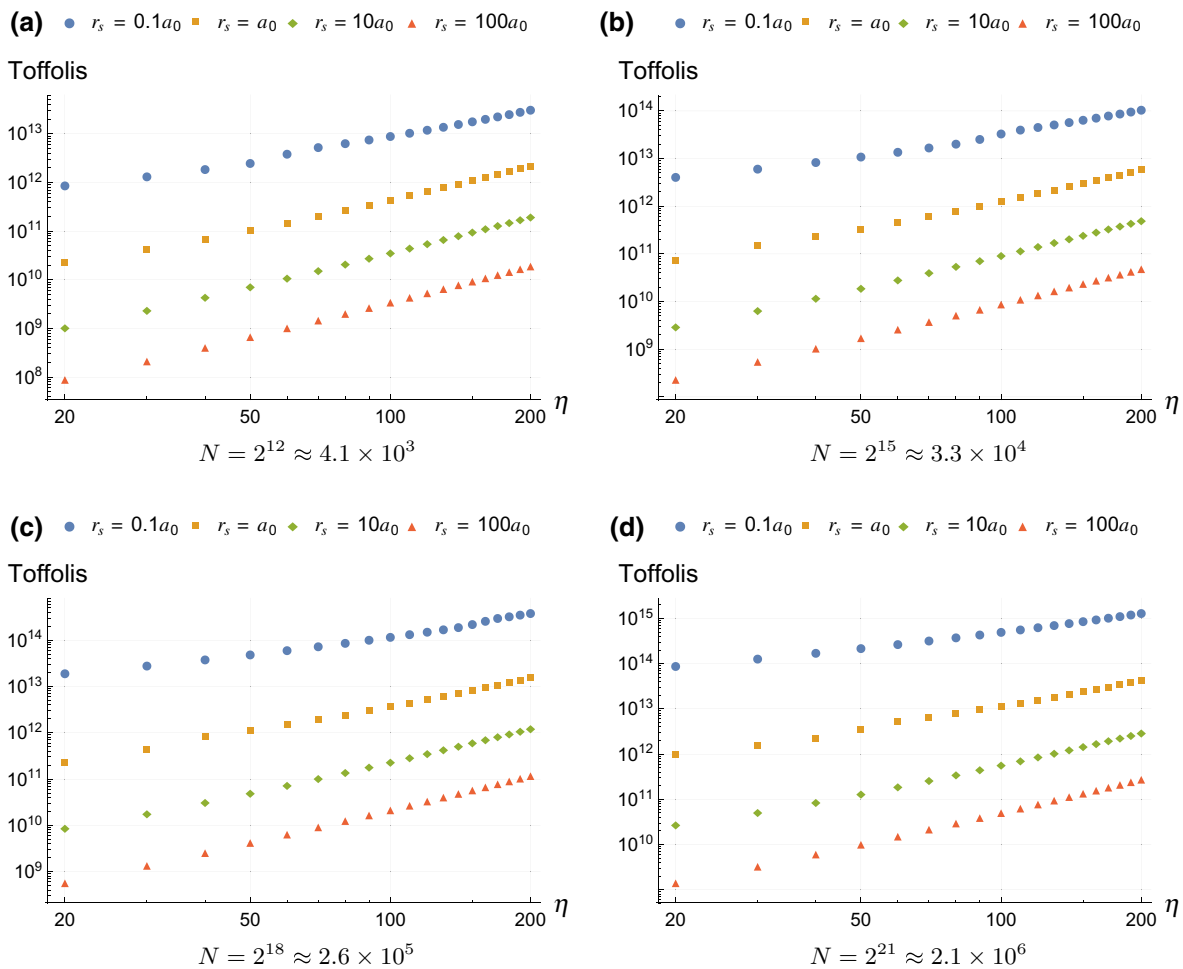


FIG. 3. Toffoli count of qubitization as a function of η using the improved techniques of this paper. In contrast to Fig. 1, we plot these costs for fixed values of the Wigner-Seitz radius r_s to give a sense of the scaling when η increases proportionally to Ω (as would be the case when growing condensed phase simulations towards their thermodynamic limit). Thus, we expect these plots to be representative of costs when simulating periodic materials. We see that the qubitization-based algorithm performs better across all regimes studied to compared to the interaction-picture-based algorithm (Fig. 4).

Fig. 3 with the requirements to simulate various materials in Table VI we see that our method can simulate classically intractable real material Hamiltonians with between 10^{10} and 10^{12} Toffoli gates. These numbers of Toffoli gates and qubits are comparable to those (for a different system) in Ref. [27], where estimates are given of the resources needed to implement the algorithm on a real device. If we adopt the same assumptions here, including a surface-code cycle time of $1 \mu\text{s}$, our proposed quantum algorithms would require on the order of a few million physical qubits and a few days of runtime. No prior quantum algorithm for chemistry has been shown to require anywhere near this few resources to simulate realistic descriptions of condensed phase molecular systems.

We now address the question of how the more practical of the two algorithms described in this work (the

qubitization one) compares to other methods for quantum-simulating chemistry. To date, the only prior methods that have been assessed for viability within fault tolerance have used second quantization. These papers fall into two categories: those that focus on plane waves (see Table I) and those that focus on Gaussian orbitals. The plane-wave algorithms assessed for fault tolerance are by Babbush *et al.* [20] and Kivlichan *et al.* [24]. Because the present work also uses plane waves it is straightforward to make a comparison. As discussed in Sec. I A, for simulating molecules these prior papers using plane waves are not particularly practical options because at least tens or hundreds of thousands of plane waves would be required for reasonably accurate simulations and those works have a space complexity scaling linearly in N . Those algorithms also have significantly worse Toffoli complexity, scaling roughly like $\tilde{O}(N^3)$.

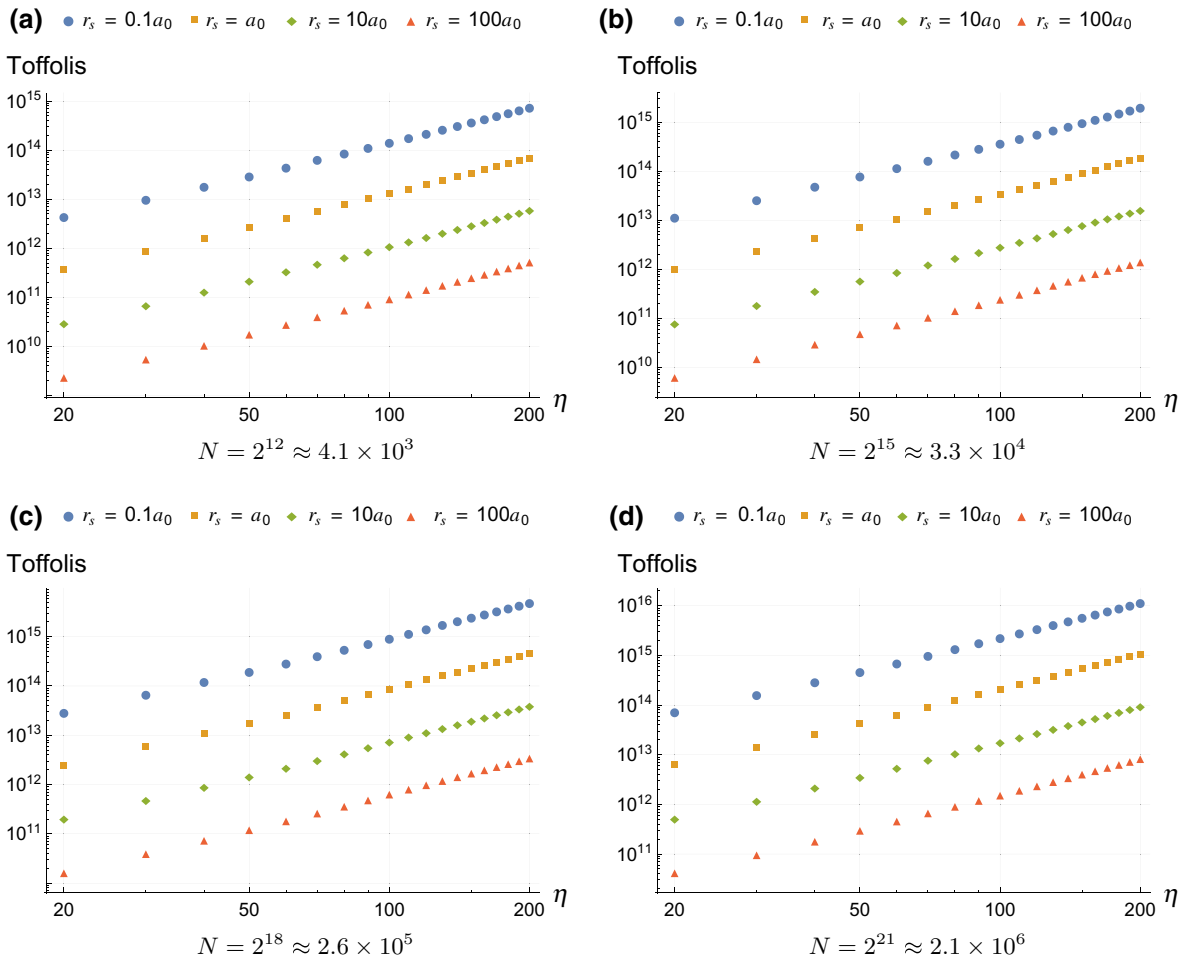


FIG. 4. Toffoli count of interaction-picture algorithm as a function of η using improved techniques. In contrast to Fig. 1, we plot these costs for fixed values of the Wigner-Seitz radius r_s to give a sense of the scaling when η increases proportionally to Ω (as would be the case when growing condensed phase simulations towards their thermodynamic limit). Thus, we expect these plots to be representative of costs when simulating periodic materials. We see that the scaling depends rather weakly on the number of plane waves, N .

The main motivation for those past papers using plane waves in second quantization is essentially a different regime than that of realistic simulations of materials. Those papers emphasize their resource estimates in terms of the first classically intractable applications of very small fault-tolerant processors (perhaps only a few hundred thousand physical qubits). In this sense, the goal is to articulate example problems on which “quantum supremacy” [82,83] might be first demonstrated for quantum chemistry in an error-corrected context, rather than to highlight applications corresponding to realistic simulations of nature. For example, at small N and high filling fraction η/N , there are a variety of model systems that are not converged to the continuum limit but provide examples of electronic structure that would be classically intractable to simulate within the basis. The value of such simulations for the jellium model (the homogeneous electron gas that occurs when $U = 0$) near $r_s = 10a_0$, is argued

in Sec. III of Ref. [29]. As a result, this model has been used as a benchmark in several subsequent papers. For example, the largest jellium simulation reported in Table III of Ref. [20] would require 2.1×10^{10} Toffoli gates [84] and 1136 logical qubits when $r_s = 10a_0$. This simulation would involve just 1024 plane-wave spin orbitals (equivalent here to $N = 512$). The largest $r_s = 10a_0$ jellium system analyzed in Table II of Kivlichan *et al.* [24] involves only $N = 125$ plane waves but would require 378 logical qubits and 2.8×10^9 Toffolis. Because those simulations occur at half-filling ($\eta = N$), they are likely more efficient than the methods of this paper at those system sizes. However, if we instead target a more physically accurate simulation (with less basis-set discretization error) where $N \geq 10^3$ and $\eta/N \leq 1/10$, we would expect the first quantized approach of this paper to require far fewer error-corrected resources than these past algorithms utilizing second quantization.

To compare the complexities reported here, the smallest value of N we consider here is for $N = 2^{12}$, which is 8 times larger than in Ref. [20]. The N^3 scaling of the second quantization approach would mean that the complexity is about 10^{13} Toffolis, nearly 100 times larger than the largest complexity we have for $r_s = 10a_0$ for the first quantized qubitization approach. For the largest value of N considered of 2^{21} , the N^3 scaling would indicate a complexity of over 10^{21} Toffolis, around a billion times larger than the complexity of the first quantized approach here.

Unlike second quantized plane-wave methods, second quantized algorithms using Gaussian basis functions seem promising options for practical quantum simulations of realistic molecular Hamiltonians. There are now a handful of papers presenting in-depth analysis of new fault-tolerant algorithms [22,26,27,38] in this representation. Dating back to the first of these papers by Reiher *et al.* [38], these papers have demonstrated algorithmic improvements by estimating the resources required to simulate a few benchmark molecules such as the FeMoCo molecule relevant to biological nitrogen fixation. The work by Lee *et al.* [27] holds the record for the Gaussian orbital-based algorithms with both the lowest constant factor and lowest asymptotic Toffoli scaling. We emphasize that all of these papers have focused on simulating molecular systems in active spaces, and none of them have developed methods for simulating periodic materials. As a result it is difficult to directly compare the algorithms developed here to those prior papers.

Active space Hamiltonians are coarse-grained versions of larger molecular Hamiltonians, designed to capture the physics of a subset of the total electronic degrees of freedom. Often the active space electrons are the valence electrons involved in the breaking or formation of chemical bonds. For example, the FeMoCo molecule ($\text{Fe}_7\text{MoS}_9\text{C}$) is a large system with 374 electrons; however, the active spaces studied by these papers require only 54 [38] or 113 [39] electrons. A variety of methods are available for computing active-space Hamiltonians [85–89] that are compatible with molecular orbital-based quantum algorithms. Most such approaches involve removing from the basis molecular orbitals that are expected to be completely occupied in the ground state and adding the potential arising from the charge density of those removed electrons to U . Unfortunately, such methods are not immediately compatible with the algorithms developed in this paper. This is because the molecular orbitals occupied by the core electrons are not orthogonal to the rest of the plane-wave basis and thus cannot be easily removed from the basis.

However, there is a way to make similar approximations for plane-wave methods. In particular, pseudopotentials [90,91] provide a well-studied methodology for removing core electrons from plane-wave calculations by modifying the external potential to account for the core. In addition

to reducing the effective system size η , pseudopotentials dramatically accelerate the convergence of the basis error. The reason for this is because the cusps in electronic wave function, which occur at the nuclei are smoothed out by pseudopotentials. Because the Fourier transform of smooth functions is exponentially convergent, the plane-wave basis can then converge dramatically faster. Thus, we suggest that the use of pseudopotentials [92,93] is likely to be essential for the practical deployment of these methods. However, as the present paper is already rather long, we leave the topic of removing core electrons via pseudopotentials to future work. Instead, we compare our algorithms to Gaussian orbital resource estimates that do not use active spaces.

A recent paper by Kim *et al.* [94] provides resource estimates based on the algorithm of von Burg *et al.* [26] (with modifications from Lee *et al.* [27]) that happens to not use active spaces, thus enabling a comparison to the present work, which we provide in Table VII. However, it is still difficult to know precisely how many plane waves would be required to match the accuracy of a given Gaussian orbital calculation. While both suppress basis error as $\mathcal{O}(1/N)$, even if a given Gaussian orbital calculation is more accurate than a given plane-wave calculation at fixed N , plane waves tend to converge more systematically than Gaussian orbitals to the continuum limit. As a result, one can more reliably predict the complete basis-set limit energy from extrapolations made from a series of calculations in progressively larger basis sets when using plane waves.

Appendix E of Ref. [29] attempts to give an idea of the relative prefactors between the number of plane waves and Gaussians required to obtain similar levels of precision. While the precise ratio of plane waves to Gaussians will depend on system-specific details, there it is suggested that periodic calculations with roughly $20N$ plane waves will often give higher precision than the same calculations with N Gaussian orbitals. It is more difficult to simulate nonperiodic systems like single molecules with plane waves and reducing this periodicity requires yet more plane waves by about a factor of 8 [29,96]. Achieving this convergence with plane waves may require softening nuclear cusps with pseudopotentials. Moreover, when targeting extremely high precision, the ratio of the required number of plane waves to Gaussians should decrease because the challenge becomes resolving the electron-electron cusp. Since the electron-electron cusp occurs at all points in space, Gaussians centered on nuclei have no special advantages for resolving it. The fact that Gaussians derive their advantage from being centered on nuclei also means that as the number of atoms in a fixed volume increases, the number of Gaussian orbitals to represent the system to fixed accuracy must also increase roughly proportionally. On the other hand, plane waves are not biased towards representing any particular nuclear locations; thus, as the

TABLE VII. Table comparing the cost of the qubitization algorithm of this paper to the cost of a state-of-the-art second quantized Gaussian basis algorithm recently analyzed by Kim *et al.* [94]. The work of Kim *et al.* estimates the resources required to simulate systems, which despite being small single molecules, are presented as relevant to the chemistry of lithium batteries. Their estimates are a natural point of comparison because they do not employ active spaces, unlike the focus of most past work using molecular orbitals [22,26,27,38,95]. We focus on two representative molecules from that work: ethylene carbonate and LiPF₆, discretized using four different Gaussian orbital basis sets. We note that the algorithm of Lee *et al.* [27] has been shown to scale better than the algorithm of von Burg *et al.* [26], on which the work by Kim *et al.* is based, but all resource estimates using that approach involve active spaces. We expect that in most contexts, if one uses a number of plane waves that is more than a thousand times the number of orbitals in a given Gaussian basis set, then one can expect similar or higher precision from the larger plane-wave calculation, even for small nonperiodic systems like the ones analyzed here. This would suggest that for $N = 2.6 \times 10^5$ plane waves, our algorithms may achieve higher precision than even the correlation consistent triple zeta (cc-pVTZ) basis set calculations studied by Kim *et al.* [94]. Thus, the results in this table may suggest that our algorithms requires significantly less surface-code space-time volume to reach similar levels of precision compared to some of the most performant Gaussian-based calculations.

Reference	System	Basis	Ω	η	N	Logical qubits	Toffoli count
Kim <i>et al.</i> [94]	Ethylene carbonate	STO-3G	N/A	46	34	2685	3.2×10^{10}
Kim <i>et al.</i> [94]	Ethylene carbonate	6-311G	N/A	46	90	14 492	8.5×10^{11}
Kim <i>et al.</i> [94]	Ethylene carbonate	cc-pVDZ	N/A	46	104	16 698	1.3×10^{12}
Kim <i>et al.</i> [94]	Ethylene carbonate	cc-pVTZ	N/A	46	236	81 958	3.1×10^{13}
This work	Ethylene carbonate	Plane waves	$10^3 a_0^3$	46	4.2×10^3	1395	2.5×10^{10}
This work	Ethylene carbonate	Plane waves	$10^5 a_0^3$	46	3.3×10^4	1701	6.6×10^{10}
This work	Ethylene carbonate	Plane waves	$10^5 a_0^3$	46	2.6×10^5	2021	1.7×10^{11}
This work	Ethylene carbonate	Plane waves	$10^5 a_0^3$	46	2.1×10^6	2355	4.2×10^{11}
Kim <i>et al.</i> [94]	LiPF ₆	STO-6G	N/A	72	44	3507	1.0×10^{11}
Kim <i>et al.</i> [94]	LiPF ₆	6-311G	N/A	72	112	18 423	1.9×10^{12}
Kim <i>et al.</i> [94]	LiPF ₆	cc-pVDZ	N/A	72	116	18 600	1.7×10^{12}
Kim <i>et al.</i> [94]	LiPF ₆	cc-pVTZ	N/A	72	244	84 721	3.3×10^{13}
This work	LiPF ₆	Plane waves	$10^5 a_0^3$	72	4.1×10^3	1758	8.0×10^{10}
This work	LiPF ₆	Plane waves	$10^5 a_0^3$	72	3.3×10^4	2150	2.1×10^{11}
This work	LiPF ₆	Plane waves	$10^5 a_0^3$	72	2.6×10^5	2556	5.1×10^{11}
This work	LiPF ₆	Plane waves	$10^5 a_0^3$	72	2.1×10^6	2976	1.3×10^{12}

N/A, not applicable.

number of atoms per unit volume is increased, the ratio of the number of plane waves to the number of Gaussians required for the same accuracy should decrease.

Considering all of these factors, Table VII suggests that our qubitized first quantized plane-wave algorithm might be more performant than the best Gaussian orbital basis algorithms applied to even small nonperiodic molecules. For example, when comparing our approach with $N = 2.6 \times 10^5$ plane waves to the approach of Kim *et al.* [94] using a triple zeta basis set for which N is between 236 and 244 for the two systems included in Table VII, see that we are using more than a thousand times the number of plane waves as Gaussian orbitals but that the first quantized plane-wave approach needs several dozen times fewer logical qubits and about a hundred times fewer Toffoli gates (amounting to approximately 3 orders of magnitude less surface-code space-time volume). Furthermore, it is clear that if the objective is significantly lower basis set error (e.g., if spectroscopic precision of 10^{-6} Hartree is desired) or if the goal is to simulate periodic systems such as crystals, metals, polymers, or surfaces, then our first quantized qubitization approach is dramatically more practical than other fault-tolerant algorithms described in the

literature. Thus, our analysis has shown that these methods are an important and state-of-the-art tool for the quantum simulation of chemistry.

B. Conclusion and outlook

We have completed the first constant factor analysis of the resources required to implement first quantized quantum simulations of quantum chemistry. Unsurprisingly, these methods appear to be the most promising approach for simulating realistic materials Hamiltonians within quantum error correction. More surprisingly, they are also highly competitive with the best second quantized approaches for simulating single molecules (often requiring fewer surface-code resources by multiple orders of magnitude). We focused on optimizing and compiling the first quantized algorithms using a plane-wave basis first described by Babbush *et al.* [42]. Despite the interaction picture of that work having better scaling, we found that the qubitization algorithm of that work tends to require fewer resources for the regimes of interest to us.

Our work goes significantly beyond just translating the prior asymptotic algorithm descriptions of Babbush *et al.* [42] to concrete circuits, even though just doing that would

have still required a similarly detailed paper. We also introduce many improvements to the implementation of both the qubitized form and the interaction-picture simulation that reduce the complexity. The innovations we introduce that lead to the most significant reduction in algorithmic complexity are summarized below.

- (i) In the case of qubitizing the Hamiltonian, we entirely eliminate the cost of the arithmetic for the kinetic energy term T (see Sec. II B). Instead of computing the sum of squares of momenta, control registers are used to select the momenta as well as the individual bits to multiply. The only arithmetic is the multiplication of bits, which is performed with a single Toffoli. This is a significant improvement, because multiplication would otherwise be the main cost.
- (ii) For the interaction-picture simulations we greatly reduce the cost of computing the total kinetic energy (see Sec. III B). Instead of computing it at each step, we compute the sum of squares of momenta at the beginning. Then, each time we block encode the potential operator, which involves updating just two momenta, we also update the register with the total momentum by adding the change in momentum.
- (iii) In both cases it is necessary to prepare a state with amplitudes $\propto 1/\|\nu\|$, where ν is a three-dimensional vector of integers. This preparation normally involves amplitude amplification, which triples the cost of state preparation [97]. For qubitizing the Hamiltonian, rather than using amplitude amplification, we can use the insight that this state preparation is only needed for the potential term, but not for the kinetic term. As a result, instead of using amplitude amplification, we can implement the kinetic term in branches of the quantum state that would otherwise correspond to failure (see Appendix D). That reduces the cost of this term by a factor of 3.
- (iv) In the case of the interaction picture one would normally require amplitude amplification on the Dyson series in order to implement a step of the time evolution deterministically, which again triples the cost. Instead, we can qubitize a single step of the Dyson series (see Sec. III A). This does not deterministically implement time evolution, but gives a step with eigenvalues related to those of the Hamiltonian. That eliminates one part where amplitude amplification was required, but in that case we still need to use amplitude amplification on the preparation of the state with amplitudes $\propto 1/\|\nu\|$.
- (v) Rather than performing arithmetic in place on the momentum registers, we swap the correct momentum registers into ancilla registers controlled on ancillae that select the registers (see Secs. III D

and III C). That means the arithmetic needs only to be performed once in those ancillae, rather than for every momentum register.

- (vi) For the discretization of the integrals, we use times in the centers of the intervals, which results in the error being higher order (see Appendix H). With a careful analysis of the error due to the time discretization, the number of bits needed to represent the time is about half of that using the methods from prior work.

As a result of these improvements, in cases where there are a large number of electrons the majority of the cost in each step is just selecting the momentum registers. There is an overhead in the interaction-picture approach corresponding to the order of the Dyson series. That means that to provide an improvement via the interaction picture, there should be a ratio between the kinetic and potential energy that is at least as large as the order of the Dyson series. That ratio is large in cases where the number of electrons is small, but in those cases the cost of preparing the state with amplitudes $1/\|\nu\|$ tends to be dominant. In those cases the direct qubitization approach has the additional advantage that one can avoid the amplitude amplification. That means the ratio needs to be even larger before there is an advantage due to using the interaction picture. We find that there is a minor improvement via the interaction picture with only $\eta = 20$ electrons and $N = 2^{21} \approx 2 \times 10^6$ plane waves, but as the number of electrons grows this advantage diminishes.

We have demonstrated that our first quantized qubitization algorithm compares favorably to all prior fault-tolerant algorithms for chemistry for both periodic, and nonperiodic systems. However, there are several obvious next steps that should be pursued in order to further refine and generalize our algorithms. While we have alluded to all of these ideas in the main text, we also summarize them below.

- (i) We should modify the algorithm for block encoding U to include various types of accurate pseudopotentials. Using pseudopotentials will allow us to remove core electrons from the simulation in order to only treat electronic degrees of freedom that are involved in chemical reactivity and bonding, thus effectively reducing η while increasing r_s . Furthermore, the screened potentials resulting from removing core electrons (even if only $1s$ electrons are removed) are smooth functions. With sharp nuclear cusps removed from the Coulomb potential (and thus, the wave function), the plane-wave basis will more rapidly converge to the continuum limit. It will be straightforward to make this modification for spherically symmetric pseudopotentials such as those that remove just the $1s$ electrons, but it

will be more involved to adapt the block encodings to more precise pseudopotentials for higher angular momenta electrons (e.g., using nonlocal pseudopotentials). Including pseudopotentials will also facilitate a more straightforward comparison to Gaussian basis approaches using active spaces.

- (ii) In Eq. (6) of this work we define the reciprocal lattice associated with our plane-wave mesh on a cubic lattice. However, for systems that lack a cubic conventional unit cell (e.g., those with hexagonal symmetry), such lattices cannot be used. Instead, we would need to define the reciprocal lattice using nonorthogonal Bravais vectors. This would change the definition of k_p from what is given in Eq. (6) to

$$k_p = \frac{2\pi}{\Omega^{1/3}} (p_1 g^{(1)} + p_2 g^{(2)} + p_3 g^{(3)})$$

$$p = (p_1, p_2, p_3) \in G, \quad (185)$$

where G is defined in Eq. (6) and $g^{(1)}$, $g^{(2)}$, and $g^{(3)}$ are the three-dimensional nonorthogonal unit vectors of the Bravais lattice defined by the *primitive* unit-cell geometry. Even for systems with cubic conventional unit cells, there are likely convergence advantages associated with being able to represent noncubic primitive unit cells. However, using this more general Bravais lattice will also come at a cost since it will make computing $\|k_p\|^2$ (a bottleneck of our approach) more costly. Further work should explore optimizations of these dot products in order to generalize our methods while retaining high efficiency.

- (iii) Related to the above point, there are different ways to converge the system physics towards its thermodynamic limit. The approach explored here involves growing the system in real space (i.e., by adding more unit cells to the computational cell). However, one can also grow the system towards the thermodynamic limit by evaluating an integral over the Brillouin zone in a fashion referred to as “k-point sampling” [60]. These schemes make use of generalized Bravais lattices such as the Monkhorst-Pack mesh, which allows one to represent systems of arbitrary symmetry while distributing the reciprocal lattice vector points homogeneously in the Brillouin zone [60]. Thus far we have only described how to perform so-called “ Γ -point” calculations. Future work should investigate the benefits of k-point sampling in the context of these algorithms and, if advantageous to do so, make the relatively simple modification of the algorithms required to support sampling outside of the Γ -point.
- (iv) We have discussed that the methods of this paper can be used to simulate nonperiodic systems despite

the underlying periodicity of plane waves. In particular, this can be accomplished by using a sufficiently large cell volume Ω so that periodic images do not interact. However, there are more efficient ways of preventing these periodic images from interacting. For example, one can use a range truncated Coulomb operators [98–100]. For an isolated molecule, the total electronic density decays exponentially quickly away from its center, and thus the molecule can be represented in a box of volume $\Omega = D^3$ with only exponentially small parts of its true wave-function density outside of the box. By using a Coulomb kernel truncated at distance D [96], such that

$$V(r, r') = \begin{cases} \frac{1}{|r-r'|} & |r-r'| \leq D \\ 0 & |r-r'| > D, \end{cases} \quad (186)$$

and by carrying out the simulation in a box of size $8\Omega = (2D)^3$, we ensure that there is no Coulomb interaction at all between the repeated images of the molecule, up to exponentially small terms in Ω arising from the density of the molecule outside of the box. While the Fourier amplitudes of the normal Coulomb kernel are $4\pi/k^2$, the Fourier amplitudes of the truncated Coulomb interaction shown above are $4\pi[1 - \cos(|k|D)]/k^2$. This would modify the potential operator coefficients of Eq. (7) to be

$$U_{pq} = \sum_{\ell=1}^L \frac{4\pi [1 - \cos(\|k_{p-q}\| D)] \zeta_{\ell} e^{ik_{q-p} \cdot R_{\ell}}}{\Omega \|k_{p-q}\|^2},$$

$$V_{pqrs}^{(\alpha, \beta)} = \delta_{p-q, r-s} \frac{4\pi [1 - \cos(\|k_{p-q}\| D)] \zeta_{\alpha} \zeta_{\beta}}{\Omega \|k_{p-q}\|^2}. \quad (187)$$

Thus, future work focused on simulating systems of reduced periodicity should adapt the algorithms presented here to block encode operators with the above modification to the coefficients.

- (v) This paper has provided first quantized algorithms for preparing molecular eigenstates given an initial state with sufficient overlap on an eigenstate of interest. However, methods for state preparation in first quantization are underdeveloped compared to their second quantized counterparts. The work of Ref. [21] provided an algorithm for symmetrizing or antisymmetrizing initial states with complexity $\mathcal{O}(\eta \log N \log \eta)$, but another common step in state preparation would be to initialize an arbitrary Slater determinant. Such states are described by $\mathcal{O}(\eta N)$ coefficients, and in second quantization can be prepared with gate complexity of $\tilde{\mathcal{O}}(\eta N)$ [23]. We would expect that this is also a lower bound on

the gate complexity required in first quantization. It would not be ideal for such state preparations to need to scale linearly in N . However, one can choose to involve fewer plane waves in the initial state preparation and so long as the number of plane waves taken is less than $M = \tilde{\mathcal{O}}(\eta^{1/3}N^{2/3}/\epsilon + \eta^{5/3}N^{1/3}/\epsilon)$ then state preparation will require a number of gates that is less than the $\mathcal{O}(\eta M)$ of our qubitization algorithm, and thus be a negligible additive cost (we also expect lower constant factors). But the most straightforward translation of this second quantized algorithm into first quantization gives an approach scaling as $\tilde{\mathcal{O}}(\eta^2 M)$, which would require taking $M = \tilde{\mathcal{O}}[N^{2/3}/(\eta^{2/3}\epsilon) + \eta^{2/3}N^{1/3}/\epsilon]$ in order to remain an additive negligible cost. This is likely fine for some applications but future work should investigate whether this scaling can be improved. Likewise, the topic of state preparation beyond single Slater determinants (even in second quantization) has often been ignored by papers providing error-corrected cost estimates for chemistry. However, when complex initial states are required (e.g., for challenging systems like FeMoCo), preparing these states may introduce a non-negligible cost that should be taken into account more carefully.

- (vi) The qubitization algorithm of this work can be used to perform time evolution, as opposed to used for preparing eigenstates for phase estimation, with the very minimal additional overhead [see Eq. (20)] of quantum signal processing [4]. We expect that time evolution of the non-Born-Oppenheimer Hamiltonian would be especially useful for computing chemical reaction rates [40,101]. There would also need to be minor modifications to the block-encoding algorithms of this paper to simulate the nuclear kinetic operators and explicit nuclei interaction terms appearing in Eq. (9). Future work should analyze the constant factors of such modifications in the context of a specific chemical dynamics problem and estimate the resources required to solve that problem within fault tolerance. This would constitute a fundamentally new type of application assessed for viability on an error-corrected quantum processor. Furthermore, we note that our interaction picture technique is trivially compatible with state-preparation methods based on adiabatic state preparation [102] as we can easily use that approach to perform time-dependent Hamiltonian evolutions. Such methods might prove important for preparing complex initial states when single Slater determinants fail to have sufficient overlap with the eigenstates of interest. However, more work is required to figure out how to best use the qubitization-based approach in the context of adiabatic state preparation.

- (vii) Finally, in Appendix K we asymptotically describe how one can instead simulate a real-space version of these algorithms using the gridlike representation employed by Kassal *et al.* [40], but employing more modern quantum-simulation methods. We introduce algorithms for block encoding U and V in this representation that lead to qubitization and interaction-picture algorithms with asymptotically the same complexity as the main algorithms explored in this work. However, the constant factors are different and the nonperiodic nature of that representation may have advantages when simulating nonperiodic systems. Thus, we propose that future work should more thoroughly explore this simulation strategy and compare the constant factors associated with it to those of the techniques in this paper.

In conclusion, there are still many aspects of these algorithms (and their applications) to explore. Nonetheless, this first study of fault-tolerant overheads of first quantized quantum simulations of chemistry suggests many potential advantages over their second quantized counterparts. While Gaussian orbital-based algorithms in second quantization will continue to compete with first quantized algorithms when simulating small nonperiodic molecules, we believe that the longer-term future of quantum computing for chemistry will focus much more on first quantized simulations. We expect this is the case due to the ease of removing the Born-Oppenheimer approximation, the fact that one can more efficiently refine the accuracy of simulations in first quantization, the fact that one does not require complicated and costly classical precomputation of active spaces and molecular orbitals, and the fact that, on a whole, these methods scale better than their second quantized counterparts, while still having reasonable constant factors (as our work has shown).

ACKNOWLEDGMENTS

The authors thank Garnet Kin-Lic Chan, Matthias Degroote, Craig Gidney, Cody Jones, Joonho Lee, Jarrod McClean, Yuval Sanders, Norm Tubman, and the quantum-computing team at BASF for helpful discussions. D.B. worked on this project under a sponsored research agreement with Google Quantum AI. D.B. is also supported by Australian Research Council Discovery Projects DP190102633 and DP210101367. N.W. is funded by a grant from Google Quantum AI as well as the US Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-Design Center for Quantum Advantage under contract number DESC0012704, which supported his work on the asymptotic analysis of the algorithms. Y.S. is partly supported by the National Science Foundation RAISE-TAQs 1839204. The

Institute for Quantum Information and Matter is an NSF Physics Frontiers Center PHY-1733907.

APPENDIX A: VARIABLE AND FUNCTION NAMES

- N : The number of reciprocal lattice vectors per electron.
- η : The number of electrons.
- m : The electron mass, and also used as an index of summation in other equations (particularly for discretized time). The difference is clear by the context.
- T : The kinetic energy operator.
- U : The potential energy between electrons and nuclei.
- V : Potential energy between electrons.
- L : The number of nuclei.
- ζ_ℓ : The nuclear charges.
- R_ℓ : Nuclear positions.
- ℓ : Usually indexes over the number of the nucleus.
- Ω : The computational cell volume.
- G : The set of N reciprocal lattice vectors defined in Eq. (6).
- G_0 : The set of reciprocal lattice vectors defined below Eq. (8).
- p, q : Plane-wave numbers, which are elements of G .
- ν : Differences of plane-wave numbers, which are elements of G_0 .
- k_p, k_ν : Reciprocal lattice vectors defined in Eq. (6).
- i, j : Used to index registers containing plane-wave numbers.
- i : The complex number $i = \sqrt{-1}$ is used in upright font, and the difference from the index i should be obvious from context.
- λ : The total sum of weights in the linear combination of unitaries.
- λ_T : The sum of the weights for T , see Eq. (25).
- λ_U : The sum of weights for U , see Eq. (25).
- λ_V : The sum of weights for V , see Eq. (25).
- λ_ζ : The sum of nuclear charges, $\sum_\ell \zeta_\ell$.
- λ_ν : The sum over $\nu \in G_0$ of $1/\|\nu\|$.
- n_p : The number of bits used to store each component of the plane-wave number.
- n_η : The number of bits needed to store the electron number, $\lceil \log \eta \rceil$.
- $n_{\mathcal{M}}$: The number of bits used in the inequality testing to prepare a state with amplitudes $1/\|\nu\|$.
- n_t : The number of bits used to store the time in each time register.
- $n_{\eta\zeta}$: The number of bits needed to store a number up to $\eta + 2\lambda_\zeta$.
- n_k : The number of qubits used for the register that we use to prepare the superposition over k , given by $\lceil \log \Sigma(0) \rceil$.
- n_T : The number of bits used in the rotation on the qubit selecting between T and $U + V$.
- n_R : The number of bits used to store each component of the nuclear position.
- K : The cutoff on the Dyson-series order used in the interaction picture.
- k : The particular order used in the Dyson series. (Some other uses of k are obvious from the context.)
- \mathcal{M} : The total number of alternative values used in the inequality testing for preparing amplitudes $1/\|\nu\|$, related to $n_{\mathcal{M}}$ as $\mathcal{M} = 2^{n_{\mathcal{M}}}$.
- M : The number of time steps in the discretization of the integrals for the time in the Dyson series, related to n_t as $M = 2^{n_t}$.
- ϵ : The allowable total rms error in the estimation of the energy.
- ϵ_{pha} : The allowable rms error in the energy due to phase estimation.
- $\epsilon_{\mathcal{M}}$: The allowable error in the estimated energy due to imprecise preparation of the state with amplitudes $1/\|\nu\|$.
- ϵ_R : The allowable error in energies due to the finite precision of the nuclear positions.
- ϵ_t : Allowable error in energies due to the finite precision in the times.
- ϵ_T : Allowable error in energies due to the finite precision in the rotation of the qubit selecting between T and $U + V$.
- w : Indexes over the x, y , and z components of the plane-wave numbers.
- r, s : Index over bits of a single component of the plane-wave number.
- a : Is given as 1 without amplitude amplification or 3 for amplitude amplification (to account for the increase in cost).
- $\text{Er}(n)$: The cost of erasing the QROM with n values, defined in Eq. (91).
- $\text{St}(n)$: The number of comparators to sort n items, listed in Table III.
- $\Sigma(n)$: A number used in the preparation of the superposition over k , defined in Eq. (159).
- $P_s(n, b_r)$: The probability of success for preparing an equal superposition state, given in Eq. (59).
- r_s : The Wigner-Seitz radius, which satisfies $4\pi r_s^3/3 = \Omega/\eta$.
- Δ : The spacing of the inverse reciprocal lattice given by $\Omega^{1/3}/N^{1/3}$.
- \mathcal{N} : The number of steps used in the phase estimation, proportional to $\lambda/\epsilon_{\text{pha}}$.
- τ : The length of time for a single step used in the phase estimation.
- A : The component of the Hamiltonian with larger norm considered in the interaction picture. We have $A = T$.
- B : The component of the Hamiltonian with smaller norm considered in the interaction picture. We take $B = U + V$.

- E_j : The energy eigenvalues.
- E_0 : An energy shift used to ensure that we can linearize the sine and arcsine in block encoding the Dyson series.
- m_1, \dots, m_k : Numbers used to index the discretized times.
- p_v : The probability of success of the state with the state with amplitudes $1/\|v\|$; see Eq. (106).
- p_v^{amp} : The probability of success when using amplitude amplification; see Eq. (117).
- μ : The numbering of the cube in the series of nested cubes used in preparing the state with amplitudes $1/\|v\|$.
- b_r : The number of bits used in rotating an ancilla used for preparing an equal superposition state.
- b_{grad} : The number of bits used in the kinetic energy phasing according to Eq. (150).
- b_T : The number of bits used for the implicit multiplication in the kinetic energy phasing.
- σ : A factor used in the phasing by T in the interaction picture.

APPENDIX B: BACKGROUND ON GALERKIN REPRESENTATIONS

1. Galerkin basis representations in first quantization

The Galerkin discretization is derived from the weak formulation of the Schrödinger equation in Hilbert space,

$$\begin{aligned}
 H_{\text{BO}} &= T + U + V + \frac{1}{2} \sum_{\ell \neq \kappa=1}^L \frac{\zeta_\ell \zeta_\kappa}{\|R_\ell - R_\kappa\|}, & H_{\text{nBO}} &= T + T_{\text{nuc}} + U_{\text{nBO}} + V + V_{\text{nuc}}, \\
 T &= \sum_{i=1}^{\eta} \sum_{p,q=1}^N T_{pq}^{(1)} |p\rangle \langle q|_i, & T_{\text{nuc}} &= \sum_{\ell=\eta+1}^{L+\eta} \sum_{p,q=1}^N T_{pq}^{(m_\ell)} |p\rangle \langle q|_\ell, \\
 U &= - \sum_{i=1}^{\eta} \sum_{p,q=1}^N U_{pq} |p\rangle \langle q|_i, & U_{\text{nBO}} &= - \sum_{i=1}^{\eta} \sum_{\ell=\eta+1}^{L+\eta} \sum_{p,q,r,s=1}^N V_{pqrs}^{(1,\zeta_\ell)} |p\rangle \langle s|_i |q\rangle \langle r|_\ell, \\
 V &= \frac{1}{2} \sum_{i \neq j=1}^{\eta} \sum_{p,q,r,s=1}^N V_{pqrs}^{(1,1)} |p\rangle \langle s|_i |q\rangle \langle r|_j, & V_{\text{nuc}} &= \frac{1}{2} \sum_{\ell \neq \kappa=\eta+1}^{L+\eta} \sum_{p,q,r,s=1}^N V_{pqrs}^{(\zeta_\ell, \zeta_\kappa)} |p\rangle \langle s|_\ell |q\rangle \langle r|_\kappa,
 \end{aligned} \tag{B1}$$

where the integrals are as defined in Eq. (2), R_ℓ are the positions of the L nuclei (under the Born-Oppenheimer approximation), m_ℓ are the masses of those nuclei and ζ_ℓ are their atomic numbers. As in the main text, the notation $|p\rangle \langle q|_i$ is a shorthand for $I_1 \otimes I_2 \cdots \otimes |p\rangle \langle q|_i \cdots \otimes I_{L+\eta}$. Finally, we note that it is not necessary to use the same basis functions for all distinguishable particles. For example, we could use a different basis (involving a fewer or greater number of functions) for the nuclei versus for the electrons. But here for simplicity, we

given by finding $|\phi\rangle$ (spanned by the basis vectors $\{|\phi_p\rangle\}$) such that $\langle \phi_p | H | \phi \rangle = E \langle \phi_p | \phi \rangle$ for all p . This is contrasted with the strong formulation of the differential equation that requires the discretization hold at all points in space, as opposed to assessing error on the restricted subspace spanned by $\{|\phi_p\rangle\}$. Energy errors are bounded from below within a Galerkin representation, meaning that for a finite basis size N , the energy of an eigenstate is always overestimated due to the discretization error. This property is beneficial when using variational methods to solve for eigenstates.

We now define both the Born-Oppenheimer and non-Born-Oppenheimer molecular Hamiltonians in a first quantized Galerkin discretization. In the Born-Oppenheimer case we store our wave function by having a computational basis that encodes configurations of the electrons in N basis functions such that a configuration is specified as $|\phi_1 \phi_2 \cdots \phi_\eta\rangle$ where each ϕ encodes the index of an occupied basis function. In the non-Born-Oppenheimer case we represent both nuclear and electronic degrees of freedom explicitly such that a configuration is specified by $|\phi_1 \phi_2 \cdots \phi_{\eta+L}\rangle$ where the first η registers encode the index of the basis function occupied by the η electrons and the last L registers encode the index of the basis function occupied by the L nuclei. Then, we can write the BO and nBO Hamiltonians as

define the Hamiltonians assuming the same basis for all particles.

2. Galerkin basis representations in second quantization

The methods and analysis of this paper focus on first quantization. However, the majority of past work in both classical quantum chemistry and quantum algorithms for chemistry focus on using second quantization. Thus, we

quickly review second quantization and its relation to first quantization for completeness and to place our work in context. In second quantization one works with computational basis states that represent symmetrized (or anti-symmetrized) configurations. For example, when encoding fermions in second quantization the computational basis state $|\phi_1\phi_2\cdots\phi_\eta\rangle$ would typically encode a real-space wave function corresponding to a Slater determinant of those basis functions; i.e.,

$$\begin{aligned} & \langle r_1 r_2 \cdots r_\eta | \phi_1 \phi_2 \cdots \phi_\eta \\ &= \sqrt{\frac{1}{\eta!}} \begin{vmatrix} \phi_{p_1}(r_1) & \phi_{p_2}(r_1) & \cdots & \phi_{p_\eta}(r_1) \\ \phi_{p_1}(r_2) & \phi_{p_2}(r_2) & \cdots & \phi_{p_\eta}(r_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{p_1}(r_\eta) & \phi_{p_2}(r_\eta) & \cdots & \phi_{p_\eta}(r_\eta) \end{vmatrix}. \end{aligned} \quad (\text{B2})$$

One can see that this function is antisymmetric by construction since swapping any two rows of a determinant incurs a negative sign. Of course, the actual computational basis states as defined on qubits do not have the same symmetries as the wave functions they encode. As a result, in second quantization one constructs operators so that they act on the wave function in a way that effectively enforces the proper symmetries.

The essential distinction between first and second quantization is that in first quantization the symmetries are encoded explicitly in the wave function whereas in second quantization the symmetries are encoded in the operators that act on the wave function. However, typically there is also a difference in how one indexes the occupied basis functions in first versus second quantization. Whereas in first quantization one tends to keep track of which basis function each particle occupies in a configuration, because the second quantized configurations encode a symmetrized representation, it is sufficient to merely keep track of which basis functions are occupied without associating a particular particle. For example, a typical way to store fermionic second quantized wave functions is to have a single qubit to represent each spin assigned basis function, which is either $|0\rangle$ if the spin orbital is unoccupied or $|1\rangle$ if the spin orbital is occupied [103]. However, because one should only symmetrize identical particles this needs to be done separately for each distinguishable type of particle (in this context, we should consider “spin-up” particles to be distinct from “spin-down” particles, etc.). As a result, in second quantization if we have J unique species of distinguishable nuclei then the number of qubits required will be proportional to J . With this notation, the Galerkin formulation of the second quantized molecular operators defining the Hamiltonians in Eq. (B1) take the explicit forms

$$\begin{aligned} T &= \sum_{\alpha \in \{\uparrow, \downarrow\}} \sum_{p,q=1}^N T_{pq}^{(1)} a_{p,\alpha}^\dagger a_{q,\alpha}, & T_{\text{nuc}} &= \sum_{\ell=1}^J \sum_{p,q=1}^N T_{pq}^{(m_\ell)} a_{p+\ell N}^\dagger a_{q+\ell N}, \\ U &= - \sum_{\alpha \in \{\uparrow, \downarrow\}} \sum_{p,q=1}^N U_{pq} a_{p,\alpha}^\dagger a_{q,\alpha}, & U_{\text{nBO}} &= \sum_{\alpha \in \{\uparrow, \downarrow\}} \sum_{\ell=1}^J \sum_{p,q,r,s=1}^N V_{pqrs}^{(1,\xi_\ell)} a_{p,\alpha}^\dagger a_{q+\ell N,\beta}^\dagger a_{r+\ell N,\beta} a_{s,\alpha}, \\ V &= \frac{1}{2} \sum_{\alpha,\beta \in \{\uparrow, \downarrow\}} \sum_{p,q,r,s=1}^N V_{pqrs}^{(1,1)} a_{p,\alpha}^\dagger a_{q,\beta}^\dagger a_{r,\beta} a_{s,\alpha}, & V_{\text{nuc}} &= \frac{1}{2} \sum_{\ell,\kappa=1}^J \sum_{p,q,r,s=1}^N V_{pqrs}^{(\xi_\ell, \xi_\kappa)} a_{p+\ell N}^\dagger a_{q+\kappa N}^\dagger a_{r+\kappa N} a_{s+\ell N}, \end{aligned} \quad (\text{B3})$$

where the indices ℓ and κ run over the J different species of distinguishable nuclei (rather than index the L different specific nuclei, as before). The nuclear orbitals do not have explicit spin assignments because we are assuming that the spin designation is summed over as one of the J nuclei types.

The operators a_p^\dagger and a_q^\dagger are either fermionic or bosonic creation and annihilation operators, depending on whether the type of particle they correspond to are fermions or bosons. For fermions we require the anticommutation relation that $\{a_p, a_q^\dagger\} = \delta_{pq}$ whereas for bosons we require the commutation relation that $[a_p, a_q^\dagger] = \delta_{pq}$. For fermions

there are many well-known methods that one can use to map these operators to qubits [30,104–106]. Such techniques yield a qubit Hamiltonian with $\mathcal{O}(JN)$ qubits where J here is the number of distinguishable fermionic species. Mapping the bosonic algebra to qubits is considerably more onerous since each mode (basis function) could be occupied by a number of bosons that is between zero and the total number of bosons in the simulation. For example, if there are b particles of a particular bosonic species then the number of qubits required to represent just that species is $\mathcal{O}(N \log b)$ in second quantization. This cost is multiplied by the number of different bosonic species.

APPENDIX C: QUBIT COSTINGS

1. Qubit costings for qubitization approach

The complete step needed for the qubitization, to give the step that one would perform phase estimation on, needs a reflection on the control ancilla. The cost of this reflection corresponds to the numbers of qubits used in the state preparation. To give this cost we first list the entire qubit cost, then describe the specific qubits that need to be reflected upon.

1. There are η registers storing the momenta, each of which has three components with n_p qubits, so the number of qubits needed is $3\eta n_p$.
2. The control register for the phase estimation needs $\lceil \log \mathcal{N} \rceil$ qubits, and there are $\lceil \log \mathcal{N} \rceil - 1$ qubits for the temporary registers, where

$$\mathcal{N} = \left\lceil \frac{\pi \lambda}{2\epsilon_{\text{pha}}} \right\rceil. \quad (\text{C1})$$

3. The phase gradient state that is used for the phase rotations. There are $n_R + 1$ bits used in the phasing, and n_T bits used in the rotation of the qubit selecting T , so the number of qubits needed for the phase gradient state is the maximum of these.
4. One qubit for the $|T\rangle$ state that is used catalytically for controlled Hadamards.
5. The qubit that is rotated to select between the T and $U + V$ components of the Hamiltonian.
6. The $n_{\eta\zeta} + 3$ qubits for the equal superposition for selecting between U and V , including the $n_{\eta\zeta}$ qubits for the superposition itself, the extra rotated qubit, the qubit flagging success of the preparation, and the qubit flagging the result of the inequality test to select between U and V .
7. Three qubits for selecting whether each of T , U , and V is performed.
8. The $2n_\eta + 5$ qubits from the preparation of the superpositions over i and j . There are η qubits for each of these registers in unary, then two qubits that are rotated for each of these preparations, two qubits that flag success of the two preparations, and one qubit that flags whether $i = j$.
9. The preparing the superposition state over ν has the following subcosts.
 - (a) Storing ν requires $3(n_p + 1)$ qubits.
 - (b) μ needs n_p qubits since it is given in unary. Note that μ takes values from 2 to $n_\mu = n_p + 1$, so there are n_p possible values to be encoded in unary.
 - (c) $n_{\mathcal{M}}$ qubits for the equal superposition state.
 - (d) $3n_p + 2$ qubits for testing if we have negative zero, including the flag qubit.

- (e) $2n_p + 1$ qubits used in signaling whether ν is outside the box, including the flag qubit. There are two qubits needed as ancillae for each of the n_p quadruple-controlled Toffolis, and the target qubit is the same each time.
 - (f) The qubits resulting from computing the sum of squares of components of ν and multiplying by \mathcal{M} are kept and used for uncomputing without Toffolis, and there are $3n_p^2 + n_p + 1 + 4n_{\mathcal{M}}(n_p + 1)$ (the same as the number of Toffolis).
 - (g) The qubit resulting from the inequality test.
 - (h) Two qubits, one flagging success of the inequality test, no negative zero and ν not outside the box, and the other an ancilla qubit used to produce the triply controlled Toffoli.
10. For the preparation of the equal superposition state over three basis states, we need four qubits. There are two to store the state itself, one qubit that is rotated, and one flagging success of the preparation.
 11. The states r and s are prepared in unary, and need n_p qubits each, for a total of $2n_p$.
 12. There are a number of temporary ancillae used in the arithmetic adding and subtracting ν into momenta. The cost here is given by items (a) and (c), giving a total of $5n_p + 1$.
 - (a) In implementing the SEL operations, we need to control a swap of a momentum register into an ancilla, which takes $3n_p$ qubits for the output. The $n_\eta - 1$ temporary ancillae for the unary iteration on the i or j register can be ignored because they are fewer than the other temporary ancillae used later.
 - (b) We use $n_p + 3$ temporary qubits to implement the block encoding of T , where we copy component w of the momentum into an ancilla, copy out two bits of this component of the momentum, then perform a controlled phase with those two qubits as control as well as the qubit flagging that T is to be performed.
 - (c) For the controlled addition or subtraction by ν in the SEL operations for U and V , we use n_p bits to copy a component of ν into an ancilla, then there are another $n_p + 1$ temporary qubits used in the addition, for a total of $2n_p + 1$ temporary qubits in this part.
 - (d) There are also temporary qubits used in converting the momentum back and forth between signed and two's complement, but these are fewer than those used in the previous step.
 13. There are two overflow qubits obtained every time we add or subtract a component of ν into a

momentum. All these qubits must be kept, giving a total of 6.

14. For the preparation of the state with the amplitudes ζ_ℓ , we do not need to output ℓ directly, because we can just output R_ℓ , which is all that is needed. The size of this output is $3n_R$ for the three components of the nuclear position.
15. There are also temporary qubits used in the arithmetic to implement $e^{-ik_v R_\ell}$. The arithmetic requires a maximum of $2(n_R - 2)$ qubits. Note that the R_ℓ can be output by the QROM, the phase factor applied, and the R_ℓ erased, after (or before) the arithmetic in item 12 is performed, so we need only take the maximum of the $5n_R - 4$ qubits used in items 14 and 15 and the $5n_p + 1$ temporary qubits used in 12.
16. A single qubit is used to control between adding and subtracting v in order to make SEL self-inverse.

Out of these qubits, the ones that need to be reflected upon are as follows.

- (i) Item 5, the qubit that is rotated to select between T and $U + V$, needs to be reflected upon.
- (ii) From item 6, the $n_{\eta\zeta}$ qubits that the equal superposition is prepared on, as well as the rotated ancilla, all need to be reflected upon. That gives a total of $n_{\eta\zeta} + 1$.
- (iii) From item 8, there are $2n_\eta$ qubits for i and j . There are also two qubits that are rotated, for a total of $2n_\eta + 2$ qubits (the flag qubits are rezeroed so do not need to be reflected upon).
- (iv) From item 9, we need to reflect upon the qubits listed in (a), (b), and (c), for a total of $4n_p + n_{\mathcal{M}} + 3$.
- (v) From item 10, there are two qubits for w , as well as one rotated qubit, for a total of three.
- (vi) From item 11, there are n_p for each of r and s , for a total of $2n_p$.
- (vii) From item 13, there are six qubits as overflow from the arithmetic that need to be reflected upon.
- (viii) The state with amplitudes $\sqrt{\zeta_\ell}$ is only prepared in an implicit way, via a QROM from the state generated in part 1. This QROM may be reversed erasing the extra ancillae, so there are no additional qubits from the state preparation that need to be reflected upon.

This gives a total of

$$n_{\eta\zeta} + 2n_\eta + 6n_p + n_{\mathcal{M}} + 16 \quad (\text{C2})$$

qubits to reflect on, with a corresponding Toffoli cost.

2. Qubit costings for interaction-picture approach

Now we consider the complete qubit costing for the interaction-picture algorithm. In the interaction-picture

algorithm, we need to block encode the potential energy $U + V$ many times (K) in succession. In a naive implementation, one would keep all the ancilla qubits used in the block encoding of each $U + V$, and reflect upon them all at the end. However, after each block encoding of $U + V$, one can use Toffolis to check that all the ancilla qubits are zero, putting the result in a flag qubit. Then for the next block encoding of $U + V$, one can reuse these ancilla qubits, because controlled on the flag qubit being zero these ancilla qubits have been rezeroed. After block encoding the entire Dyson series with the product of K block encodings of $U + V$ one can reflect upon these $K - 1$ flag qubits as well. The number of flag qubits is $K - 1$ rather than K , because the checking is done in between each of the K block encodings of $U + V$. It should also be noted that when using ancilla for the purpose of arithmetic, these ancilla qubits can *not* be used. This is because in the arithmetic Toffoli gates are inverted using measurements, and that requires that the ancillae used are guaranteed to be rezeroed.

For the computation of the time difference, multiplication of the time by the kinetic energy and phasing, at each step when a Toffoli is performed an extra ancilla qubit is introduced, so the number of ancillae is the same as the number of Toffolis for these three parts [the first three items in the list following Eq. (150) plus $n_t - 1$ for computing the absolute values of the times],

$$\begin{aligned} & 2(n_t - 1) + 2n_t(n_\eta + 2n_p) - n_t + b_{\text{grad}} - 2 \\ & = 2n_t(n_\eta + 2n_p) + n_t + b_{\text{grad}} - 4. \end{aligned} \quad (\text{C3})$$

For updating the kinetic energy register, there are two main steps.

- (i) Computing one component of $q \cdot v$ or $p \cdot v$ has cost $2n_p^2 - 3n_p$, and takes that many ancillae. The addition into the kinetic energy register uses $n_\eta + 2n_p - 1$ qubits. That gives a total

$$n_p(2n_p - 1) + n_\eta - 1. \quad (\text{C4})$$

- (ii) For controlled addition of $\|v\|$ we use $2n_p$ ancilla qubits and $n_\eta + 2n_p$ qubits for the addition itself, which takes less ancilla qubits than the previous step provided $n_p \geq 3$.

The dominant ancilla cost in updating the kinetic energy is that in Eq. (C4).

To quantify the total number of ancillae, we consider the qubit costs specific to the interaction-picture algorithm, then consider other qubit costs that are analogous to those for the qubitization approach. The specific qubit costs are as follows.

11. In preparing the equal superposition state for the preparation for k , we need n_k qubits to store the

equal superposition, another qubit that is rotated, and a success flag qubit, for a total of $n_k + 2$.

12. In completing the preparation of the superposition over k , we are keeping temporary ancillae to reduce the cost of inverting the inequality tests. As discussed in Appendix E, the number of qubits including the unary representation of k output is $n_k + 1 + \sum_{k=2}^{K-1} \lceil \log[\Sigma(k)] \rceil$.
13. The number of ancillae needed to store the K times is Kn_t .
14. In the reversible sort, a qubit is generated for every comparator, so there are $\text{St}(K)$ qubits.
15. We use an ancilla to store the total kinetic energy. This total kinetic energy has a maximum value of $3\eta(2^{n_p-1} - 1)^2$, so the number of qubits needed is $n_\eta + 2n_p$.
16. In phasing by T , $2n_t(n_\eta + 2n_p) + n_t + b_{\text{grad}} - 4$ qubits are used according to Eq. (C3), where b_{grad} is as in Eq. (150).
17. For updating the total kinetic energy the maximum number of ancillae used is $n_p(2n_p - 1) + n_\eta - 1$ according to Eq. (C4).
18. An extra phase gradient state with b_{grad} bits is used.
19. There are $K - 1$ flag qubits used to indicate success of each of the first $K + 1$ block encodings of $U + V$.
110. One qubit for controlling between forward and reverse time evolution for obtaining the sine.

The following qubit costs are analogous to the qubitization, so we use the same numbering. An exception is that we include I7 above as 13 (e) below, because it is needed at that step in the procedure.

1. There are η registers storing the momenta, each of which has three components with n_p qubits, so the number of qubits needed is $3\eta n_p$.
2. The control register for the phase estimation needs $\lceil \log \mathcal{N} \rceil$ qubits, and there are $\lceil \log \mathcal{N} \rceil - 1$ qubits for the temporary registers.
3. The phase gradient state that is used for the phase rotations (except those for kinetic energy phasing). The maximum precision of the rotations is that in the phasing for U , and needs $n_R + 1$ bits, so this is the number of qubits needed for the phase gradient state.
4. One qubit for the $|T\rangle$ state.
5. No qubit needs to be rotated to select between the T and $U + V$ components of the Hamiltonian, since we are not block encoding T .
6. For each of the K block encodings of $U + V$, there are $n_{\eta\zeta} + 3$ qubits for the equal superposition, including the $n_{\eta\zeta}$ qubits for the superposition itself, the extra rotated qubit, the qubit flagging success of the preparation, and the qubit flagging the result

of the inequality test to select between U and V . A feature here is that $n_{\eta\zeta} + 1$ qubits need to be kept and checked to be equal to zero, whereas the two flag qubits are temporary and rezeroed. This means that the total qubit cost is $n_{\eta\zeta} + 1$ plus 2 temporary ancillae.

7. Because we are not block encoding T , we do not need an extra qubit for the result of the inequality here.
8. For the superpositions over i and j , there are $2n_\eta$ binary qubits. In addition, for each there are two qubits rotated in the state preparation, giving a total of $2(n_\eta + 1)$ ancillae to be checked. There are also three temporary ancillae flagging success of the two state preparations and $i \neq j$, for a total of three temporary ancillae.
9. For preparing the superposition state over ν , the first three subcosts we list below are qubits that are imperfectly erased, so need to be checked to be equal to zero. For the rest of the subcosts they are temporary.
 - (a) Storing ν requires $3(n_p + 1)$ qubits.
 - (b) μ needs n_p qubits.
 - (c) $n_{\mathcal{M}}$ qubits for the equal superposition state.
 - (d) $3n_p + 2$ qubits for testing if we have negative zero, including a flag qubit.
 - (e) $2n_p + 1$ qubits used in signaling whether ν is outside the box, including a flag qubit. There are two qubits needed as ancillae for each of the n_p quadruple-controlled Toffolis, and the target qubit is the same each time.
 - (f) The qubits resulting from computing the sum of squares of components of ν and multiplying by \mathcal{M} are kept and used for uncomputing without Toffolis, and there are $3n_p^2 + n_p + 1 + 4n_{\mathcal{M}}(n_p + 1)$ (the same as the number of Toffolis).
 - (g) The qubit resulting from the inequality test.
 - (h) Two qubits, one flagging success of the inequality test, no negative zero and ν not outside the box, and the other an ancilla qubit used to produce the triply controlled Toffoli to set the first qubit.
10. For the preparation of the state with the amplitudes ζ_ℓ , we do not need to output ℓ directly, because we can just output R_ℓ , which is all that is needed. The size of this output is $3n_R$.
11. We are no longer preparing the equal superposition state over three basis states for w .
12. We are no longer preparing superposition states for r and s .
13. There are a number of temporary ancillae used in the arithmetic adding and subtracting ν into momenta.

The cost here is given by items (a) and (e), giving a total of $2n_p^2 + 5n_p + n_\eta$.

- (a) In implementing the SEL operations, we need to control a swap of a momentum register into an ancilla, which takes $3n_p$ qubits. There are another $n_\eta - 1$ temporary ancillae used in the unary iteration, but these may be ignored because they are not used at the same time as other temporary ancillae below.
- (b) We are no longer block encoding T .
- (c) For the controlled addition or subtraction by ν in the SEL operations for U and V , we use n_p bits to copy a component of ν into an ancilla, then there are another $n_p + 1$ temporary qubits used in the addition, for a total of $2n_p + 1$ temporary qubits in this part.
- (d) There are also temporary qubits used in converting the momentum back and forth between signed and two's complement, but these are fewer than those used in the previous step.
- (e) It is at this point that we also need to update the kinetic energy register, which takes $2n_p(n_p + 1) + n_\eta$ qubits. This cost is needed at the same time as that in (a), but not those in (c) or (d).

14. There are two overflow qubits obtained every time we add or subtract a component of ν into a momentum, for a total of six.
15. There are $2(n_R - 2)$ temporary qubits used in the arithmetic to implement $e^{-ik\nu \cdot R_\ell}$.

Now we need to consider the number of qubits that need to be checked to be zero to implement the block encoding of $U + V$. We have costs from items 6, 8, 9(a) to (c), and 14. Adding these gives a total

$$\begin{aligned} n_{\eta\zeta} + 1 + 2n_\eta + 2 + 3n_p + 3 + n_p + n_{\mathcal{M}} + 6 \\ = n_{\eta\zeta} + 2n_\eta + 4n_p + n_{\mathcal{M}} + 12. \end{aligned} \quad (\text{C5})$$

The cost of checking whether these qubits are all zero is one less Toffoli, and it is done $K - 1$ times, for a total of

$$(K - 1)(n_{\eta\zeta} + 2n_\eta + 4n_p + n_{\mathcal{M}} + 11) \quad (\text{C6})$$

Toffolis.

Next, we consider how many qubits we need to reflect upon for the complete qubitization. We then have the qubits listed above, as well as $n_k + 2$ from I1, Kn_t from I3, the $K - 1$ flag qubits from I9, and one qubit from I10. This gives a total of

$$n_{\eta\zeta} + 2n_\eta + 4n_p + n_{\mathcal{M}} + 12 + n_k + 2 + Kn_t + K - 1 + 1 \quad (\text{C7})$$

qubits to reflect upon, with a corresponding cost in Toffolis for the reflection. Again, we are including 2 Toffolis for

the control on the qubits for phase estimation, as well as the iteration of that register. That gives a total number of Toffolis

$$\begin{aligned} (K - 1)(n_{\eta\zeta} + 2n_\eta + 4n_p + n_{\mathcal{M}} + 11) + n_{\eta\zeta} + 2n_\eta \\ + 4n_p + n_{\mathcal{M}} + 12 + n_k + 2 + Kn_t + K - 1 + 1 \\ = K(n_{\eta\zeta} + 2n_\eta + 4n_p + n_{\mathcal{M}} + n_t + 12) + n_k + 3. \end{aligned} \quad (\text{C8})$$

The analysis of the total number of ancillae needed is complicated by the fact that there are temporary ancillae used in various parts of the procedure. The temporary qubits used in parts 13 and 15 are not needed at the same time, so we can take the maximum of those ancilla counts. But, the temporary ancillae used in part 9 are needed at the same time (because it is for preparing ν), so we need to add that temporary ancilla cost to the maximum of those in 13 and 15. From part 8, 3 of the qubits are temporary, and also need to be added to the total temporary ancilla costs, as are two qubits from part 7. But, all those temporary ancilla costs are *not* needed during the phasing by T . We also need ancilla qubits in order to check whether the number of qubits given in Eq. (C5) are all zero, but these are not needed at the same time as the other temporary ancillae. Therefore we take the maximum of these three temporary ancilla counts.

APPENDIX D: SELECTION BETWEEN T , U , AND V

Here we give detail for how the selection between T , U , and V is performed, and show how to calculate the precision required for the qubit for selecting between T and $U + V$. We are using a qubit to select between T and $U + V$, but are also applying SEL for T in some cases where state preparations for $U + V$ fail. To show this explicitly, we have a rotation on an ancilla selecting between T and $U + V$, to give

$$\cos \theta |0\rangle + \sin \theta |1\rangle, \quad (\text{D1})$$

and a preparation on the register selecting between $V(|0\rangle)$ and $U(|1\rangle)$ of

$$\frac{1}{\sqrt{\eta + 2\lambda_\zeta}} \left(\sqrt{\eta} |0\rangle + \sqrt{2\lambda_\zeta} |1\rangle \right), \quad (\text{D2})$$

where we abstract the state preparation as being just on a single qubit, rather than a superposition over $\eta + 2\lambda_\zeta$ basis states as proposed above. The $|0\rangle$ and $|1\rangle$ here would be equivalent to a superposition of subsets of those basis states, but can be described in the way presented here for clarity. The inequality test for $i \neq j$ has the effect of

preparing the state

$$\frac{1}{\sqrt{\eta}} \left(\sqrt{\eta-1} |0\rangle + |1\rangle \right). \quad (\text{D3})$$

Again we are using notation with a single qubit as a way of describing sets of basis states. Here $|0\rangle$ would be equivalent to the subset of i and j where $i \neq j$. The preparation of $1/\|v\|$ with some probability of failure can be written as

$$|0\rangle \sqrt{p_v/\lambda_v} \sum_{v \in G_0} \frac{1}{\|v\|} |v\rangle + |1\rangle \sqrt{1-p_v} |v^\perp\rangle, \quad (\text{D4})$$

where λ_v is needed for normalization for the state, p_v is the probability of success of the state preparation, and $|v^\perp\rangle$ is a state that is obtained in the case of failure. In the case of amplitude amplification, we can simply replace p_v with p_v^{amp} . We are adopting the convention that 0 corresponds to success and 1 flags failure.

Now there are two alternatives we can choose.

1. We can apply T if the ancilla for selecting T is 0 OR we have failure of the state preparation or the inequality test for $i \neq j$ with V .
2. We can apply T if the ancilla for selecting T is 0 AND we have failure of the state preparation or the inequality test for $i \neq j$ with V .

There are four bits that need to be checked for either of these alternatives. We order them in the order they were listed above; the qubits for selecting between T and $U+V$, the qubit for selecting between U and V , the qubit flagging $i \neq j$, and the qubit flagging success of preparation of $1/\|v\|$. So, for the first case (OR), the SEL for T is applied in the following cases. Below we use “x” to indicate arbitrary values.

1. $|0xxx\rangle$: That is, we have $|0\rangle$ on the first qubit, and any values for the other qubits.
2. $|1xx1\rangle$: We have $|1\rangle$ on the first qubit, but the preparation of $1/\|v\|$ failed.
3. $|1010\rangle$: We have $|1\rangle$ on the first qubit, the preparation of $1/\|v\|$ succeeded, but the second qubit selected V and the third qubit flagged $i = j$.

Because we perform SEL_T , then invert these state preparations, the total weight for T is the sum of the squares of the amplitudes for these alternatives. The total weight for T is then

$$\cos^2 \theta + \sin^2 \theta \left(1 - p_v + \frac{p_v}{\eta + 2\lambda_\zeta} \right). \quad (\text{D5})$$

The case where SEL_V is applied is just $|1000\rangle$, where the first qubit flags $U+V$, the second qubit flags V , the

third qubit flags $i \neq j$, and the fourth flags success of the preparation. As a result, the weight on V is

$$p_v \sin^2 \theta \frac{\eta - 1}{\eta + 2\lambda_\zeta}. \quad (\text{D6})$$

The cases where SEL_U are applied are $|11x0\rangle$, where the first qubit flags $U+V$, the second qubit flags U , and the fourth flags success of the preparation. As a result, the weight for U is

$$p_v \sin^2 \theta \frac{2\lambda_\zeta}{\eta + 2\lambda_\zeta}. \quad (\text{D7})$$

These relative weights for T , U , and V need to correspond to the relative weights between the λ values. First, note that the ratio of the weights between U and V is $2\lambda_\zeta/(\eta - 1)$, which is exactly what we need for λ_U/λ_V . This was the result of the inequality test $i \neq j$ adjusting the weights between U and V . We also need for the relative weight on T

$$\cos^2 \theta + \sin^2 \theta \left(1 - p_v + \frac{p_v}{\eta + 2\lambda_\zeta} \right) = \frac{\lambda_T}{\lambda_T + \lambda_U + \lambda_V}. \quad (\text{D8})$$

The value of θ needs to be chosen to make this expression hold. If λ_T is large, then it is always possible to choose θ , because the lhs can be arbitrarily close to 1. On the other hand, the minimum value of the lhs is the expression in brackets. If that is *greater* than the rhs, then we need to perform the state preparation with the second alternative, where we use the AND. Before we proceed to that alternative, note that because we always implement SEL for T , U , or V , the net λ value is $\lambda_T + \lambda_U + \lambda_V$.

In the case of AND, we implement SEL_T in the following cases.

1. $|0xx1\rangle$: We have $|0\rangle$ on the first qubit, and the preparation of $1/\|v\|$ failed.
2. $|0010\rangle$: We have $|0\rangle$ on the first qubit, and the preparation of $1/\|v\|$ succeeded, but the second qubit selected V and the third qubit flagged $i = j$.

As a result, the net weight on T is

$$\begin{aligned} & \cos^2 \theta \left(1 - p_v + \frac{p_v}{\eta + 2\lambda_\zeta} \right) \\ &= \cos^2 \theta - p_v \cos^2 \theta \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right). \end{aligned} \quad (\text{D9})$$

The cases where SEL_U and SEL_V are applied are changed in that we no longer check the first qubit. As a result, the

weights for V and U are changed to

$$p_v \frac{\eta - 1}{\eta + 2\lambda_\zeta}, \quad p_v \frac{2\lambda_\zeta}{\eta + 2\lambda_\zeta}, \quad (\text{D10})$$

respectively. We also need to apply the identity in the remaining cases. These are as follows.

1. $|1xx1\rangle$: We have $|1\rangle$ on the first qubit, but the preparation of $1/\|v\rangle$ failed.
2. $|1010\rangle$: We have $|1\rangle$ on the first qubit, the preparation of $1/\|v\rangle$ succeeded, but the second qubit selected V and the fourth qubit flagged $i = j$.

Adding these weights together gives

$$\sin^2 \theta \left(1 - p_v + \frac{p_v}{\eta + 2\lambda_\zeta} \right). \quad (\text{D11})$$

It is easily verified that the weights add to 1. However, the total λ is larger than $\lambda_T + \lambda_U + \lambda_V$ now, because there is also λ_I for performing the identity. Adding the weights for U and V gives

$$p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right), \quad (\text{D12})$$

so the total weight including T , U , and V is

$$\cos^2 \theta + p_v \sin^2 \theta \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right). \quad (\text{D13})$$

We need the relative weight for T

$$\frac{\cos^2 \theta \left(1 - p_v + \frac{p_v}{\eta + 2\lambda_\zeta} \right)}{\cos^2 \theta + p_v \sin^2 \theta \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right)} = \frac{\lambda_T}{\lambda_T + \lambda_U + \lambda_V}. \quad (\text{D14})$$

In this case the *largest* value on the lhs is given by the expression in brackets in the numerator, which is the same as the smallest value before. Therefore, one should choose one of the two alternatives for applying T depending on the relative values of $1 - p_v + p_v/(\eta + 2\lambda_\zeta)$ and $\lambda_T/(\lambda_T + \lambda_U + \lambda_V)$, and in either case it is possible to choose θ to get the appropriate weightings.

To determine the total value of λ in the second case (AND), we can use that the relative weights for U and V should satisfy

$$p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right) = \frac{\lambda_U + \lambda_V}{\lambda}, \quad (\text{D15})$$

so

$$\lambda = \frac{\lambda_U + \lambda_V}{p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right)}. \quad (\text{D16})$$

Now we have $\lambda_V/\lambda_U = (\eta - 1)/(2\lambda_\zeta)$, so

$$\begin{aligned} \frac{\lambda_U + \lambda_V}{\lambda_U + \lambda_V/(1 - 1/\eta)} &= \frac{1 + \lambda_V/\lambda_U}{1 + \eta\lambda_V/\lambda_U/(\eta - 1)}, \\ &= \frac{1 + (\eta - 1)/(2\lambda_\zeta)}{1 + \eta/(2\lambda_\zeta)}, \\ &= 1 - \frac{1}{\eta + 2\lambda_\zeta}. \end{aligned} \quad (\text{D17})$$

Hence, the λ value can be given as

$$\lambda = \frac{1}{p_v} \left(\lambda_U + \frac{\lambda_V}{1 - 1/\eta} \right). \quad (\text{D18})$$

This λ value necessarily must be at least $\lambda_T + \lambda_U + \lambda_V$, because λ_I is non-negative.

In the case where

$$1 - p_v + \frac{p_v}{\eta + 2\lambda_\zeta} < \frac{\lambda_T}{\lambda_T + \lambda_U + \lambda_V}, \quad (\text{D19})$$

then we need to apply the OR. In that case we find that

$$1 - \frac{\lambda_T}{\lambda_T + \lambda_U + \lambda_V} < p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right), \quad (\text{D20})$$

so

$$\begin{aligned} \frac{\lambda_U + \lambda_V}{\lambda_T + \lambda_U + \lambda_V} &< p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right) \\ \frac{\lambda_U + \lambda_V}{p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right)} &< \lambda_T + \lambda_U + \lambda_V. \end{aligned} \quad (\text{D21})$$

Therefore, in this case $\lambda_T + \lambda_U + \lambda_V$ is *larger* than the expression we use for λ in the AND case. To account for the two cases, we can therefore give the effective λ value as

$$\max \left[\lambda_T + \lambda_U + \lambda_V, \frac{1}{p_v} \left(\lambda_U + \frac{\lambda_V}{1 - 1/\eta} \right) \right]. \quad (\text{D22})$$

Next we consider the error due to the rotation θ being given to finite precision. Let us define

$$p_v' := p_v \left(1 - \frac{1}{\eta + 2\lambda_\zeta} \right). \quad (\text{D23})$$

In the case where OR is used in the implementation of SEL_T , the amplitude on $U + V$ relative to the entire Hamiltonian is

$$p_v' \sin^2 \theta, \quad (\text{D24})$$

and the weight on T relative to the entire Hamiltonian is

$$1 - p_v' \sin^2 \theta. \quad (\text{D25})$$

When there is error $\Delta\theta$ in the value of θ , the relative size in either can be changed by no more than

$$p_v' \Delta\theta < \Delta\theta. \quad (\text{D26})$$

To bound the error in the overall energy, it can be no more than $\lambda_T \Delta\theta$ for T , plus $(\lambda_U + \lambda_V) \Delta\theta$ for $U + V$, for a total error no larger than $\lambda \Delta\theta$.

Similarly, in the case where AND is used in the implementation of SEL_T , the relative weight for T can be given as

$$\cos^2 \theta (1 - p_v'). \quad (\text{D27})$$

The weight on $U + V$ is independent of θ . The error in the relative weight for T is then no larger than $\Delta\theta(1 - p_v') < \Delta\theta$, so the error in the energy is no larger than $\lambda_T \Delta\theta$. Taking into account both cases, the error in the energy is no larger than $\lambda \Delta\theta$. Now say we have n_T bits used in the rotation of the bit for selecting between T and $U + V$. The error in θ is then no greater than $\pi/2^{n_T}$. We therefore require

$$\pi \lambda / 2^{n_T} \leq \epsilon_T, \quad (\text{D28})$$

where ϵ_T is allowable error due to finite precision of this rotation. We can then take

$$n_T \geq \log(\pi \lambda / \epsilon_T). \quad (\text{D29})$$

For our block encodings, we replace λ_T with the slightly larger λ_T' , which is due to the preparation of the registers for selecting the digits, and modify λ_U and λ_V to $\lambda_U^\alpha, \lambda_V^\alpha$ to account for the finite precision of preparation of the superposition of ν . These modified values can be used in the above reasoning unchanged. Moreover, because $\lambda_U^\alpha, \lambda_V^\alpha$ are proportional to α , we can select α such that the rotation required for θ uses a finite number of bits, and we can just consider imprecision in the preparation of the superposition over ν with that value of α , rather than the finite number of bits for θ .

To be more specific,

$$\lambda_T' = \frac{6\eta\pi^2}{\Omega^{2/3}} 2^{2n_p-2}, \quad (\text{D30})$$

$$\lambda_U^\alpha + \lambda_V^\alpha = \frac{\eta}{2\pi\Omega^{1/3}} (\eta - 1 + 2\lambda_\zeta) \lambda_v^\alpha. \quad (\text{D31})$$

Therefore the ratio is

$$\begin{aligned} \frac{\lambda_U^\alpha + \lambda_V^\alpha}{\lambda_T'} &= \frac{\Omega^{1/3}}{12\pi^3} \frac{(\eta - 1 + 2\lambda_\zeta)}{2^{2n_p-2}} \lambda_v^\alpha, \\ &= \alpha \frac{4\Omega^{1/3}}{3\pi^3} \frac{(\eta - 1 + 2\lambda_\zeta)}{2^{2n_p-2}} 2^{n_p+2} p_v, \\ &= \alpha \frac{4\Omega^{1/3}}{3\pi^3} (\eta - 1 + 2\lambda_\zeta) 2^{-n_p} p_v. \end{aligned} \quad (\text{D32})$$

In practice, we find the best values of α are around $1 - 3/2\mathcal{M}$ to $1 - 1/\mathcal{M}$, with only a few percent variation in the error in that range. Therefore, if we have enough precision to give α in that range, then the error will be negligibly increased. The precision required for α approximately corresponds to the number of bits in \mathcal{M} , so we can typically take approximately $n_{\mathcal{M}}$ bits for the rotation of α .

APPENDIX E: DETAILED COSTING OF PREPARATION OF k STATE

To explain the sequence of operations to prepare the superposition over k in more detail, we first prepare an equal superposition state of the form

$$\frac{1}{\sqrt{\Sigma(0)}} \sum_{w=0}^{\Sigma(0)-1} |w\rangle. \quad (\text{E1})$$

Introducing a zeroed ancilla and performing the inequality test with $\Sigma(2)$ gives the state

$$\frac{1}{\sqrt{\Sigma(0)}} \left(\sum_{w=\Sigma(2)}^{\Sigma(0)-1} |w\rangle |0\rangle + \sum_{w=0}^{\Sigma(2)-1} |w\rangle |1\rangle \right). \quad (\text{E2})$$

Introducing a $|0\rangle$ state and performing a Hadamard gives $|+\rangle$. Then if we have zero on this qubit and the result of the inequality test, then we set zero in another ancilla; that requires just one Toffoli. This qubit is $|0\rangle$ for $k = 0$, and $|1\rangle$ for all other values of k . The state is now

$$\begin{aligned} \frac{1}{\sqrt{\Sigma(0)}} \left(\frac{1}{\sqrt{2}} \sum_{w=\Sigma(2)}^{\Sigma(0)-1} |w\rangle |0\rangle_2 (|0\rangle |0\rangle_1 + |1\rangle |1\rangle_1) \right. \\ \left. + \sum_{w=0}^{\Sigma(2)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 \right), \end{aligned} \quad (\text{E3})$$

where the subscripts 1 and 2 denote the qubits with the unary encoding of k , which are $|1\rangle$ for $k \geq 1$ and $k \geq 2$, respectively. Next, perform an inequality test with $\Sigma(3)$. This gives

$$\frac{1}{\sqrt{\Sigma(0)}} \left(\frac{1}{\sqrt{2}} \sum_{w=\Sigma(2)}^{\Sigma(0)-1} |w\rangle |0\rangle_2 (|0\rangle |0\rangle_1 + |1\rangle |1\rangle_1) |0\rangle_3 + \sum_{w=\Sigma(3)}^{\Sigma(2)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 |0\rangle_3 + \sum_{w=0}^{\Sigma(3)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 |1\rangle_3 \right). \tag{E4}$$

Now note that the inequality test here is performed on all the qubits storing w . We need only to perform an inequality test on the qubits that are sufficient to store values up to $\Sigma(2) - 1$, which gives

$$\frac{1}{\sqrt{\Sigma(0)}} \left(\frac{1}{\sqrt{2}} \sum_{w=\Sigma(2)}^{\Sigma(0)-1} |w\rangle |0\rangle_2 (|0\rangle |0\rangle_1 + |1\rangle |1\rangle_1) |?\rangle + \sum_{w=\Sigma(3)}^{\Sigma(2)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 |0\rangle + \sum_{w=0}^{\Sigma(3)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 |1\rangle \right). \tag{E5}$$

The question mark is used to indicate a value that may be 0 or 1 depending on w , but whose value we do not care about. Now we set a new qubit to $|1\rangle$ if the result of that inequality test *and* the subsystem with subscript 2 is in the state $|1\rangle$. That takes one Toffoli, and gives the state

$$\frac{1}{\sqrt{\Sigma(0)}} \left(\frac{1}{\sqrt{2}} \sum_{w=\Sigma(2)}^{\Sigma(0)-1} |w\rangle |0\rangle_2 (|0\rangle |0\rangle_1 + |1\rangle |1\rangle_1) |?\rangle |0\rangle_3 + \sum_{w=\Sigma(3)}^{\Sigma(2)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 |0\rangle |0\rangle_3 + \sum_{w=0}^{\Sigma(3)-1} |w\rangle |1\rangle_2 |+\rangle |1\rangle_1 |1\rangle |1\rangle_3 \right). \tag{E6}$$

That new qubit is now the third qubit in the unary representation of k . In this way we generate the unary representation of k directly, without needing to perform binary to unary conversion. The final inequality test we perform is that with $\Sigma(K)$ to generate the unary qubit flagging $k = K$.

To explain how to perform the inequality test more explicitly, see Fig. 18 in Ref. [74]. That shows how to implement addition with a (classically given) constant without needing to store that constant in qubits. Simply by reversing the circuit, we can subtract a constant, which can be used to perform an inequality test with that constant. That is shown in Fig. 5. In this figure, the inequality test would be on 4 bits, with $i_4 = t_4 = 0$ and the result of the inequality test being given in the bottom register.

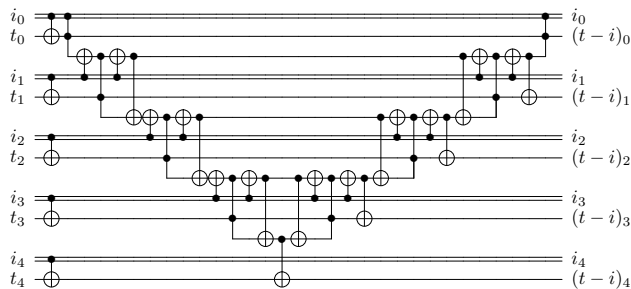


FIG. 5. A circuit to subtract a constant from a quantum register without storing that constant in qubits. The double lines indicate classical registers. The qubit at the bottom provides the output of an inequality test.

In that case, the result of the inequality test is given in the temporary ancilla where the second CNOT on t_4 is performed. That means that we can use the result of the inequality test immediately without needing to perform the rest of the quantum circuit. We also want to be able to use the input qubits for other purposes, without them being altered. The only operations that are performed with them as the target are the initial column of CNOTs. Therefore,

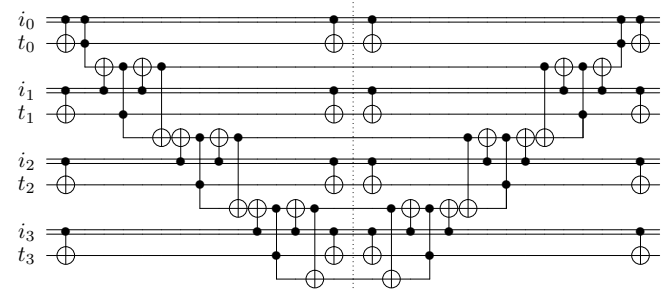


FIG. 6. A circuit to perform an inequality test with a constant giving the result in the bottom register at the dotted line. At the position of the dotted line the input registers are returned to their original values. For application of this circuit, operations that use the result of the inequality test would be put in the position of the dotted line. For the state preparation over k , the part of the circuit to the left of the dotted line would be used for preparation, then the SEL operations would be performed, then the part of the circuit on the right would be used to invert the state preparation.

those CNOTs can be performed on those qubits again to return them to their initial values, giving the circuit on the left side of the dotted line in Fig. 6.

At the point of the dotted line, the result of the inequality test has been output, and the other qubits have been returned to their initial values. Then one can simply reverse the circuit in order to erase the result of the inequality test. In this example, there is an inequality test with four qubits, and only 3 Toffolis are needed. Recall that the left and right elbows in these diagrams correspond to Toffolis, but the right elbows can be performed with measurements and Clifford gates. The initial Toffoli has one control that is a classical register, so is a Clifford gate. This circuit also indicates that four temporary ancillae are used, but the top one is produced with Clifford gates, so can be erased with Clifford gates and need not be kept through the rest of the calculation. As a result, there is a Toffoli cost of one less than the number of qubits, and an ancilla cost of one less than the number of qubits.

For the number of qubits, first there are n_k to store w . Then for the first inequality test we have $n_k - 1$ to perform the inequality test, including the output. That result of the inequality test is the first unary qubit for k . Then we have an additional $|+\rangle$ state, and generate the second unary qubit for k , with a total of $2n_k + 1$ qubits so far. Then for the other $K - 2$ inequality tests, we have $\lceil \log \Sigma(k) \rceil - 1$ qubits to make the inequality test reversible, then one qubit for the unary representation of k . That is distinct from the result of the inequality test because we are not using all the qubits in the inequality test. As a result, the total number of ancillae is $n_k + 1 + \sum_{k=2}^{K-1} \lceil \log \Sigma(k) \rceil$ including the unary representation of k , plus n_k to store w .

Next we describe how to properly prepare the times based on k . Let us first consider the case where the times are given in ascending order for implementing $e^{-iH\tau}$. For each unary register for k , the controlled Hadamard on n_t qubits gives an equal superposition. For each register beyond the range of k , we set the n_t qubits to 1. Then we have $m_{k+1} = 2^{n_t-1} - 1 = M - 1$, and subtracting m_k gives $M - 1 - m_k$. We apply an extra phasing corresponding to $1/2$ at the end, yielding $M - 1/2 - m_k$. Similarly, we apply a phasing of $1/2$ at the beginning, so we do not need to explicitly add $1/2$ to m_1 .

For register K , which is storing m_K if $k = K$ and $M - 1$ otherwise, we flip all the bits. This is equivalent to taking the number $M - 1$ (all ones) and subtracting our number, so we would get $M - 1 - m_k$ if $k = K$, or 0 otherwise. In the case $k = K$ we have the value $M - 1 - m_k$ required.

Now we consider the case that we wish to control between forward and time evolutions. This means the roles of the first and last time differences are reversed. The only difference in the procedure we have used is that for the last time difference we have flipped n_t bits. Therefore we can control this part by simply using CNOTs, and there is no Toffoli cost.

For the remaining $K - 1$ time differences, when we have the times in reverse order we obtain negative numbers in two's complement or zero. To find the absolute value, we can perform bit flips and add 1. To make this controlled, we can use CNOTs, and add the bit that flags that we are giving the times in reverse order. The cost of addition on n_t bits is $n_t - 1$ Toffolis. Combined with the $n_t - 1$ Toffolis needed for each subtraction, we have a cost of $2(K - 1)(n_t - 1)$ Toffolis.

A further subtlety in this procedure is that in the case of failure of the equal superposition state, we should perform no operations on the target system. However, we need only leave the qubit flagging failure of this state preparation flipped (not erasing it at the end). In the block encoding, this part is automatically eliminated, since postselection on this qubit being zero eliminates the failure.

APPENDIX F: THE MODIFIED PHASE GRADIENT STATE

To account for multiplying factors in phase shifts, the phase gradient state can be modified to something of the form

$$\frac{1}{\sqrt{2^{b_{\text{grad}}}}} \sum_{\ell=0}^{2^{b_{\text{grad}}}-1} e^{-2m\pi i \ell / 2^{b_{\text{grad}}}} |\ell\rangle, \quad (\text{F1})$$

for some integer m . As discussed in Sec. III B, the kinetic energy is given by $2\pi^2/\Omega^{2/3}$ times the sum of squares of the integers stored in the momentum registers, and this energy is multiplied by $\tau = 1/(\lambda_U + \lambda_V)$ to give the phase required. Now say that

$$\frac{2\pi^2}{(\lambda_U + \lambda_V)\Omega^{2/3}} = \sigma \frac{2\pi}{2^{b_{\text{grad}}}}, \quad (\text{F2})$$

so σ is the multiplying factor on $2\pi/2^{b_{\text{grad}}}$ that we need. We choose b_{grad} so that σ is between 2^{b_T-1} and $2^{b_T} - 1$ for some integer b_T , and we aim to round σ to an integer with b_T bits. The worst case is when σ is between 2^{b_T-1} and $2^{b_T-1} + 1$ (in the case of $b_T = 10$, this is 512 and 513). Then the reduction in $\tau/\exp[\tau(\lambda_U + \lambda_V)]$ in rounding to 2^{b_T-1} would be

$$\frac{e}{1/(\lambda_U + \lambda_V)} \frac{\sigma/2^{b_T-1}/(\lambda_U + \lambda_V)}{\exp(\sigma/2^{b_T-1})} = \frac{\sigma/2^{b_T-1}}{\exp(\sigma/2^{b_T-1} - 1)}. \quad (\text{F3})$$

The reduction in $\tau/\exp[\tau(\lambda_U + \lambda_V)]$ in rounding to $2^{b_T-1} + 1$ would be

$$\begin{aligned} & \frac{e}{1/(\lambda_U + \lambda_V)} \frac{\sigma/(2^{b_T-1} + 1)/(\lambda_U + \lambda_V)}{\exp[\sigma/(2^{b_T-1} + 1)]} \\ &= \frac{\sigma/(2^{b_T-1} + 1)}{\exp[\sigma/(2^{b_T-1} + 1) - 1]}. \end{aligned} \quad (\text{F4})$$

Because there is freedom in which way σ is rounded, we should consider the maximum of these two. The worst case (the smallest maximum) is when they are equal, which gives

$$\begin{aligned} \frac{1/2^{b_T-1}}{\exp(\sigma/2^{b_T-1} - 1)} &= \frac{1/(2^{b_T-1} + 1)}{\exp[\sigma/(2^{b_T-1} + 1) - 1]} \\ 1 + 1/2^{b_T-1} &= \frac{\exp(\sigma/2^{b_T-1} - 1)}{\exp[\sigma/(2^{b_T-1} + 1) - 1]} \\ 1 + 1/2^{b_T-1} &= \exp[\sigma/(2^{b_T-1} + 2^{2b_T-2})] \\ \sigma &= (2^{b_T-1} + 2^{2b_T-2}) \ln(1 + 1/2^{b_T-1}). \end{aligned} \quad (\text{F5})$$

Substituting that value of σ gives

$$\begin{aligned} \frac{\sigma/(2^{b_T-1} + 1)}{\exp[\sigma/(2^{b_T-1} + 1) - 1]} &= \frac{e2^{b_T-1} \ln(1 + 1/2^{b_T-1})}{\exp[2^{b_T-1} \ln(1 + 1/2^{b_T-1})]} \\ &\geq \frac{1}{1 + 1/2^{2b_T+1}}. \end{aligned} \quad (\text{F6})$$

Following this reasoning, in evaluating the complexity we can compute the values of λ_U and λ_V , choose a value of b_{grad} and find σ as

$$\sigma = \frac{\pi 2^{b_{\text{grad}}}}{(\lambda_U + \lambda_V) \Omega^{2/3}}, \quad (\text{F7})$$

then take

$$b_T = \lceil \log \sigma \rceil = b_{\text{grad}} + \left\lceil \log \left(\frac{\pi}{(\lambda_U + \lambda_V) \Omega^{2/3}} \right) \right\rceil, \quad (\text{F8})$$

where b_T is the number of bits being used for the approximation of σ . Then the increase in the effective λ from adjusting τ is no more than a factor of $1 + 1/2^{2b_T+1}$. The choice of b_{grad} can alternatively be made by choosing b_T first, then taking

$$b_{\text{grad}} = b_T - \left\lceil \log \left(\frac{\pi}{(\lambda_U + \lambda_V) \Omega^{2/3}} \right) \right\rceil. \quad (\text{F9})$$

The cost of the addition into the phase gradient register is $b_{\text{grad}} - 2$. Rather than using the upper bound $1 + 1/2^{2b_T+1}$ on the increase in the cost, it is also possible to just compute the effective time as

$$\frac{[\sigma]/\sigma}{(\lambda_U + \lambda_V) \exp([\sigma]/\sigma)}, \quad (\text{F10})$$

where $[\sigma]$ is being used to indicate the rounded value of σ .

APPENDIX G: COMPLEXITY OF BINARY SQUARING

In this Appendix we provide a sequence of results on the Toffoli complexity of computing squares of numbers. The first result uses a relatively simple technique based on repeated addition, but taking account of the fact that some bits are repeated.

Lemma 6 (Binary squaring). *An n -bit binary number can be squared using $n^2 - 2$ Toffoli gates.*

Proof. Consider a binary number $b_{n-1} \dots b_0$ stored in the input register. Our goal is to compute its square in a result register of length $2n$. In the initial step, we copy $b_0 = b_0 \times b_0$ to the result register using one CNOT gate. We also create $b_0 \times b_{n-1}, \dots, b_0 \times b_1$ using $n - 1$ Toffolis at positions $3, \dots, n + 1$ from the right. We then uncompute all the intermediate results.

In the second step, we copy $b_1 = b_1 \times b_1$ to an ancilla register of length n using one CNOT gate. We also create $b_1 \times b_{n-1}, \dots, b_1 \times b_2$ in the ancilla register using $n - 2$ Toffolis at positions $3, \dots, n$. We now add this number to the corresponding positions of the result register and we need an $(n + 1)$ -bit adder (including the carry), which by Ref. [107] can be implemented using n Toffolis. We uncompute all the intermediate results.

In the next step, we copy $b_2 = b_2 \times b_2$ to an ancilla register of length $n - 1$ using one CNOT gate. We also create $b_2 \times b_{n-1}, \dots, b_2 \times b_3$ using $n - 3$ Toffolis at positions $3, \dots, n - 1$. We now add this number to the result register and we need an n -bit adder (including the carry), which can be implemented using $n - 1$ Toffolis. We uncompute all the intermediate results.

This procedure can be recursed. In the penultimate step, we copy $b_{n-2} = b_{n-2} \times b_{n-2}$ to an ancilla register of length 3 using one CNOT gate. We also create $b_{n-2} \times b_{n-1}$ using 1 Toffolis at the left-most position. We now add this number to the result register and we need a 4-bit adder (including the carry), which can be implemented using 3 Toffolis. We uncompute all the intermediate results.

In the last step, we copy $b_{n-1} = b_{n-1} \times b_{n-1}$ to a single-qubit ancilla register. We use a 2-bit adder to add this value back to the result register. This costs 1 Toffoli gate.

Overall, we have used a total number

$$\frac{[1 + (n - 1)](n - 1)}{2} + \frac{(3 + n)(n - 2)}{2} + 1 = n^2 - 2 \quad (\text{G1})$$

of Toffoli gates. ■

However, it is possible to improve the complexity if we apply the approach for computing Hamming sums from Ref. [24]. We also used this method for improved arithmetic in Ref. [74]. The general principle is that bits are

grouped into those at different levels, where there are those multiplied by 1, those multiplied by 2, those multiplied by 2^2 and so forth. Three (or two) bits can be summed with a single Toffoli, giving a carry bit at the next level.

Therefore, one should sum bits at the first level, giving carry bits at the next level, then sum bits at the next level giving carry bits at the level after that, and so forth. In computing Hamming weights as in Ref. [24] there were only carry bits, but when computing products or sums there are bits at each level as well as carry bits. The bits obtained at each level will correspond to products of bits from the original numbers that can be computed with Toffolis.

In the following we refer to the complexity of “summing” bits at a level to mean the complexity of simply summing bits and giving carry bits at the next level, and refer to “completely summing” bits to mean that we sum the carry bits, then the carry bits from that sum, and so forth, as in the scheme for computing Hamming sums from Ref. [24]. Then the complexity of summing m bits at a level is equal to $\lfloor m/2 \rfloor$, so for even numbers it is just $m/2$ whereas for odd numbers there is rounding down. The number of carry bits is *equal* to the number of Toffolis used. The complexity of *completely* summing m bits (the Hamming sum) is no more than $m - 1$, though it is often less in specific instances [24].

First we consider the simpler case of computing a square, then proceed to sums of squares.

Lemma 7 ((Binary squaring improved)). *An n -bit binary number can be squared using $n(n - 1)$ Toffoli gates.*

Proof. Writing an n -bit number p in terms of its bits

$$p = \sum_{j=0}^{n-1} p_j 2^j, \quad (\text{G2})$$

we have

$$\begin{aligned} p^2 &= \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} p_j p_k 2^{j+k}, \\ &= \sum_{j=0}^{n-1} p_j 2^{2j} + 2 \sum_{j=0}^{n-1} \sum_{k=0}^{j-1} p_j p_k 2^{j+k}, \\ &= \sum_{j=0}^{n-1} p_j 2^{2j} + \sum_{\ell=0}^{2n-3} 2^{\ell+1} \sum_{j=\max(0, \ell-n+1)}^{\lfloor (\ell-1)/2 \rfloor} p_{\ell-j} p_j, \end{aligned} \quad (\text{G3})$$

where $\lfloor (\ell - 1)/2 \rfloor$ is to ensure that $\ell - j > j$. Breaking this up into odd and even gives

$$p^2 = p_0 + \sum_{\ell=1}^{n-2} 2^{2\ell+1} \sum_{j=\max(0, 2\ell-n+1)}^{\ell-1} p_{2\ell-j} p_j + \sum_{\ell=1}^{n-1} 2^{2\ell} \left(p_{\ell} + \sum_{j=\max(0, 2\ell-n)}^{\ell-1} p_{2\ell-1-j} p_j \right). \quad (\text{G4})$$

Now say n is even, so that for $\ell < n/2$ we have $\max(0, 2\ell - n + 1) = 0$ and $\max(0, 2\ell - n) = 0$ and for $\ell \geq n/2$ we have $\max(0, 2\ell - n + 1) = 2\ell - n + 1$ and $\max(0, 2\ell - n) = 2\ell - n$. Then we get

$$\begin{aligned} p^2 &= p_0 + \sum_{\ell=1}^{n/2-1} 2^{2\ell} \left(p_{\ell} + \sum_{j=0}^{\ell-1} p_{2\ell-1-j} p_j \right) + \sum_{\ell=1}^{n/2-1} 2^{2\ell+1} \sum_{j=0}^{\ell-1} p_{2\ell-j} p_j \\ &\quad + \sum_{\ell=n/2}^{n-1} 2^{2\ell} \left(p_{\ell} + \sum_{j=2\ell-n}^{\ell-1} p_{2\ell-1-j} p_j \right) + \sum_{\ell=n/2}^{n-2} 2^{2\ell+1} \sum_{j=2\ell-n+1}^{\ell-1} p_{2\ell-j} p_j. \end{aligned} \quad (\text{G5})$$

Alternatively, if n is odd, for $\ell < (n + 1)/2$ we have $\max(0, 2\ell - n + 1) = 0$ and $\max(0, 2\ell - n) = 0$ and for $\ell \geq (n + 1)/2$ we have $\max(0, 2\ell - n + 1) = 2\ell - n + 1$ and $\max(0, 2\ell - n) = 2\ell - n$. Then we get

$$\begin{aligned} p^2 &= p_0 + \sum_{\ell=1}^{(n-1)/2} 2^{2\ell} \left(p_{\ell} + \sum_{j=0}^{\ell-1} p_{2\ell-1-j} p_j \right) + \sum_{\ell=1}^{(n-1)/2} 2^{2\ell+1} \sum_{j=0}^{\ell-1} p_{2\ell-j} p_j \\ &\quad + \sum_{\ell=(n+1)/2}^{n-1} 2^{2\ell} \left(p_{\ell} + \sum_{j=2\ell-n}^{\ell-1} p_{2\ell-1-j} p_j \right) + \sum_{\ell=(n+1)/2}^{n-2} 2^{2\ell+1} \sum_{j=2\ell-n+1}^{\ell-1} p_{2\ell-j} p_j. \end{aligned} \quad (\text{G6})$$

To write both odd and even cases together we can write

$$p^2 = p_0 + \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4, \quad (G7)$$

where

$$\Sigma_1 = \sum_{\ell=1}^{\lceil n/2 \rceil - 1} 2^{2\ell} \left(p_\ell + \sum_{j=0}^{\ell-1} p_{2\ell-1-j} p_j \right), \quad (G8)$$

$$\Sigma_2 = \sum_{\ell=1}^{\lceil n/2 \rceil - 1} 2^{2\ell+1} \sum_{j=0}^{\ell-1} p_{2\ell-j} p_j, \quad (G9)$$

$$\Sigma_3 = \sum_{\ell=\lceil n/2 \rceil}^{n-1} 2^{2\ell} \left(p_\ell + \sum_{j=2\ell-n}^{\ell-1} p_{2\ell-1-j} p_j \right), \quad (G10)$$

$$\Sigma_4 = \sum_{\ell=\lceil n/2 \rceil}^{n-2} 2^{2\ell+1} \sum_{j=2\ell-n+1}^{\ell-1} p_{2\ell-j} p_j. \quad (G11)$$

In the following we refer to Σ_j by these symbols, and use “sums” to refer to sums of bits from individual terms of Σ_j .

The terms in each Σ_j provide bits at various levels that need to be summed. There is initially a bit from p_0 , which is on its own and does not need to be summed. Then the next level with bits to be summed corresponds to $\ell = 1$ in Σ_1 , then $\ell = 1$ in Σ_2 , then $\ell = 2$ in Σ_1 , and so forth. Summing each set of bits leads to carry bits at the next level. This proceeds up to bits at a level corresponding to $\ell = \lfloor n/2 \rfloor - 1$ for Σ_2 , then we go to $\ell = \lfloor n/2 \rfloor$ for Σ_3 . We now alternate between Σ_3 and Σ_4 , until the last level is that with $\ell = n - 1$ in Σ_3 (because Σ_4 ends at $\ell = n - 2$).

Now, for each of the Σ_j we have the following numbers of bits to sum given ℓ .

1. Σ_1 : There are $\ell + 1$ bits to sum.
2. Σ_2 : There are ℓ bits to sum.
3. Σ_3 : There are $n - \ell + 1$ bits to sum.
4. Σ_4 : There are $n - \ell - 1$ bits to sum.

Therefore, for $\ell = 1$ with Σ_1 , we have 2 bits to sum, giving one carry bit to $\ell = 1$ with Σ_2 , so there are again 2 bits to sum, giving one carry bit to $\ell = 2$ with Σ_1 . There are then 4 bits to sum including the carry bit, with complexity 2 and 2 carry bits to $\ell = 2$ with Σ_2 .

In general, there are 2ℓ bits to sum. We have seen this at the starting values of ℓ . To prove by induction, we assume it is true for some ℓ for Σ_1 . Then there are ℓ carry bits from Σ_1 to Σ_2 at the same value of ℓ . That gives a total number of bits to sum 2ℓ as required. Then there are $(\ell + 1) + 1 + \ell$ bits to sum for Σ_1 with $\ell' = \ell + 1$. That is equal to $2\ell'$, so the formula is correct for the next value of ℓ , this proving the formula by induction.

Passing from Σ_2 to Σ_3 with $\ell = \lceil n/2 \rceil$, there are $\lceil n/2 \rceil - 1$ carry bits, giving a total number of bits to

sum $n - \lceil n/2 \rceil + 1 + \lceil n/2 \rceil - 1 = n$. Now we consider the cases of odd and even n separately. For even n , there are $n/2$ bits carried to Σ_4 , and then a total number of bits to sum $n - n/2 - 1 + n/2 = n - 1$. Then there are $n/2 - 1$ bits carried to Σ_3 at $\ell = n/2 + 1$, giving a total number of bits to sum $n - (n/2 + 1) + 1 + n/2 - 1 = n - 1$. Then there are $n/2 - 1$ bits carried to Σ_4 at $\ell = n/2 + 1$, giving a total number of bits to sum $n - (n/2 + 1) - 1 + n/2 - 1 = n - 3$. These bits can be summed with $n/2 - 2$ Toffolis, giving the same number of carry bits.

In general, the number of bits to sum is $2(n - \ell) + 1$ for Σ_3 and $2(n - \ell) - 1$ for Σ_4 , except for $2(n - \ell)$ with $\ell = n/2$ for Σ_3 . We have seen this for the starting values, and for the iteration there are $n - \ell$ carry bits from Σ_3 to Σ_4 , giving a total number of bits to sum $2(n - \ell) - 1$. Then there are $n - \ell - 1$ carry bits from Σ_4 to Σ_3 at $\ell' = \ell + 1$, so there is then $n - (\ell + 1) + 1 + n - \ell - 1 = 2(n - \ell') + 1$ bits to sum. This gives the iteration required.

Then at the end we have $2\lceil n - (n - 1) \rceil + 1 = 3$ bits to sum with $\ell = n - 1$ for Σ_3 , and there are no further bit sums needed. Therefore, for the even n case the total number of Toffolis is

$$\begin{aligned} & \sum_{\ell=1}^{n/2-1} \ell + \sum_{\ell=1}^{n/2-1} \ell + \sum_{\ell=n/2}^{n-1} (n - \ell) + \sum_{\ell=n/2}^{n-2} (n - \ell - 1) \\ & = n(n - 1)/2. \end{aligned} \quad (G12)$$

The four sums here correspond to Σ_1 to Σ_4 , and the special case $\ell = n/2$ for Σ_3 can be combined with the others because the cost of the bit sums is still $n/2$. There are also $n(n - 1)/2$ Toffolis needed to compute the products of bits that are to be summed, giving a total complexity $n(n - 1)$ as required.

Then for the odd n case, there are $(n - 1)/2$ carry bits to Σ_4 at $\ell = (n + 1)/2$, giving a total number of bits to sum $n - (n + 1)/2 - 1 + (n - 1)/2 = n - 2$. That is equal to $2(n - \ell) - 1$, and so we are finding that the same formula as before holds. Similarly we have $2(n - \ell) + 1$ bits to sum for Σ_3 , and in fact we find that this formula holds for the initial $\ell(n + 1)/2$ for Σ_3 as well, so it does not need to be treated as a special case. The iteration to prove the formula is exactly the same as for the case of even n , so will not be repeated. Then the total Toffoli complexity is

$$\begin{aligned} & \sum_{\ell=1}^{(n-1)/2} \ell + \sum_{\ell=1}^{(n-1)/2} \ell + \sum_{\ell=(n+1)/2}^{n-1} (n - \ell) \\ & + \sum_{\ell=(n+1)/2}^{n-2} (n - \ell - 1) = n(n - 1)/2. \end{aligned} \quad (G13)$$

Again there is the complexity $n(n - 1)/2$ to compute the products of bits, giving a total complexity $n(n - 1)$ as required. ■

Next we consider the case where we need to sum three squares. This is a task that we need repeatedly, in order to calculate the norms of vectors. Summing the bits all together is more efficient than simply separately computing the sums then adding them together. The result is as follows.

Lemma 8 ((Summing triple squares)). *The sum of three n -bit binary numbers can be computed using $3n^2 - n - 1$ Toffolis.*

Proof. We use the same expression for the squares as before from Eq. (G7), again using Σ_j . In this case there are 3 bits at the first level, corresponding to p_0 for the 3 value of p to square. These may be summed with complexity 1, giving 1 carry bit at the next level (bits multiplied by 2). However, the next level with bits to sum is that for bits multiplied by 4, corresponding to $\ell = 1$ in Σ_1 . Therefore the carry bits are not included at that level, and there are $3(\ell + 1) = 6$ bits to sum, with complexity 3. That gives three carry bits to the next level, which corresponds to Σ_2 with $\ell = 1$. There are again $3\ell + 3 = 6$ bits to sum, with complexity 3. Then these 3 carry bits are taken back to Σ_1 with $\ell = 2$, and then we have $3(\ell + 1) = 12$ bits to sum, with complexity 6. We have 6 carry bits for the next level, Σ_2 with $\ell = 2$, which means there are $3\ell + 6 = 12$ bits to sum, again with complexity 6.

In general, for Σ_1 and Σ_2 there are 6ℓ bits to sum, which has complexity 3ℓ . Note that we have already shown this for $\ell = 1$ and 2. For induction, if it is true for some ℓ with Σ_1 , then there are 3ℓ carry bits being taken to term ℓ in Σ_2 , so the number of bits to sum is 6ℓ again. Then 3ℓ bits are being taken to term $\ell + 1$ in Σ_1 , and there are then $3(\ell + 2) + 3\ell = 6(\ell + 1)$ bits to sum. This demonstrates that our formula is correct via induction.

Next we consider the special case of even n , where we have the following special starting cases for $\ell \geq n/2$.

1. For Σ_3 with $\ell = n/2$, we have $3n/2 - 3$ carry bits from the previous level (Σ_2 with $\ell = n/2 - 1$). The number of bits to sum is therefore $3(n - \ell + 1) + 3n/2 - 3 = 3n$, which can be summed with $3n/2$ Toffolis.
2. There are $3n/2$ carry bits for Σ_4 with $\ell = n/2$, so there are $3(n - \ell - 1) + 3n/2 = 3n - 3$ bits to sum. These are summed with complexity $3n/2 - 2$.
3. There are $3n/2 - 2$ carry bits for Σ_3 at $\ell = n/2 + 1$. That gives $3(n - \ell + 1) + 3n/2 - 2 = 3n - 2$ bits to sum, which is done with complexity $3n/2 - 1$.
4. There are $3n/2 - 1$ carry bits for Σ_4 with $\ell = n/2 + 1$, so there are $3(n - \ell - 1) + 3n/2 - 1 = 3n - 7$ bits to sum. These are summed with complexity $3n/2 - 4$.

5. There are $3n/2 - 4$ carry bits for Σ_3 at $\ell = n/2 + 2$. That gives $3(n - \ell + 1) + 3n/2 - 4 = 3n - 7$ bits to sum, which is done with complexity $3n/2 - 4$.

From this point on, we find that there are $6(n - \ell) - 1$ bits to sum for Σ_4 with $\ell \geq n/2 + 1$, and $6(n - \ell) + 5$ bits to sum for Σ_3 with $\ell \geq n/2 + 2$. The last two items above we use as the starting point for the induction, where we see that the formula is correct.

To prove the iteration for induction, say there are $6(n - \ell) + 5$ bits to sum for Σ_3 with some ℓ . Then there are $3(n - \ell) + 2$ carry bits to Σ_4 . That gives $3(n - \ell - 1) + 3(n - \ell) + 2 = 6(n - \ell) - 1$ bits for Σ_4 as required. Then those can be summed with $3(n - \ell) - 1$ Toffolis, giving this number of carry bits to Σ_3 at $\ell + 1$. Then the number of bits to sum is $3(n - \ell) + 3(n - \ell) - 1 = 6(n - \ell) - 1 = 6[n - (\ell + 1)] + 5$. This is the formula for Σ_3 for $\ell + 1$ as required.

Finally, for $\ell = n - 1$, there are 11 bits to sum for Σ_3 . These give 5 bits at the next level, but there is no matching term from Σ_4 , so these can be completely summed with just another 3 Toffolis (there are two Toffolis and carry bits to the next level, then the two bits at the next level can be summed with one Toffoli). As a result, the overall Toffoli complexity is

$$\begin{aligned}
 & 1 + \sum_{\ell=1}^{n/2-1} 3\ell + \sum_{\ell=1}^{n/2-1} 3\ell + 3n/2 + (3n/2 - 2) + (3n/2 - 1) \\
 & + \sum_{\ell=n/2+2}^{n-1} [3(n - \ell) + 2] + \sum_{\ell=n/2+1}^{n-2} [3(n - \ell) - 1] + 3 \\
 & = n(3n + 1)/2 - 1. \tag{G14}
 \end{aligned}$$

Now, we also note that to compute the $n(n - 1)/2$ bit products $p_j p_k$ for each of the three components of p requires $3n(n - 1)/2$ Toffolis. Adding that to the Toffoli count above gives $3n^2 - n - 1$, as required.

Next we consider the case of odd n with Σ_3 and Σ_4 . Then for $\ell \geq (n + 1)/2$, we have the following starting cases.

1. For Σ_3 with $\ell = (n + 1)/2$ we have $3(n - 1)/2$ carry bits from Σ_2 . There is then a total of $3(n - \ell + 1) + 3(n - 1)/2 = 3n$, which can be summed with $(3n - 1)/2$ Toffolis.
2. For Σ_4 with $\ell = (n + 1)/2$ we have $3(n - 1)/2$ carry bits, giving a total $3(n - \ell - 1) + (3n - 1)/2 = 3n - 5$. This can be summed with $(3n - 5)/2$ Toffolis.
3. For Σ_3 with $\ell = (n + 3)/2$ we have $(3n - 5)/2$ carry bits from Σ_4 . The total number of bits to sum is $3(n - \ell + 1) + (3n - 5)/2 = 3n - 4$, which can be summed with $(3n - 5)/2$ Toffolis.

4. For Σ_4 with $\ell = (n + 3)/2$ we have $3(n - 5)/2$ carry bits, giving a total $3(n - \ell - 1) + (3n - 5)/2 = 3n - 10$. This can be summed with $(3n - 11)/2$ Toffolis.
5. For Σ_3 with $\ell = (n + 5)/2$ we have $(3n - 11)/2$ carry bits from Σ_4 . The total number of bits to sum is $3(n - \ell + 1) + (3n - 5)/2 = 3n - 10$, which can be summed with $(3n - 11)/2$ Toffolis.

From these last two cases on, we have the same formulae for the numbers of bits to sum as before, and again they can be used as the starting point for the induction. The induction proceeds in exactly the same way as before, so will not be repeated.

Lastly, for $\ell = n - 1$, there are again 11 bits to sum for Σ_3 , and the 5 carry bits can be completely summed with 3 Toffolis. As a result, the overall Toffoli complexity is

$$\begin{aligned}
 & 1 + \sum_{\ell=1}^{(n-1)/2} 3\ell + \sum_{\ell=1}^{(n-1)/2} 3\ell + (3n - 1)/2 + (3n - 5) \\
 & + \sum_{\ell=(n+5)/2}^{n-1} [3(n - \ell) + 2] + \sum_{\ell=(n+3)/2}^{n-2} [3(n - \ell) - 1] + 3 \\
 & = n(3n + 1)/2 - 1. \tag{G15}
 \end{aligned}$$

Again, adding the $3n(n - 1)/2$ Toffolis needed for computing the bit products gives the total of $3n^2 - n - 1$, as required. ■

Next we consider a more general case where we are summing k squares. In this case the bound we prove an upper bound on the complexity that is more efficient than computing the squares then summing, but numerical testing indicates that the actual complexity will be slightly lower.

Lemma 9 ((Binary sums of squares)). *The sum of the squares of k binary numbers, each of n bits, can be computed using kn^2 Toffoli gates.*

Proof. Again we can use Eq. (G7) with Σ_j for a square of a single integer. Similar to the previous cases, we need to consider the sum of bits that are multiplied by 1, giving carry bits to those multiplied by 2, then sum the bits that are multiplied by 2, and so forth. In the above case where we summed the squares of three numbers, there were 3 bits to sum that were multiplied by 1 (the least significant bit for each of the three numbers). That gave a single carry bit that was multiplied by 2. That was the only bit that was multiplied by 2, then the next level was bits that were multiplied by 4, corresponding to those for $\ell = 1$ in Σ_1 . Therefore no sum was needed for the single bit, and one could simply consider the sum of the bits for $\ell = 1$ for Σ_1 .

In the case where we are summing the squares for k numbers where k may be larger than 3, there may be more than one carry bit multiplied by 2. Those then need to be summed, which will introduce carry bits that are multiplied by 4, and will need to be summed together with the other bits from $\ell = 1$ in Σ_1 . The complexity of the sum of m bits (giving carry bits at the next level) is generally $\lfloor m/2 \rfloor$ Toffolis, so if m is even it is just $m/2$, but if m is odd then the complexity is $m/2$ rounded down to the nearest integer. This feature can be used to bound the complexity due to summing bits in terms in Σ_j together with possible carry bits that were obtained from the initial sum of k bits (the least significant bits from each of the k numbers). If we have m bits to sum and another c bits to sum together with those, then the number of Toffolis needed is no larger than $(m + c)/2$, regardless of whether $m + c$ is odd or even.

What this means is that we can upper bound the complexity by considering the sums of the bits in Σ_j separately from the carry bits from the initial sum of k bits. If we simply upper bound the complexity by dividing by 2, then the complexity due to summing both together can be upper bounded by upper bounding the complexities independently then adding the two upper bounds.

With that in mind, let us consider the numbers of bits to sum for Σ_1 and Σ_2 disregarding any carry bits from the initial sum over k bits. Here, the number of bits to sum (including carry bits from Σ_1 and Σ_2) will be shown to be $2k\ell$. Recall that we sum bits with $\ell = 1$ in Σ_1 , then with $\ell = 1$ for Σ_2 , then for $\ell = 2$ for Σ_1 , and so forth. To show this expression by induction, we first consider Σ_1 and $\ell = 1$, where we have $k(\ell + 1) = 2k = 2k\ell$ bits to sum. Therefore the formula is true for $\ell = 1$. Then assuming the formula is true for some ℓ for Σ_1 , we have $k\ell$ carry bits to Σ_2 . The total number of bits to sum is then $k\ell + k\ell = 2k\ell$ as required. Then, given the formula for Σ_2 , there are $k\ell$ bits to carry over to Σ_1 at $\ell' = \ell + 1$, giving $k\ell + k(\ell + 2) = 2k\ell'$ bits to sum, which is the correct formula with $\ell' = \ell + 1$, thereby establishing the iteration for induction. That proves that the formula $2k\ell$ is correct.

Now, considering the case of even n , we need to consider Σ_3 with $\ell = n/2$. Then we have $k(n/2 - 1)$ carry bits and $k(n/2 + 1)$ bits from Σ_3 , giving a total number of bits to sum kn . We write this number of bits as $2k(n - \ell + 1 - 2^{n-2\ell})$ for reasons that will become evident shortly. Then, given this formula, the number of carry bits to Σ_4 is no more than $k(n - \ell + 1 - 2^{n-2\ell})$. Here we are now ignoring any rounding down for odd cases so we can later introduce the complexity of summing the initial k bits.

The total number of bits to sum at ℓ for Σ_4 is then $k(n - \ell - 1) + k(n - \ell + 1 - 2^{n-2\ell}) = 2k(n - \ell - 2^{n-2\ell-1})$. Then there are $k(n - \ell - 2^{n-2\ell-1})$ carry bits to Σ_3 with $\ell' = \ell + 1$. The total number of bits to sum is then $k(n - \ell) + k(n - \ell - 2^{n-2\ell-1}) = 2k(n - \ell - 2^{n-2\ell-2}) = 2k(n - \ell' + 1 - 2^{n-2\ell'})$. This iteration means

that for $\ell > n/2$ the number of bits to sum may be upper bounded as $2k(n - \ell + 1 - 2^{n-2\ell})$ and $2k(n - \ell - 2^{n-2\ell-1})$ for Σ_3 and Σ_4 , respectively.

The highest level in Σ_3 and Σ_4 is $\ell = n - 1$ for Σ_3 , which corresponds to bits multiplied by $2^{2^{n-2}}$. Using the formula we just proved, we have an upper bound on the number of bits to sum $2k(2 - 2^{2-n})$, and the number of carried bits is $k(2 - 2^{2-n})$. To completely sum those bits the complexity would usually be given as the number of bits minus 1. However, that is using rounding down, and in order to account for the initial k bits summed we are not rounding down. Then the complexity of summing this number of bits is just a number of Toffolis equal to the number of bits. Similarly, we give the complexity of summing the k bits from the first level as k . This gives the overall complexity as

$$\begin{aligned} k + \sum_{\ell=1}^{n/2-1} k\ell + \sum_{\ell=1}^{n/2-1} k\ell + \sum_{\ell=n/2}^{n-1} k(n - \ell + 1 - 2^{n-2\ell}) \\ + \sum_{\ell=n/2}^{n-2} k(n - \ell - 2^{n-2\ell-1}) + k(2 - 2^{2-n}) \\ = kn(n+1)/2. \end{aligned} \quad (\text{G16})$$

In this expression the initial term with k is the upper bound on the contribution to the complexity from summing the initial k bits. Terms 2 to 5 are then complexities of the sums for Σ_1 to Σ_4 , respectively. The final term is the upper bound on the complexity for summing the final $k(2 - 2^{2-n})$ bits. There are $kn(n-1)/2$ Toffolis needed to compute the bit products of the form $p_j p_k$, so the total Toffoli complexity is upper bounded by kn^2 .

Then for the case of odd n , for Σ_3 at $\ell = (n+1)/2$, we have $k(n-1)/2$ carry bits and $k(n-\ell+1) = k(n+1)/2$ bits from Σ_3 . The total number of bits to sum is kn again. We can again write this as $2k(n-\ell+1-2^{n-2\ell})$, and we obtain the formula $2k(n-\ell-2^{n-2\ell-1})$ for the number of bits to sum at each level for Σ_4 again, by exactly the same process of logic as for the even n case. In the sum for the upper bound on the number of Toffolis needed, only the bounds on the sums are changed, and we obtain

$$\begin{aligned} k + \sum_{\ell=1}^{(n-1)/2} k\ell + \sum_{\ell=1}^{(n-1)/2} k\ell + \sum_{\ell=(n+1)/2}^{n-1} k(n - \ell + 1 - 2^{n-2\ell}) \\ + \sum_{\ell=(n+1)/2}^{n-2} k(n - \ell - 2^{n-2\ell-1}) + k(2 - 2^{2-n}) \\ = kn(n+1)/2. \end{aligned} \quad (\text{G17})$$

Again adding the $kn(n-1)/2$ Toffolis for computing bit products gives a total of kn^2 . ■

Despite this result being quite neat, it is not as tight as what we find when computing the Toffoli costs for particular examples. Numerically we find that the number of Toffolis needed is $kn^2 - n$, which is n lower than was proven above. In the case where k is a multiple of 3, numerical testing indicates the even tighter bound $kn^2 - n - 1$.

Lastly we give the complexity for products. The result is as follows.

Lemma 10 ((Binary products)). *The cost of multiplying two numbers, one with n bits and the other with m bits, is $2nm - n$ Toffolis for $n \geq m$.*

Proof. Without loss of generality we take p to be of length n and q to be of length m bits. To consider the cost of products, we can write out the product as

$$\begin{aligned} \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} p_j q_k 2^{j+k} = \sum_{\ell=0}^{m-1} \sum_{k=0}^{\ell} p_{\ell-k} q_k 2^{\ell} + \sum_{\ell=m}^{n-1} \sum_{k=0}^{m-1} p_{\ell-k} q_k 2^{\ell} \\ + \sum_{\ell=n}^{n+m-2} \sum_{k=\ell-n+1}^{m-1} p_{\ell-k} q_k 2^{\ell}. \end{aligned} \quad (\text{G18})$$

Then for the first few levels we have the following.

1. At $\ell = 0$, there is just a single bit and no sum needed.
2. At $\ell = 1$ there are two bits with a cost of 1 and one carry bit.
3. At $\ell = 2$ there are three bits and one carry bit to sum giving cost 2 and 2 carry bits.

The general formula is that at ℓ the cost of summing is ℓ . To see that, if there are ℓ carry bits to $\ell' = \ell + 1$, the number of bits to sum is then $2\ell + 2$, which has cost $\ell + 1 = \ell'$, so the iteration needed for induction holds.

Then for $\ell = m$ to $n-1$ there are always m bits to sum and the carry bits. So, at $\ell = m$ there are $m-1$ carry bits and m bits giving cost $m-1$. After that there are always $m-1$ carry bits and m bits to sum, giving cost $m-1$.

After that, for $\ell \geq n$, there are $n+m-1-\ell$ bits to sum plus the carry bits. At $\ell = n$ there are $m-1$ carry bits and another $m-1$, giving cost $m-1$. Then at $\ell = n+1$ there are $m-1$ carry bits and another $m-2$, giving cost $m-2$. In general, the cost of summing at ℓ is $n+m-1-\ell$. The total cost is then

$$\sum_{\ell=0}^{m-1} \ell + \sum_{\ell=m}^{n-1} (m-1) + \sum_{\ell=n}^{n+m-2} (n+m-1-\ell) = nm - n. \quad (\text{G19})$$

The cost of computing the bit products is then nm Toffolis, giving a total cost of $2nm - n$ as required. ■

APPENDIX H: BOUND ON TIME-DISCRETIZATION ERROR

For the time evolution under the Dyson series, we have the time-evolution operator that evolves an initial state as

$$|\psi(t)\rangle = \mathcal{T} \exp\left(-i \int_0^t H(s) ds\right) |\psi(0)\rangle. \quad (\text{H1})$$

Here

$$\begin{aligned} \partial_t \mathcal{T} \exp\left(-i \int_0^t H(s) ds\right) \\ = -iH(t) \mathcal{T} \exp\left(-i \int_0^t H(s) ds\right). \end{aligned} \quad (\text{H2})$$

We now consider the error due to the discretization of this integral. In each short time interval, we have

$$\int_{\tau}^{\tau+\delta} H(s) ds \approx H(\tau)\delta. \quad (\text{H3})$$

We can approximate H within the interval by

$$H(s) \approx H(\tau) + (s - \tau)H'(\tau). \quad (\text{H4})$$

The error is upper bounded by

$$\|H(\tau) - H(s)\| \leq \delta \max_{v \in [\tau, \tau+\delta]} \|H'(v)\|. \quad (\text{H5})$$

Then we can use the triangle inequality on unitary operators in the form

$$\begin{aligned} \|ABC - A'B'C'\| &= \|ABC - A'BC + A'BC - A'B'C + A'B'C - A'B'C'\| \\ &\leq \|(A - A')BC\| + \|A'(B - B')C\| + \|A'B'(C - C')\| \\ &= \|A - A'\| + \|B - B'\| + \|C - C'\|. \end{aligned} \quad (\text{H6})$$

That means, for the time evolution, which is a product of evolutions for infinitesimal time periods, we can bound the difference by the integral of the differences between the Hamiltonians. So, if we have Hamiltonians H_1 and H_2 , then

$$\begin{aligned} &\left\| \mathcal{T} \exp\left(-i \int_0^t H_1(s) ds\right) - \mathcal{T} \exp\left(-i \int_0^t H_2(s) ds\right) \right\| \\ &= \left\| \int_0^t \partial_{\tau} \mathcal{T} \exp\left(-i \int_0^{\tau} H_1(s) ds\right) d\tau - \int_0^t \partial_{\tau} \mathcal{T} \exp\left(-i \int_0^{\tau} H_2(s) ds\right) d\tau \right\| \\ &= \left\| \int_0^t (-iH_1(\tau)) \mathcal{T} \exp\left(-i \int_0^{\tau} H_1(s) ds\right) d\tau - \int_0^t (-iH_2(\tau)) \mathcal{T} \exp\left(-i \int_0^{\tau} H_2(s) ds\right) d\tau \right\| \\ &\leq \int_0^t \|H_1(s) - H_2(s)\| ds. \end{aligned} \quad (\text{H7})$$

In the case of the interaction picture, we have

$$H(t) = e^{itA} B e^{-itA}, \quad (\text{H8})$$

and thus the product rule implies

$$H'(t) = i e^{itA} [A, B] e^{-itA}. \quad (\text{H9})$$

As a result, using the triangle inequality and submultiplicative property of the spectral norm we can bound

$\|H'(t)\|$ as

$$\|H'(t)\| \leq 2\|A\| \|B\|. \quad (\text{H10})$$

For a single time step of length δ we have

$$\begin{aligned} \int_{\tau}^{\tau+\delta} \|H(\tau) - H(s)\| ds &\leq \int_{\tau}^{\tau+\delta} |s - \tau| 2\|A\| \|B\| ds \\ &= \delta^2 \|A\| \|B\|. \end{aligned} \quad (\text{H11})$$

Therefore, if the integral is approximated by M time steps of size $\delta = \tau/M$, then the error in the integral is

bounded by

$$M(\tau/M)^2 \|A\| \|B\| = \frac{\tau^2}{M} \|A\| \|B\|. \quad (\text{H12})$$

As we are working in the interaction frame of the kinetic operator, $\|A\|$ can be upper bounded by λ_T , and $\|B\|$ upper bounded by $\lambda_U + \lambda_V$. We are block encoding the Dyson series for the time-evolution operator over a time interval $\tau = 1/(\lambda_U + \lambda_V)$, and the number of time steps is (by the definition of n_t) $M = 2^{n_t}$. As a result, the error in the block-encoded operator due to time discretization is no larger than

$$\frac{\tau^2}{M} \|A\| \|B\| \leq \frac{\lambda_T}{2^{n_t}(\lambda_U + \lambda_V)}. \quad (\text{H13})$$

This error translates to an error in the estimated eigenvalue of $H/(\lambda_U + \lambda_V)$. Therefore, if ϵ_t is the allowable error due to the time discretization, then we need

$$\lambda_T 2^{-n_t} \leq \epsilon_t. \quad (\text{H14})$$

Therefore, n_t should be chosen as

$$n_t = \lceil \log(\lambda_T/\epsilon_t) \rceil. \quad (\text{H15})$$

Note that, in the above, the approximation of the time-dependent Hamiltonian is made at one side of the interval. A more accurate approximation can be made by using the middle of the interval, though that reduces only the number of bits for the time by 1 in this analysis. A more careful analysis can be made where the second-order correction to the Hamiltonian is considered. That can be used to show that the error in the approximation of the integral over the Hamiltonian is higher order.

Consider $U_1(\delta)$ to be the time-evolution operator over the interval $[\tau - \delta/2, \tau + \delta/2)$ for the Hamiltonian H , and $U_0(\delta)$ to be the time evolution operator for the Hamiltonian in the center of the interval, $H(\tau)$. We can find an expression for the derivative of $U_1(\delta)$ by integrating the differential equation for the time-evolution operator. As appropriate, we also use the notation that the time-evolution operator from times a to b is $U_1(b, a)$. Then we obtain

$$\begin{aligned} \partial_\delta U_1(\delta) &= \partial_\delta [U(\tau + \delta/2, \tau)U(\tau, \tau - \delta/2)], \\ &= -\frac{i}{2}H(\tau + \delta/2)U_1(\delta) + U(\tau + \delta/2, \tau)[\partial_\delta U(\tau, \tau - \delta/2)], \\ &= -\frac{i}{2}H(\tau + \delta/2)U_1(\delta) + U(\tau + \delta/2, \tau)[\partial_\delta U^\dagger(\tau - \delta/2, \tau)], \\ &= -\frac{i}{2}H(\tau + \delta/2)U_1(\delta) + U(\tau + \delta/2, \tau)[\partial_\delta U(\tau - \delta/2, \tau)]^\dagger, \\ &= -\frac{i}{2}[H(\tau + \delta/2)U_1(\delta) + U_1(\delta)H(\tau - \delta/2)]. \end{aligned} \quad (\text{H16})$$

Next, it follows immediately that the derivative of U_0 is

$$\frac{\partial}{\partial \delta} U_0(\delta) = -\frac{i}{2} [H(\tau)U_0(\delta) + U_0(\delta)H(\tau)]. \quad (\text{H17})$$

We can then upper bound the rate of change of the norm of the difference of U_0 and U_1 by

$$\begin{aligned} &\lim_{h \rightarrow 0} \frac{1}{h} \left| \|U_1(\delta + h) - U_0(\delta + h)\| - \|U_1(\delta) - U_0(\delta)\| \right| \\ &\leq \left\| \lim_{h \rightarrow 0} \frac{1}{h} [U_1(\delta + h) - U_0(\delta + h) - U_1(\delta) - U_0(\delta)] \right\| \\ &= \left\| \frac{\partial}{\partial \delta} [U_1(\delta) - U_0(\delta)] \right\| \\ &= \frac{1}{2} \|U_1(\delta)H(\tau - \delta/2) + H(\tau + \delta/2)U_1(\delta) - U_0(\delta)H(\tau) - H(\tau)U_0(\delta)\| \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \left\| U_1(\delta)H(\tau - \delta/2) + U_1(\delta)H(\tau + \delta/2) - 2U_1(\delta)H(\tau) \right. \\
 &\quad - U_1(\delta)H(\tau + \delta/2) + U_1(\delta)H(\tau) + H(\tau + \delta/2)U_1(\delta) - H(\tau)U_1(\delta) \\
 &\quad \left. + U_1(\delta)H(\tau) - U_0(\delta)H(\tau) + H(\tau)U_1(\delta) - H(\tau)U_0(\delta) \right\| \\
 &\leq \frac{1}{2} \left(\|U_1(\delta)[H(\tau - \delta/2) + H(\tau + \delta/2) - 2H(\tau)]\| \right. \\
 &\quad + \|[H(\tau + \delta/2) - H(\tau), U_1(\delta)]\| \\
 &\quad \left. + \|[U_1(\delta) - U_0(\delta)]H(\tau)\| + \|H(\tau)[U_1(\delta) - U_0(\delta)]\| \right). \tag{H18}
 \end{aligned}$$

Using an upper bound Λ on $\|H\|$, we have that $\|U(\delta) - I\| \leq \delta\Lambda$. We can therefore say that

$$\begin{aligned}
 \|[H(\tau + \delta/2) - H(\tau), U_1(\delta)]\| &= \|[H(\tau + \delta/2) - H(\tau), U_1(\delta) - I]\| \\
 &\leq 2\|U_1(\delta) - I\| \|H(\tau + \delta/2) - H(\tau)\| \\
 &\leq 2\delta\Lambda \|H(\tau + \delta/2) - H(\tau)\|. \tag{H19}
 \end{aligned}$$

If there is an upper limit Γ_1 on the norm of the derivative of H , then

$$\|[H(\tau + \delta/2) - H(\tau), U_1(\delta)]\| \leq \delta^2\Lambda\Gamma_1. \tag{H20}$$

If there is a constant upper bound Γ_2 on the second derivative of H , we get

$$\|U_1(\delta)[H(\tau + \delta/2) + H(\tau - \delta/2) - 2H(\tau)]\| \leq \Gamma_2\delta^2/4. \tag{H21}$$

These bounds give us

$$\|U_1(\delta) - U_0(\delta)\| \leq \int_0^\delta \left\{ \frac{1}{2}(\Lambda\Gamma_1 + \Gamma_2/4)s^2 + \Lambda \|U_1(s) - U_0(s)\| \right\} ds. \tag{H22}$$

A solution to the above integral equation is

$$\|U_1(\delta) - U_0(\delta)\| \leq \frac{4\Lambda\Gamma_1 + \Gamma_2}{(2\Lambda)^3} [2(e^{\Lambda\delta} - 1) - \Lambda\delta(2 + \Lambda\delta)]. \tag{H23}$$

For the interaction picture, we have

$$H(t) = e^{itA} B e^{-itA}, \tag{H24}$$

$$H'(t) = ie^{itA} [A, B] e^{-itA}, \tag{H25}$$

$$H''(t) = -e^{itA} [A, [A, B]] e^{-itA}. \tag{H26}$$

Therefore we can take $\Lambda = \|B\|$, $\Gamma_1 = 2\|A\| \|B\|$ and $\Gamma_2 = 4\|A\|^2 \|B\|$, so we get

$$\|U_1(\delta) - U_0(\delta)\| \leq \left(\frac{\|A\|}{\|B\|} + \frac{\|A\|^2}{2\|B\|^2} \right) \times [2(e^{\|B\|\delta} - 1) - \|B\|\delta(2 + \|B\|\delta)]. \tag{H27}$$

For our application, we are approximating the integral by 2^{n_t} time steps of size $\delta = \tau/2^{n_t}$. Then the error in the time-ordered exponential over time step δ is bounded by

$$2^{-n_t} \left(\frac{2\|A\|}{\|B\|} + \frac{\|A\|^2}{\|B\|^2} \right) [2^{n_t} (e^{\|B\|\tau/2^{n_t}} - 1) - \|B\|\tau(1 + \|B\|\tau/2^{n_t+1})]. \tag{H28}$$

Multiplying by 2^{n_t} gives the bound on the error for the time-ordered exponential over the entire region

$$\left(\frac{2\|A\|}{\|B\|} + \frac{\|A\|^2}{\|B\|^2} \right) [2^{n_t} (e^{\|B\|\tau/2^{n_t}} - 1) - \|B\|\tau(1 + \|B\|\tau/2^{n_t+1})]. \tag{H29}$$

This gives us the error in approximating the time-evolution operator with the centered difference formula and this leads us to the following lemma.

Lemma 11. Let $H(t) = e^{iA} B e^{-iA}$ for bounded Hermitian operators A and B and $t \in \mathbb{R}$. We then have that if n_t is a positive integer and $\delta = \tau/2^{n_t}$ then

$$\begin{aligned} & \left\| \mathcal{T} e^{-i \int_{t_0-\delta/2}^{t_0+\delta/2} H(t) dt} - e^{-iH(t_0)\delta} \right\| \\ & \leq 2^{-n_t} \left(\frac{2\|A\|}{\|B\|} + \frac{\|A\|^2}{\|B\|^2} \right) \left[2^{n_t} \left(e^{\|B\|\tau/2^{n_t}} - 1 \right) \right. \\ & \quad \left. - \|B\|\tau(1 + \|B\|\tau/2^{n_t+1}) \right]. \end{aligned}$$

In practice, for the interaction picture one should have $\|A\| \ll \|B\|$, so that it is the second term given in the round brackets above that should be dominant. Moreover, expanding the exponential gives

$$2(e^{\Lambda\delta} - 1) - \Lambda\delta(2 + \Lambda\delta) = \frac{1}{3}(\Lambda\delta)^3 + \mathcal{O}[(\Lambda\delta)^4]. \quad (\text{H30})$$

Therefore, we expect that the error should be approximately

$$\|U_1(\delta) - U_0(\delta)\| \lesssim \frac{\delta^3 \|A\|^2 \|B\|}{6} \quad (\text{H31})$$

for a time step, or

$$\frac{\tau^3 \|A\|^2 \|B\|}{6 \times 2^{2n_t}}, \quad (\text{H32})$$

for the entire time interval.

A similar result can be obtained for the approximation of just the integral by the discretized integral.

$$\begin{aligned} & \|H(\tau) + (s - \tau)H'(\tau) - H(s)\| \\ & \leq \frac{(s - \tau)^2}{2} \max_{v \in [\tau - \delta/2, \tau + \delta/2]} \|H''(v)\|. \end{aligned} \quad (\text{H33})$$

The bound on the difference between the integral over time δ and the value of H is therefore

$$\begin{aligned} & \left\| H(\tau)\delta - \int_{\tau-\delta/2}^{\tau+\delta/2} H(s) ds \right\| \\ & \leq \int_{\tau-\delta/2}^{\tau+\delta/2} \frac{(s - \tau)^2}{2} 4\|A\|^2 \|B\| ds \\ & = \frac{\delta^3}{24} 4\|A\|^2 \|B\| = \frac{\delta^3}{6} \|A\|^2 \|B\|. \end{aligned} \quad (\text{H34})$$

In taking the time-ordered exponential of the integral, we cannot simply use this bound because the term $(s - \tau)H'(\tau)$ does not average out, and there are contributions from higher-order terms in the exponential. The above derivation is able to take account of that.

APPENDIX I: INTERACTION-PICTURE ALGORITHM FOR SIMULATING GENERIC HAMILTONIANS

We have given an implementation of the interaction-picture algorithm in Sec. III for simulating first quantized quantum chemistry. In this Appendix, we describe a related implementation for simulating a generic Hamiltonian.

As in Sec. III, we assume that the target Hamiltonian takes the form $H = A + B$, where $\|B\|$ is much smaller than $\|A\|$. Then, the evolution under H for a short time τ can be represented by the Dyson series as

$$\begin{aligned} e^{-i(A+B)\tau} &= e^{-i\tau A} \mathcal{T} \exp \left(-i \int_0^\tau ds e^{isA} B e^{-isA} \right) \\ &= \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} \int_0^\tau d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{k-1}} d\tau_k e^{-i(\tau-\tau'_k)A} B e^{-i(\tau'_k-\tau'_{k-1})A} B \cdots B e^{-i(\tau'_2-\tau'_1)A} B e^{-i\tau_1 A} \\ &= \lim_{\substack{K \rightarrow \infty \\ M \rightarrow \infty}} \sum_{k=0}^K \frac{(-i\tau)^k}{M^k k!} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \cdots \sum_{m_k=0}^{M-1} e^{-i\tau(M-1/2-m'_k)A/M} B e^{-i\tau(m'_k-m'_{k-1})A/M} B \cdots \\ & \quad \cdots B e^{-i\tau(m'_2-m'_1)A/M} B e^{-i\tau(m'_1+1/2)A/M} \end{aligned} \quad (\text{I1})$$

where τ'_1, \dots, τ'_k are sorted times from τ_1, \dots, τ_k , and m'_1, \dots, m'_k are sorted integers from m_1, \dots, m_k . Here, we assume that the operator B is given by the block encoding

$$\langle 0 | \text{PREP}_B^\dagger \cdot \text{SEL}_B \cdot \text{PREP}_B | 0 \rangle = B/\lambda_B, \quad (\text{I2})$$

and A can be directly exponentiated on a quantum computer. This gives the linear combination of unitaries

$$e^{-i(A+B)\tau} \approx \left(\langle 0 | \text{PREP}_B^\dagger \right)^{\otimes K} \sum_{k=0}^K \frac{(\lambda_B \tau)^k}{M^k k!} \sum_{m_1, \dots, m_k=0}^{M-1} \left(e^{-i\tau(M-1/2-m'_k)A/M} (-i\text{SEL}_B) e^{-i\tau(m'_k-m'_{k-1})A/M} \right. \\ \left. \cdot (-i\text{SEL}_B) \dots \dots (-i\text{SEL}_B) e^{-i\tau(m'_2-m'_1)A/M} (-i\text{SEL}_B) e^{-i\tau(m'_1+1/2)A/M} \right) (\text{PREP}_B |0\rangle)^{\otimes K}, \quad (13)$$

where PREP_B prepares initial states in K different ancilla registers and, for each fixed value of k , the k selection operators SEL_B act on k different registers and the target system register.

To implement this on a quantum computer, we first prepare the state that encodes the coefficients from the Dyson-series expansion

$$\frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(\lambda_B \tau)^k}{k!}} |k\rangle, \quad (14)$$

where β is the normalization factor and $|k\rangle$ is represented in unary as $|1^k 0^{K-k}\rangle$. We would like to prepare this state in a similar way as in Sec. III D. To this end, we assume that the time interval τ can be taken as a rational multiple of $1/\lambda_B$:

$$\tau = \frac{t}{\mathcal{N}} = \frac{c}{d\lambda_B}, \quad (15)$$

where \mathcal{N} is the number of simulation steps and c and d are positive integers that are relatively prime. We can see that the squared amplitudes will be integers if we multiply by

$\Sigma(0)$, where we redefine $\Sigma(k)$ as

$$\Sigma(k) = \sum_{\ell=k}^K \frac{K! c^\ell d^{K-\ell}}{\ell!}. \quad (16)$$

This means that we can perform state preparation using the same inequality testing procedure as in Sec. III D except replacing the definition of $\Sigma(k)$. One minor change we need to make is that the first two amplitudes ($k=0$ and 1) are no longer equal. This means we need to perform the inequality test with the number $\Sigma(k+1)$ for $0 \leq k < K$, with complexity $\lceil \log[\Sigma(k)] \rceil - 1$. For $0 < k < K$ we also need a Toffoli for the conditioning on the result of the previous inequality test. Combining this with the cost in Eq. (161), the total Toffoli complexity for this state preparation (including inverse preparation) is

$$2(3n_k + 2b_r - 9) - 1 + \sum_{k=0}^{K-1} \lceil \log[\Sigma(k)] \rceil. \quad (17)$$

where $n_k = \lceil \log[\Sigma(0)] \rceil$ is the number of qubits for the equal superposition state.

Next, we consider the preparation of the register:

$$\frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(\lambda_B \tau)^k}{M^k k!}} |k\rangle \sum_{m_1, \dots, m_k=0}^{M-1} |m_1, \dots, m_k, M-1, \dots, M-1\rangle \\ \mapsto \frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(\lambda_B \tau)^k}{M^k k!}} |k\rangle \sum_{m_1, \dots, m_k=0}^{M-1} |m'_1, \dots, m'_k, M-1, \dots, M-1\rangle \\ \mapsto \frac{1}{\sqrt{\beta}} \sum_{k=0}^K \sqrt{\frac{(\lambda_B \tau)^k}{M^k k!}} |k\rangle \sum_{m_1, \dots, m_k=0}^{M-1} |m'_1, m'_2 - m'_1, \dots, m'_k - m'_{k-1}, M-1 - m'_k, 0, \dots, 0\rangle. \quad (18)$$

Specifically, we prepare k equal superposition states for each order k of the Dyson series using Hadamard gates controlled by the $|k\rangle$ register. We initialize the remaining

$K-k$ ancilla registers to state $|M-1\rangle$ corresponding to the maximal discrete time. This state preparation has Toffoli complexity Kn_t and will be inverted with the same cost.

We then quantumly sort the K registers with cost $2n_t \text{St}(K)$ and compute the time differences with cost $(K-1)(n_t-1)$. The same complexity will be needed to invert the sorting while no Toffoli gate is required to uncompute the time differences.

At the beginning and the end we also need to apply evolution by time $\tau/2M$. There are $(K+1)n_t$ bits controlling evolution under A . For simplicity we combine these bits with the two uncontrolled evolutions, giving a total of $(K+1)n_t+2$. Therefore we implement exponentials of A controlled by the time registers and perform PREP_B and SEL_B to block encode B with cost

$$[(K+1)n_t+2] \text{ controlled-}e^{isA} + 2K \text{ PREP}_B + K \text{ SEL}_B. \quad (19)$$

Note that this represents only an upper bound on the cost and can often be improved by using specific properties of the target Hamiltonian; see Sec. III for improvements for performing the first quantized quantum chemistry simulation. This implements the interaction-picture algorithm for time τ .

The success amplitude of this procedure can be made larger than $1/2$. To make the amplitude equal to $1/2$, we introduce an ancilla register with state

$$\cos \theta |0\rangle + \sin \theta |1\rangle, \quad (110)$$

so that the success amplitude is now multiplied by $\cos \theta$. Letting n_θ be the number of bits used to represent θ , we can then prepare this state with $n_\theta-3$ Toffoli gates.

To boost the success amplitude close to unity, we need to perform a single step of oblivious amplitude amplification, which triples the implementation of the truncated Dyson series and introduces two additional reflections. Specifically, as follows.

1. We implement the preamplified truncated Dyson series and prepare the θ state with a complexity as described above.
2. We perform a reflection on $n_k + K(n_t + n_B) + 2$ qubits.
3. We invert the truncated Dyson series and unprepare the θ state.
4. We perform a reflection again on $n_k + K(n_t + n_B) + 2$ qubits.
5. We implement the truncated Dyson series without preparing the θ state.

The $+2$ accounts for the qubit rotated by θ here, as well as a qubit rotated for the preparation of the equal superposition state for preparing k . Each reflection can be implemented with Toffoli complexity

$$n_k + K(n_t + n_B), \quad (111)$$

where n_B is the number of ancilla qubits for block-encoding operator B .

We now analyze the normalization factor introduced by the above steps. Such a factor will come from three sources: (i) the preparation of the Dyson-expansion coefficients; (ii) the preparation of the equal superposition state on the $|k\rangle$ register; and (iii) the implementation of the block encoding of operator B . Same as in Sec. III D, we have that the success probability of item (ii) is $p = P_s[\Sigma(0), b_r]$, whereas items (i) and (iii) together contribute a normalization factor of $\beta = \sum_{k=0}^K c^k / (d^k k!) \leq \exp(c/d)$. To boost this close to unity with a single step of oblivious amplitude amplification, we choose coprime integers c and d so that

$$\frac{P_s[\Sigma(0), b_r]}{\exp(c/d)} \geq \frac{1}{2}, \quad (112)$$

which implies

$$\mathcal{N} = \frac{\lambda_B t d}{c}. \quad (113)$$

The error of our circuit implementation is due to the truncation of the Dyson-series expansion (ϵ_K) and the finite binary representation of the time steps (ϵ_t), as well as the preparation of θ state (ϵ_θ). With n_θ digits to represent the angle θ , the error in the amplitude $\cos \theta$ is upper bounded by

$$\epsilon_\theta = \frac{\pi}{2^{n_\theta}}, \quad (114)$$

which causes an overall error of at most $\epsilon_K + \epsilon_\theta$ in the preamplified truncated Dyson series. After a single step of oblivious amplitude amplification, the error is bounded as [108, Lemma G.2]

$$(\epsilon_K + \epsilon_\theta) \frac{(\epsilon_K + \epsilon_\theta)^2 + 3(\epsilon_K + \epsilon_\theta) + 4}{2}. \quad (115)$$

To ensure that the entire simulation has accuracy ϵ , we require that

$$\mathcal{N}(\epsilon_K + \epsilon_\theta) \frac{(\epsilon_K + \epsilon_\theta)^2 + 3(\epsilon_K + \epsilon_\theta) + 4}{2} + \mathcal{N}\epsilon_t \leq \epsilon. \quad (116)$$

We summarize the cost of the circuit implementation in the following theorem.

Theorem 12. *Using the interaction-picture algorithm, the time evolution under the Hamiltonian $H = A + B$ can be simulated for time t that is a rational multiple of $1/\lambda_B$ to within error ϵ by choosing positive integers K , n_t , n_θ , b_r , coprime c and d , and*

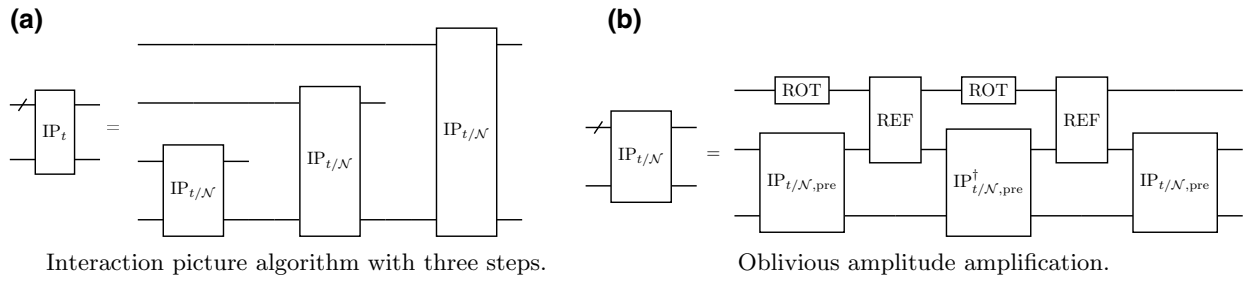


FIG. 7. Quantum circuits for the interaction-picture algorithm for simulating generic Hamiltonians.

using $3[\mathcal{N}(K+1)n_t + 2]$ controlled- e^{isA} , $6\mathcal{N}K$ PREP_B , $3\mathcal{N}K$ SEL_B , and a number of Toffoli gates

$$6\mathcal{N}(3n_k + 2b_r - 9) + 3\mathcal{N} \left(-1 + \sum_{k=0}^{K-1} \lceil \log[\Sigma(k)] \rceil \right) + 6\mathcal{N}Kn_t + 12\mathcal{N}(n_t - 1)St(K) + 6\mathcal{N}(K-1)(n_t - 1) + 2\mathcal{N}(n_\theta - 3) + 2\mathcal{N}[n_k + K(n_t + n_B)], \quad (I17)$$

where $St(K)$ is the cost of sorting K items, $\tau = c/(d\lambda_B)$, $\mathcal{N} = t/\tau$ is an integer, and

$$\frac{P_s[\Sigma(0), b_r]}{\exp(c/d)} \geq \frac{1}{2}, \quad (I18)$$

$$\mathcal{N}(\epsilon_K + \epsilon_\theta) \frac{(\epsilon_K + \epsilon_\theta)^2 + 3(\epsilon_K + \epsilon_\theta) + 4}{2} + \mathcal{N}\epsilon_t \leq \epsilon, \quad (I19)$$

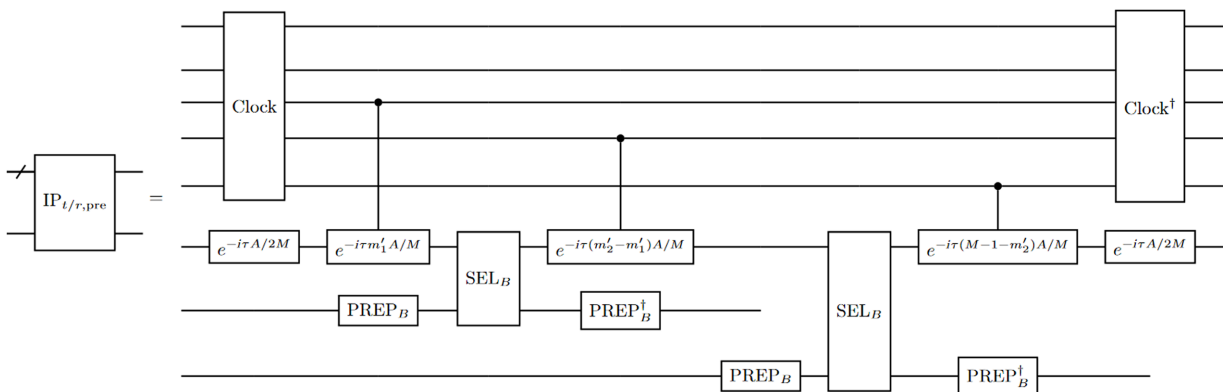
$$\epsilon_K = \exp(c/d) - \sum_{k=0}^K \frac{c^k}{d^k k!}, \quad \epsilon_\theta = \frac{\pi}{2^{n_\theta}}, \quad (I20)$$

$$\epsilon_t = \left(\frac{2\|A\|}{\lambda_B} + \frac{\|A\|^2}{\lambda_B^2} \right) \left[2^{n_t} \left(e^{\lambda_B \tau / 2^{n_t}} - 1 \right) - \lambda_B \tau (1 + \lambda_B \tau / 2^{n_t+1}) \right]. \quad (I21)$$

Proof. The complete interaction-picture simulation is indicated in Fig. 7, with the interaction-picture operator to be amplified shown in Fig. 8 and the clock preparation shown in Fig. 9. The total costs for each of the parts of the simulation are shown in Table VIII. The sum of the costs in entries 1 to 7 gives the cost in Eq. (I17). The costs in terms of the evolution under A and the block encoding of B are given in entries 8 and 9 of Table VIII.

Provided $\lambda_B t$ is rational, for any integer \mathcal{N} , $\lambda_B t / \mathcal{N}$ must be rational, and we can choose c and d to satisfy $c/d = \lambda_B \tau$ for $\tau = t/\mathcal{N}$. Then Eq. (I18) comes from Eq. (I12), and constrains the value of c/d . Since c/d is set by \mathcal{N} , we need a way to select \mathcal{N} so that c/d satisfies Eq. (I18). Choosing some value of b_r , we can find a lower bound P_{\min} on $P_s(\Sigma(0), b_r)$ (for example as given in Appendix J below). Then we can set $\mathcal{N} = \lceil \lambda_B t / \ln(2P_{\min}) \rceil$, and $\lambda_B \tau = \lambda_B t / \mathcal{N}$ is strictly less than $\ln(2P_{\min})$. Then $c/d = \lambda_B \tau < \ln(2P_{\min})$ implies that Eq. (I18) is satisfied.

The error bound in Eq. (I19) is from Eq. (I16). The quantities ϵ_K , ϵ_θ , and ϵ_t are bounds on the error due to truncating the Dyson series, the rotation on the ancilla qubit for the amplitude amplification, and the time discretization, respectively. The norm of each successive term in the Dyson series is bounded as $(\tau \lambda_B)^k / k!$, and $\lambda_B \tau = c/d$. Therefore the error in truncating the Dyson series to order


 FIG. 8. Circuit implementation of the preamplified operator ($K = 2$).

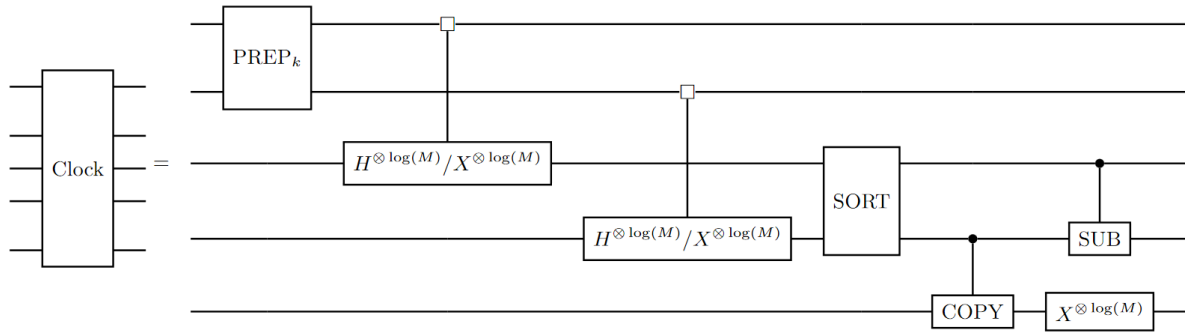


FIG. 9. Quantum circuit for preparing the clock state ($K = 2$). Note that the last time difference $M - 1 - m'_2$ can be computed with a copying operation followed by flipping the qubits.

K can be bounded by the error in truncating the exponential of c/d , giving the expression for ϵ_K in Eq. (I20). The expression for ϵ_θ in Eq. (I20) is from Eq. (I14). The bound in Eq. (I21) is from Eq. (H28), where we use λ_B as the upper bound on $\|B\|$. ■

The algorithm will be most efficient when c/d is a close rational approximation of $\ln 2$, so \mathcal{N} is minimized. This is because c/d is upper bounded by $\ln\{2P_s[\Sigma(0), b_r]\}$ according to Eq. (I18), $P_s[\Sigma(0), b_r]$ is close to 1, and $\mathcal{N} = \lambda_B t d / c$. It would be expected that for a real application one can choose the exact time to be simulated, so that c and d are relatively small (for example $9/13$ is less than but close to $\ln 2$). If evolution for a specific t is needed where $\lambda_B t$ is not rational, we prepare an ancilla register in the state

$$\cos \theta_B |0\rangle + \sin \theta_B |1\rangle \quad (I22)$$

and use the amplitude $\cos \theta_B$ to artificially adjust the normalization factor λ_B . This completes our costing of the interaction-picture algorithm for generic Hamiltonians.

APPENDIX J: BOUND ON THE SUCCESS PROBABILITY FOR PREPARING EQUAL SUPERPOSITION

Here we consider the probability of success for preparation of an equal superposition state and place a lower bound on it for given b_r . The probability for success was given in Ref. [74], and is

$$P_s(n, b_r) = \frac{n}{2^{\lceil \log n \rceil}} \left\{ \left[1 + \left(2 - \frac{4n}{2^{\lceil \log n \rceil}} \right) \sin^2 \theta(n, b_r) \right]^2 + \sin^2 [2\theta(n, b_r)] \right\}, \quad (J1)$$

where $\theta(n, b_r)$ is the angle of rotation on an ancilla qubit. It turns out that there is a simple lower bound for the success probability, though the expression for θ is adjusted slightly from the formula used before. The result is as in the following theorem.

Theorem 13. For all $n, b_r \in \mathbb{N}$, taking

$$x = \frac{n}{2^{\lceil \log n \rceil}}, \quad (J2)$$

TABLE VIII. The costs of the interaction-picture algorithm for time evolution of a general Hamiltonian. The costs are Toffoli costs unless otherwise specified (the last two entries).

Procedure	Cost
The preparation of the equal superposition over $\Sigma(0)$ numbers; see Eq. (I61).	$6\mathcal{N}(3n_k + 2b_r - 9)$
Inequality tests used for the preparation of the superposition over k ; see Eq. (I7).	$3\mathcal{N} \left(-1 + \sum_{k=0}^{K-1} \lceil \log(\Sigma(k)) \rceil \right)$
Controlled Hadamards for preparing the superpositions over times; see Eq. (I66).	$6\mathcal{N}K n_t$
The sorting of the time registers; see Eq. (I67).	$12\mathcal{N} n_t \text{St}(K)$
The differences of the times; see Eq. (I51).	$6\mathcal{N}(K-1)(n_t - 1)$
The preparation of θ state; see Eq. (I10).	$2\mathcal{N}(n_\theta - 3)$
The reflection cost for oblivious amplitude amplification; see Eq. (I11).	$2\mathcal{N} \lceil n_k + K(n_t + n_B) \rceil$
The exponential of A needs to be performed $K + 1$ times; see Eq. (I9).	$3 \lceil \mathcal{N}(K+1)n_t + 2 \rceil$ controlled- e^{isA}
Perform the controlled block encoding of operator B K times; see Eq. (I9).	$6\mathcal{N}K \text{PREP}_B + 3\mathcal{N}K \text{SEL}_B$

$$\theta_0 = \arcsin\left(\frac{1}{2\sqrt{x}}\right), \tag{J3}$$

$$(2x - 1)(3 - x) \geq 0,$$

$$(2x - 1)(3 - x) + 2 > 0,$$

$$4(1 - x)(2x - 1)[(2x - 1)(3 - x) + 2] \geq 0,$$

$$4(1 - x)(2x - 1)[(2x - 1)(3 - x) + 2] + 1 > 0,$$

$$[2(4x - 1 - 2x^2)]^2 - [(2x - 1)\sqrt{4x - 1}]^2 > 0. \tag{J8}$$

$$\theta(n, b_r) = \frac{2\pi}{2^{b_r}} \text{round} \left[\frac{2\pi}{2^{b_r}} \left(\theta_0 - \frac{(\pi/2^{b_r})^2(2x - 1)}{\sqrt{4x - 1}} \right) \right] \tag{J4}$$

gives the success probability for preparation of an equal superposition state lower bounded as

$$P_s(n, b_r) \geq 1 - \left(\frac{3}{2}\right)^2 \left(\frac{\pi}{2^{b_r}}\right)^2. \tag{J5}$$

Proof. Now with $b_r = 1$ or $b_r = 2$ the result is trivial, because the right-hand side is negative. We therefore aim to prove the result for $b_r \geq 3$. Because the expression for $P_s(n, b_r)$ depends only on the ratio $x = n/2^{\lceil \log n \rceil}$, we can prove the lower bound for all values of x in the range $x \in (1/2, 1)$, and that will give us the result for all $n \in \mathbb{N}$. Then the probability of success can be written as

$$P_s(n, b_r) = x \{ [1 + (2 - 4x) \sin^2 \theta]^2 + \sin^2(2\theta) \}. \tag{J6}$$

Let us take $\Delta\theta = \theta - \theta_0$. Using elementary trigonometric identities, the probability of success can be rewritten as

$$\begin{aligned} P_s(n, b_r) &= 1 - 4(1 - x)(4x - 1) \sin^2(\Delta\theta) \\ &\quad + 4(1 - x) [2(4x - 1 - 2x^2) \sin^2(\Delta\theta) \\ &\quad - (2x - 1)\sqrt{4x - 1} \sin(2\Delta\theta)] \sin^2(\Delta\theta). \end{aligned} \tag{J7}$$

To prove lower bounds for $x \in (1/2, 1)$, we use

$$2(4x - 1 - 2x^2) \sin^2(\Delta\theta) \geq (2x - 1)\sqrt{4x - 1} [\sin(2\Delta\theta) - 2\Delta\theta], \tag{J12}$$

which in turn implies

$$2(4x - 1 - 2x^2) \sin^2(\Delta\theta) - (2x - 1)\sqrt{4x - 1} \sin(2\Delta\theta) \geq -2(2x - 1)\sqrt{4x - 1} \Delta\theta \tag{J13}$$

for $x \in (1/2, 1)$ and $\Delta\theta \in [-0.7, 0.7]$. This means that the probability is lower bounded by

$$\begin{aligned} P_s(n, b_r) &\geq 1 - 4(1 - x)(4x - 1) \sin^2(\Delta\theta) - 8(1 - x)(2x - 1)\sqrt{4x - 1} \Delta\theta \sin^2(\Delta\theta) \\ &= 1 - 4(1 - x) [(4x - 1) + 2(2x - 1)\sqrt{4x - 1} \Delta\theta] \sin^2(\Delta\theta). \end{aligned} \tag{J14}$$

Next, $4x - 1 - 2x^2 = (2x - 1)(3/2 - x) + 1/2$ is positive for $x \in (1/2, 1)$, and $2x - 1$ and $4x - 1$ are non-negative. As a result, we have

$$2(4x - 1 - 2x^2) > (2x - 1)\sqrt{4x - 1}. \tag{J9}$$

We also find that

$$\sin^2 \Delta\theta - [\sin(2\Delta\theta) - 2\Delta\theta] = (\Delta\theta)^2 + \mathcal{O}[(\Delta\theta)^3], \tag{J10}$$

so it is non-negative for small $\Delta\theta$. Moreover this expression has zeros of its derivative for $n\pi$ and $n\pi - \arctan(1/2)$. This means the only zeros of its derivative for $\Delta\theta \in [-0.7, 0.7]$ are at 0 and $-\arctan(1/2) \approx -0.46$. This means that the function $\sin^2 \Delta\theta - [\sin(2\Delta\theta) - 2\Delta\theta]$ can turn around towards negative values for $\Delta\theta < -0.46$, but computing it at $\Delta\theta = 0.7$ gives about 0.000 466, which is still positive. This means that for $\Delta\theta \in [-0.7, 0.7]$ we have

$$\sin^2 \Delta\theta - [\sin(2\Delta\theta) - 2\Delta\theta] \geq 0, \tag{J11}$$

because negative values would require an additional turning point. That implies

Next, we have for $x \in (1/2, 1)$ and $\Delta\theta \in [-0.7, 0.7]$

$$\begin{aligned} 4(1-x)(3x-1) &\geq 0, \\ 4x-1-3(2x-1)^2 &\geq 0, \\ \sqrt{4x-1} &\geq \sqrt{3}(2x-1), \\ 4x-1 &\geq \sqrt{3}(2x-1)\sqrt{4x-1}, \\ 4x-1 &\geq -2(2x-1)\sqrt{4x-1}\Delta\theta. \end{aligned} \quad (\text{J15})$$

As a result, we can lower bound the success probability by

$$\begin{aligned} P_s(n, b_r) &\geq 1 - 4(1-x) \\ &\quad \times \left[(4x-1) + 2(2x-1)\sqrt{4x-1}\Delta\theta \right] (\Delta\theta)^2. \end{aligned} \quad (\text{J16})$$

Note also that the derivative of this expression with respect to $\Delta\theta$ is

$$-8(1-x) \left[(4x-1) + 3(2x-1)\sqrt{4x-1}\Delta\theta \right] \Delta\theta. \quad (\text{J17})$$

Now $4x-1 \geq \sqrt{3}(2x-1)\sqrt{4x-1}$ implies the expression in square brackets must be non-negative for $\Delta\theta \in [-1/\sqrt{3}, 1/\sqrt{3}]$. This implies that, provided we are considering $\Delta\theta \in [-1/\sqrt{3}, 1/\sqrt{3}]$, we can lower bound the probability of success in any region by the probabilities of success at the bounds of the region.

Now, using the formula in the theorem for θ , we have

$$\Delta\theta + \frac{(\pi/2^{b_r})^2(2x-1)}{\sqrt{4x-1}} \in [-\pi/2^{b_r}, \pi/2^{b_r}]. \quad (\text{J18})$$

Note that $\sqrt{4x-1} \geq \sqrt{3}(2x-1)$ implies

$$\frac{2x-1}{\sqrt{4x-1}} \leq \frac{1}{\sqrt{3}}. \quad (\text{J19})$$

Therefore the maximal value of $\Delta\theta$ with $b_r \geq 3$ is

$$\frac{(\pi/8)^2}{\sqrt{3}} + \pi/2^8 \approx 0.481734 < \frac{1}{\sqrt{3}}. \quad (\text{J20})$$

Because $\Delta\theta \in [-1/\sqrt{3}, 1/\sqrt{3}]$, we can lower bound the probability by the lower bounds at the extremal points in

the range of $\Delta\theta$, which give

$$\begin{aligned} P_s(n, b_r) &\geq 1 - 4(1-x)(4x-1) \left(\frac{\pi}{2^{b_r}} \right)^2 \\ &\quad + (2x-1)^2(1-x) \left(\frac{\pi}{2^{b_r}} \right)^4 \\ &\quad \pm \frac{24(2x-1)^3(1-x)}{\sqrt{4x-1}} \left(\frac{\pi}{2^{b_r}} \right)^5 \\ &\quad + \frac{8(2x-1)^4(1-x)}{4x-1} \left(\frac{\pi}{2^{b_r}} \right)^6. \end{aligned} \quad (\text{J21})$$

Next, we have that $\sqrt{4x-1} \geq \sqrt{3}(2x-1)$ implies

$$\begin{aligned} 20(2x-1)^2(1-x)\sqrt{4x-1} &\geq 20\sqrt{3}(2x-1)^3(1-x) \\ 20(2x-1)^2(1-x) &\geq \frac{24(2x-1)^3(1-x)}{\sqrt{4x-1}} \\ (2x-1)^2(1-x) \left(\frac{\pi}{2^{b_r}} \right)^4 &\geq \frac{24(2x-1)^3(1-x)}{\sqrt{4x-1}} \\ &\quad \times \left(\frac{\pi}{2^{b_r}} \right)^5 \end{aligned} \quad (\text{J22})$$

because $\pi/2^{b_r} < 1$. Hence we have

$$\begin{aligned} (2x-1)^2(1-x) \left(\frac{\pi}{2^{b_r}} \right)^4 \\ \pm \frac{24(2x-1)^3(1-x)}{\sqrt{4x-1}} \left(\frac{\pi}{2^{b_r}} \right)^5 &\geq 0, \end{aligned} \quad (\text{J23})$$

and the sixth-order term in Eq. (J21) is non-negative, so we obtain the lower bound

$$P_s(n, b_r) \geq 1 - 4(1-x)(4x-1) \left(\frac{\pi}{2^{b_r}} \right)^2. \quad (\text{J24})$$

Now $(1-x)(4x-1)$ has its maximal value of 9/16 for $x = 5/8$, so the lower bound becomes

$$P_s(n, b_r) \geq 1 - \left(\frac{3}{2} \right)^2 \left(\frac{\pi}{2^{b_r}} \right)^2 \quad (\text{J25})$$

as required. \blacksquare

If we take $b_r = 8$ for example, then the probability of success must be at least 0.999661. That is useful for the general interaction-picture simulations described in Appendix I, because then $\ln\{2P_s[\Sigma(0), 8]\} > 0.6928$. That is larger than 9/13, so one can use $c/d = 9/13$ as a good approximation of $\ln 2$ and still satisfy Eq. (I18).

APPENDIX K: SIMULATING FIRST QUANTIZED REAL-SPACE GRID REPRESENTATIONS

Whether the goal is to perform non-Born-Oppenheimer dynamics or to solve the electronic structure problem,

in order to simulate molecular or material systems on a quantum computer one must discretize the wave function and map it to qubits. Perhaps the most obvious representation involves projecting the wave function onto a real-space grid and encoding the positions of each particle explicitly in a quantum register. In this case, the computational basis states that span the wave function of interest might be stored in the quantum computer as $|r_1 r_2 \cdots r_\eta R_1 R_2 \cdots R_L\rangle$ where the r_i and R_ℓ index the coordinates of the η electrons and L nuclei, respectively, using binary registers with N grid points (each requiring $\log N$ qubits). Thus, this encoding would require $(\eta + L) \log N$ qubits.

Representing the potential part of the Hamiltonian is straightforward when using a real-space grid [the operators U , V , and V_{nuc} would take the same diagonal forms as they do in Eq. (1)]. However, representing the kinetic operators will require some thought. One can see that if basis functions are not overlapping in space (e.g., delta functions or finite elements with disjoint support) the integrals from the Galerkin discretization would suggest that the kinetic operator is also diagonal. To avoid this, one idea would be to take a finite-difference approach and approximate the Laplacian using a k -point stencil; this is exactly the approach taken in first quantization in Ref. [109] and described in second quantization in Appendix A of [29]. Another idea, similar to using what is referred to as a “discrete value representation” (DVR) [110], would be to exactly define the kinetic operator in its eigenbasis and approximate a transformation that relates the grid representation to the eigenbasis of the kinetic operator. For example, if using a regularly spaced grid then one might make the approximation that the quantum Fourier transform (QFT) diagonalizes the kinetic operator. After all, the kinetic operator is diagonal in the momentum basis

(plane waves) and plane waves are obtained as the continuous Fourier transform of delta functions. However, this is an approximation when using a finite grid; the QFT is a discrete Fourier transform (which is linear) and a linear combination of N plane waves cannot sum to a delta function. This sort of representation is employed by the first paper to consider quantum simulations of molecular systems using a grid representation by Kassal *et al.* [40]. Components of that algorithm were assessed for fault tolerance in the work of Jones *et al.* [111] but that work stopped short of estimating all of the constant factors associated with the algorithm realization.

In addition to the necessity of the approximations already mentioned there are other unfavorable characteristics of real-space grid representations for use in quantum simulation. For example, basis errors are no longer variationally bounded when deviating from a Galerkin discretization and extrapolations to the continuum limit might be less reliable. Furthermore, ways of softening the potential and removing core electrons (akin to pseudopotentials for plane-wave methods) are less developed for grids. Thus, the majority of work on simulating molecular systems focuses on Galerkin discretizations involving basis functions rather than grids. Nonetheless, in this section we briefly describe how one can block encode a real-space representation of the Hamiltonian with the same asymptotic complexity as the main approaches pursued in this paper. Since these grid representations are not periodic, it is possible that they might converge faster than plane waves when simulating nonperiodic systems. Also, since the block-encoding algorithms are entirely different, it is possible that this approach might have lower constant factors.

We can express the aforementioned first quantized real-space Hamiltonian as

$$H_{\text{BO}} = T + U + V + \frac{1}{2} \sum_{\ell \neq \kappa=1}^L \frac{\zeta_\ell \zeta_\kappa}{\|R_\ell - R_\kappa\|}, \quad H_{\text{nBO}} = T + T_{\text{nuc}} + U_{\text{nBO}} + V + V_{\text{nuc}}, \quad (\text{K1})$$

$$T = \sum_{i=1}^{\eta} \text{QFT}_i \left(\sum_{p \in G} \frac{\|k_p\|^2}{2} |p\rangle \langle p|_i \right) \text{QFT}_i^\dagger, \quad T_{\text{nuc}} = \sum_{\ell=\eta+1}^{L+\eta} \text{QFT}_\ell \left(\sum_{p \in G} \frac{\|k_p\|^2}{2 m_\ell} |p\rangle \langle p|_\ell \right) \text{QFT}_\ell^\dagger, \quad (\text{K2})$$

$$U = - \sum_{i=1}^{\eta} \sum_{\ell=1}^L \sum_{p \in G} \frac{\zeta_\ell}{\|R_\ell - r_p\|} |p\rangle \langle p|_i, \quad U_{\text{nBO}} = - \sum_{i=1}^{\eta} \sum_{\ell=\eta+1}^{L+\eta} \sum_{p, q \in G} \frac{\zeta_\ell}{2 \|r_p - r_q\|} |p\rangle \langle p|_i |q\rangle \langle q|_\ell, \quad (\text{K3})$$

$$V = \sum_{i \neq j=1}^{\eta} \sum_{p, q \in G} \frac{1}{2 \|r_p - r_q\|} |p\rangle \langle p|_i |q\rangle \langle q|_j, \quad V_{\text{nuc}} = \sum_{\ell \neq \kappa=\eta+1}^{L+\eta} \sum_{p, q \in G} \frac{\zeta_\ell \zeta_\kappa}{2 \|r_p - r_q\|} |p\rangle \langle p|_\ell |q\rangle \langle q|_\kappa, \quad (\text{K4})$$

where QFT_i is the usual quantum Fourier transform applied to register i [112]. As in the main text, we use atomic units where \hbar as well as the mass and charge of the electron are unity and $\|\cdot\|$ denotes a function taking the 2-norm of its argument. In the above expression ℓ and κ index nuclear degrees of freedom; thus, R_ℓ represents the positions of nuclei and ζ_ℓ the atomic numbers of nuclei. Furthermore, we have the following definition of grid points and their frequencies in the dual space defined by the QFT:

$$r_p = \frac{p \Omega^{1/3}}{N^{1/3}} \quad k_p = \frac{2\pi p}{\Omega^{1/3}} \quad p \in G$$

$$G = \left[-\frac{N^{1/3}-1}{2}, \frac{N^{1/3}-1}{2} \right]^3 \subset \mathbb{Z}^3, \quad (\text{K5})$$

where Ω is the volume of our simulation cell and N is the number of grid points in the cell. Again, this is the same representation used in the first work on quantum-simulating chemistry in first quantization, by Kassal *et al.* [40], well over a decade ago. We note that like with the plane waves, it would be possible to use different grids for the different particles but we avoid doing that here for simplicity.

We now show that one can simulate the Hamiltonian of Eq. (K1) with the same complexity as the approaches of this paper using either qubitization, or interaction-picture simulation frameworks. Here we discuss only the Born-Oppenheimer case for simplicity, but the techniques readily generalize to the non-Born-Oppenheimer Hamiltonian. Focusing first on the qubitization approach, the goal is to block encode the operators T , U , and V . The block encoding of T can be accomplished using almost the same techniques used for plane waves. The only difference comes in the implementation of the SEL operator. Whenever the ancilla register flags that a term from T should be applied, one must apply the QFT, then phase the register by the appropriate amount, then apply the inverse QFT. The QFT needs to be performed independently on η registers each of $\log N$ qubits, so has gate complexity $\tilde{\mathcal{O}}(\eta \log N)$, where here the $\tilde{\mathcal{O}}$ indicates terms double logarithmic in N are omitted [113]. That is complexity in terms of one- and two-qubit gates, and requires rotations up to $\mathcal{O}(\log N)$ bits of precision. The complexity in terms of Toffoli gates is $\tilde{\mathcal{O}}(\eta \log^2 N)$.

Here the block encoding of U and V will deviate from the other strategies of this paper. Our approach will be to break up the terms as a sum of many terms with equal magnitude coefficients that can be computed on the fly and applied as phases, mirroring a strategy first introduced in Ref. [114]. First, we begin by re-expressing the V operator

as

$$V = \frac{v_{\max}}{M} \sum_{i \neq j=1}^{\eta} \sum_{p,q \in G} \sum_{m=1}^M v(m, r_p, r_q) \times |p\rangle \langle p|_i |q\rangle \langle q|_j, \quad (\text{K6})$$

$$v(m, r_p, r_q) = \begin{cases} 1 & \frac{m v_{\max}}{M} < \frac{1}{2\|r_p - r_q\|}, \\ 0 & \text{otherwise} \end{cases}, \quad (\text{K7})$$

$$v_{\max} = \max_{p \neq q} \left(\frac{1}{2\|r_p - r_q\|} \right) = \frac{N^{1/3}}{2\Omega^{1/3}}, \quad (\text{K8})$$

where in the inequality above we take advantage of the fact that the repulsive Coulomb kernel is always greater than zero. From this definition we can see that

$$\left| \frac{1}{2\|r_p - r_q\|} - \frac{v_{\max}}{M} \sum_{m=1}^M v(m, r_p, r_q) \right| \leq \frac{v_{\max}}{M}. \quad (\text{K9})$$

Hence, we can make sure that the error in the Hamiltonian coefficients is less than ϵ if $M > v_{\max}/\epsilon$. This error can be suppressed exponentially since the cost of the comparison test is quadratic in the number of qubits used to represent the number (because we need to compare squares) [74], which is logarithmic in M , and thus logarithmic $1/\epsilon$.

We can now block encode V in the following way. We start by applying PREP as a series of Hadamard gates, which realize the equal superposition state over the ancilla registers $|i, j\rangle |m\rangle$ (essentially all of the complexity of our approach will enter through SEL). Then, controlled on register $|i\rangle$ we copy the momentum held in the i th system register to an ancilla register and translate it into r_p using the relation in Eq. (K5). We obtain r_q in the same way, but this time controlled on the register $|j\rangle$. Then, we compute $v(m, r_p, r_q)$ and record its value in an ancilla register. Because this ancilla register is initialized to $|0\rangle$ by PREP, if it is changed to the $|1\rangle$ state by SEL then it will be removed from the linear combination of unitaries. Indeed, we wish to remove the term from the linear combination of unitaries whenever the inequality in Eq. (K7) is false. Therefore, the final step of our procedure should be to apply a bit-flip operator to the ancilla holding $v(m, r_p, r_q)$ and to uncompute the r_p and r_q ancillae. Thus, SEL will act with the following steps (with p_i and p_j equivalent to p and q , respectively):

$$\text{SEL } |i, j\rangle |m\rangle |0\rangle^{\otimes \mathcal{O}(\log(N/\epsilon))} \underbrace{|p_0 \cdots p_i \cdots p_j \cdots p_\eta\rangle}_{|\psi\rangle}$$

$$\mapsto |i, j\rangle |m\rangle |r_{p_i}, r_{p_j}\rangle |0\rangle |p_0 \cdots p_i \cdots p_j \cdots p_\eta\rangle$$

$$\begin{aligned}
 & \mapsto |i, j\rangle |m\rangle |r_{p_i}, r_{p_j}\rangle |v(m, r_{p_i}, r_{p_j})\rangle \\
 & \otimes |p_0 \cdots p_i \cdots p_j \cdots p_\eta\rangle \\
 & \mapsto |i, j\rangle |m\rangle |0\rangle^{\otimes \mathcal{O}(\log(N/\epsilon))} |\neg v(m, r_{p_i}, r_{p_j})\rangle \\
 & \otimes |p_0 \cdots p_i \cdots p_j \cdots p_\eta\rangle
 \end{aligned} \tag{K10}$$

where \neg implies the negation of the following bit.

One might be concerned that we would need to compute a reciprocal and a square root as part of the 2-norm in the inequality for $v(m, r_p, r_q)$. However, we can much more efficiently compute Eq. (K7) by multiplying the norm through to the other side and squaring it; in other words we can evaluate $v(m, r_p, r_q)$ by testing

$$4 m^2 v_{\max}^2 \|r_p - r_q\|^2 < M^2. \tag{K11}$$

The cost of evaluating this expression is also $\mathcal{O}(\log^2 M) = \mathcal{O}[\log^2(1/\epsilon)]$ gates using the result of Lemma 9, which again contributes only a polylogarithmic contribution to the gate complexity. Apart from simplicity, this form of a comparison test is much easier (will have lower constant factors) to implement because the squared 2-norm is equivalent to just computing the sum of squares; whereas the reciprocal square root can be a costly function to evaluate directly [115,116]. We note that another approach to avoiding the computation of reciprocal square roots in the context of quantum simulating first quantized real-space Hamiltonians is discussed in Ref. [117], where it suggested that one take advantage of structure arising from the inverse Laplace transformed version of the Coulomb kernel represented in a Cartesian component-separated approach. While that approach may have some advantages for product-formula-based methods, it is unclear whether it might also be useful for further reducing the cost of block encoding the Coulomb operator.

The block encoding of U follows a similar pattern. In this case, we re-express the U operator as

$$\begin{aligned}
 U &= -\frac{u_{\max}}{M} \sum_{i=1}^{\eta} \sum_{\ell=1}^L \zeta_\ell \sum_{p \in G} \sum_{m=1}^M u(m, r_p, R_\ell) \\
 & \times |p\rangle \langle p|_i,
 \end{aligned} \tag{K12}$$

$$u(m, r_p, R_\ell) = \begin{cases} 1 & \text{if } \frac{m u_{\max}}{M} < \frac{1}{\|r_p - R_\ell\|}, \\ 0 & \text{otherwise} \end{cases}, \tag{K13}$$

$$u_{\max} = \max_{p, \ell} \left(\frac{1}{\|r_p - R_\ell\|} \right) = \mathcal{O} \left(\frac{N^{1/3}}{\Omega^{1/3}} \right). \tag{K14}$$

where we require that the R_ℓ do not overlap with any of the grid points and we once again take advantage of the fact

that the right-hand side of the inequality in $u(m, r_p, R_\ell)$ is always greater than zero. Similar to the strategy employed for V , here the SEL operation will compute $u(m, r_p, R_\ell)$ into an ancilla that is then negated [thus, excluding the cases where $u(m, r_p, R_\ell) = 0$ from the linear combination of unitaries]. But there are two main differences: (1) the value R_ℓ will need to be loaded using a QROM and (2) we have not incorporated the ζ_ℓ into the definition of the inequality function that is computed as part of SEL.

Instead, we include the ζ_ℓ through the PREP operation. At the same time we directly load the $|R_\ell\rangle$ values into superposition (rather than computing it from ℓ). Thus, controlled on a flag ancilla specifying that the U term is to be applied, our PREP circuit will act as

$$\text{PREP } |0\rangle^{\otimes \mathcal{O}(\log(N/\epsilon))} \mapsto \sum_{\ell=1}^L \sum_{i=1}^{\eta} \sum_{m=1}^M \sqrt{\frac{\zeta_\ell}{\eta M}} |R_\ell\rangle |i\rangle |m\rangle |0\rangle. \tag{K15}$$

Using the coherent aliasing strategies developed in Ref. [20], the weighted superposition over R_ℓ can be prepared with no more than ϵ error in the amplitudes at Toffoli complexity $L + \mathcal{O}[\log(1/\epsilon)]$. That will be a negligible additive cost compared to the rest of the SEL operation, which will have Toffoli complexity $\mathcal{O}(\eta)$ for charge neutral systems. Thus, up to exponentially small errors in PREP we can see that

$$\left| \sum_{\ell=1}^L \frac{\zeta_\ell}{\|r_p - R_\ell\|} - \sum_{\ell=1}^L \frac{u_{\max} \zeta_\ell}{M} \sum_{m=1}^M u(m, r_p, R_\ell) \right| \leq \frac{u_{\max} \lambda_\zeta}{M}, \tag{K16}$$

with $\lambda_\zeta = \sum_{\ell} \zeta_\ell$, and again we can make sure that the error in the Hamiltonian coefficients is less than ϵ if $M > u_{\max} \lambda_\zeta / \epsilon$. Once again, this ϵ will only enter the gate complexity inside of logarithms.

Therefore, we see that we can implement the block encoding of $(T + U + V)/\lambda$ with gate complexity $\tilde{\mathcal{O}}(\eta)$. Given our scheme, the λ_T value is exactly the same as in (25), and

$$\begin{aligned}
 \lambda_V &= \eta^2 v_{\max} = \mathcal{O} \left(\frac{\eta^2 N^{1/3}}{\Omega^{1/3}} \right) \\
 \lambda_U &= \eta \left(\sum_{\ell=1}^L \zeta_\ell \right) u_{\max} = \mathcal{O} \left(\frac{\eta^2 N^{1/3}}{\Omega^{1/3}} \right),
 \end{aligned} \tag{K17}$$

where these values of λ_V and λ_U involve slightly different constant factors compared to the forms in Eq. (25), but the same asymptotic scaling. Accordingly, we see that the overall asymptotic complexity of the qubitization approach

using this Hamiltonian would be

$$\begin{aligned} \tilde{\mathcal{O}}\left(\frac{(\lambda_T + \lambda_U + \lambda_V)\eta}{\epsilon}\right) &= \tilde{\mathcal{O}}\left(\frac{\eta^2 N^{2/3}}{\epsilon \Omega^{2/3}} + \frac{\eta^3 N^{1/3}}{\epsilon \Omega^{1/3}}\right) \\ &= \tilde{\mathcal{O}}\left(\frac{\eta^{4/3} N^{2/3} + \eta^{8/3} N^{1/3}}{\epsilon}\right) \end{aligned} \quad (\text{K18})$$

for $\eta \propto \Omega$, which is the same as in Eq. (26). It is likely that the QFT will be the dominant cost in this method. The complexity for the plane-wave approach tends to be dominated by the cost of controlled selection of the momentum registers, which has complexity $\mathcal{O}(\eta \log N)$ and is the only part that is polynomial in η . The QFT has scaling as $\tilde{\mathcal{O}}(\eta \log^2 N)$, which is larger and so should dominate the complexity.

Finally, we discuss how the interaction-picture scheme can be combined with the real space first quantized representation in order to suppress the complexity term proportional to $N^{2/3}$. To do this, we follow a similar approach as in the main paper with the block encodings of U and V as described in this Appendix. The main difference is that before evolving under the kinetic operator we must apply the QFT. The gate complexity of performing the QFT in each time step is $\tilde{\mathcal{O}}(\eta \log^2 N)$ Toffoli gates. A second difference is that it is no longer possible to keep the total kinetic energy in an ancilla register and update it when we block encode U and V . At each step the kinetic energy will need to be recomputed, which has complexity $\mathcal{O}(\eta \log^2 N)$. That is almost as large as the complexity scaling for the QFT. As a result there is the same asymptotic scaling as the interaction-picture algorithm in the main paper up to a logarithmic factor,

$$\tilde{\mathcal{O}}\left(\frac{(\lambda_U + \lambda_V)\eta}{\epsilon}\right) \tilde{\mathcal{O}}\left(\frac{\eta^3 N^{1/3}}{\epsilon \Omega^{1/3}}\right) = \tilde{\mathcal{O}}\left(\frac{\eta^{8/3} N^{1/3}}{\epsilon}\right). \quad (\text{K19})$$

Again, the complexity will be increased over the plane-wave approach primarily due to the QFT.

-
- [1] Richard P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
 - [2] Seth Lloyd, Universal quantum simulators, *Science* **273**, 1073 (1996).
 - [3] Alan Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon, Simulated quantum computation of molecular energies, *Science* **309**, 1704 (2005).
 - [4] Guang Hao Low and Isaac L. Chuang, Hamiltonian simulation by qubitization, *Quantum* **3**, 163 (2019).
 - [5] Guang Hao Low and Nathan Wiebe, Hamiltonian Simulation in the Interaction Picture, [arXiv:1805.00675](https://arxiv.org/abs/1805.00675) (2018).
 - [6] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma, Simulating

Hamiltonian Dynamics with a Truncated Taylor Series, *Phys. Rev. Lett.* **114**, 90502 (2015).

- [7] Nathan Wiebe and Christopher E. Granade, Efficient Bayesian Phase Estimation, [arXiv:1508.00869](https://arxiv.org/abs/1508.00869) (2015).
- [8] David Poulin, Alexei Y. Kitaev, Damian Steiger, Matthew Hastings, and Matthias Troyer, Fast Quantum Algorithm for Spectral Properties, *Phys. Rev. Lett.* **121**, 010501 (2017).
- [9] Jeongwan Haah, Matthew B. Hastings, Robin Kothari, and Guang Hao Low, Quantum algorithm for simulating real time evolution of lattice Hamiltonians, *SIAM Journal on Computing*, FOCS18–250–FOCS18–284 (2018).
- [10] Earl Campbell, Random Compiler for Fast Hamiltonian Simulation, *Phys. Rev. Lett.* **123**, 070503 (2019).
- [11] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu, Theory of Trotter Error with Commutator Scaling, *Phys. Rev. X* **11**, 011020 (2021).
- [12] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O’Rourke, Erika Ye, Austin J. Minnich, Fernando G. S. L. Brandao, and Garnet Kin-Lic Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, *Nat. Phys.* **16**, 205 (2020).
- [13] Burak Şahinoğlu and Rolando D. Somma, Hamiltonian simulation in the low energy subspace, [arXiv:2006.02660](https://arxiv.org/abs/2006.02660) (2020).
- [14] Minh C. Tran, Yuan Su, and Daniel Carney, and Jacob M. Taylor, Faster digital quantum simulation by symmetry protection, *PRX Quantum* **2**, 010323 (2021).
- [15] James D. Whitfield, Jacob Biamonte, and Alan Aspuru-Guzik, Simulation of electronic structure hamiltonians using quantum computers, *Mol. Phys.* **109**, 735 (2011).
- [16] David Wecker, Bela Bauer, Bryan K. Clark, Matthew B. Hastings, and Matthias Troyer, Gate-count estimates for performing quantum chemistry on small quantum computers, *Phys. Rev. A* **90**, 022305 (2014).
- [17] Matthew B. Hastings, Dave Wecker, Bela Bauer, and Matthias Troyer, Improving quantum algorithms for quantum chemistry, *Quantum Inf. Comput.* **15**, 1 (2015).
- [18] David Poulin, M. 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 Inf. Comput.* **15**, 361 (2015).
- [19] Ryan Babbush, Dominic W. Berry, Ian D. Kivlichan, Annie Y. Wei, Peter J. Love, and Alan Aspuru-Guzik, Exponentially more precise quantum simulation of fermions in second quantization, *New J. Phys.* **18**, 33032 (2016).
- [20] Ryan Babbush, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven, Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity, *Phys. Rev. X* **8**, 041015 (2018).
- [21] Dominic W. Berry, Maria Kieferová, Artur Scherer, Yuval R. Sanders, Guang Hao Low, Nathan Wiebe, Craig Gidney, and Ryan Babbush, Improved techniques for preparing eigenstates of fermionic hamiltonians, *Npj Quantum Inf.* **4**, 22 (2018).
- [22] Dominic W. Berry, Craig Gidney, Mario Motta, Jarrod McClean, and Ryan Babbush, Qubitization of arbitrary

- basis quantum chemistry leveraging sparsity and low rank factorization, *Quantum* **3**, 208 (2019).
- [23] Ian D. Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alan Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, *Phys. Rev. Lett.* **120**, 110501 (2018).
- [24] Ian D. Kivlichan, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Wei Sun, Zhang Jiang, Nicholas Rubin, Austin Fowler, Alán Aspuru-Guzik, Hartmut Neven, and Ryan Babbush, Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization, *Quantum* **4**, 296 (2020).
- [25] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [26] Vera von Burg, Guang Hao Low, Thomas Häner, Damian S. Steiger, Markus Reiher, Martin Roetteler, and Matthias Troyer, Quantum computing enhanced computational catalysis, [arXiv:2007.14460](https://arxiv.org/abs/2007.14460) (2020).
- [27] Joonho Lee, Dominic W. Berry, Craig Gidney, William J. Huggins, Jarrod R. McClean, Nathan Wiebe, and Ryan Babbush, Even more efficient quantum computations of chemistry through tensor hypercontraction, *PRX Quantum* **2**, 030305 (2021).
- [28] Yuan Su, and Hsin-Yuan Huang, and Earl T. Campbell, Nearly tight trotterization of interacting electrons, *Quantum* **5**, 495 (2021).
- [29] Ryan Babbush, Nathan Wiebe, Jarrod McClean, James McClain, Hartmut Neven, and Garnet Kin-Lic Chan, Low-Depth Quantum Simulation of Materials, *Phys. Rev. X* **8**, 011044 (2018).
- [30] Jacob T. Seeley, Martin J. Richard, and Peter J. Love, The Bravyi-Kitaev transformation for quantum computation of electronic structure, *J. Chem. Phys.* **137**, 224109 (2012).
- [31] Sergey Bravyi, Jay M. Gambetta, Antonio Mezzacapo, and Kristan Temme, Tapering off qubits to simulate fermionic Hamiltonians, [arXiv:1701.08213](https://arxiv.org/abs/1701.08213) (2017).
- [32] Kanav Setia and James D. Whitfield, Bravyi-Kitaev superfast simulation of fermions on a quantum computer, *J. Chem. Phys.* **148**, 164104 (2018).
- [33] Mario Motta, Erika Ye, Jarrod R. McClean, Zhendong Li, Austin J. Minnich, Ryan Babbush, and Garnet Kin-Lic Chan, Low rank representations for quantum simulation of electronic structure, *Npj Quantum Inf.* **7**, 83 (2021).
- [34] Nicholas P. Bauman, Eric J. Bylaska, Sriram Krishnamoorthy, Guang Hao Low, Nathan Wiebe, Christopher E. Granade, Martin Roetteler, Matthias Troyer, and Karol Kowalski, Downfolding of many-body hamiltonians using active-space models: Extension of the sub-system embedding sub-algebras approach to unitary coupled cluster formalisms, *J. Chem. Phys.* **151**, 014107 (2019).
- [35] Tyler Takeshita, Nicholas C. Rubin, Zhang Jiang, Eunseok Lee, Ryan Babbush, and Jarrod R. McClean, Increasing the Representation Accuracy of Quantum Simulations of Chemistry Without Extra Quantum Resources, *Phys. Rev. X* **10**, 011004 (2020).
- [36] Jarrod R. McClean, Fabian M. Faulstich, Qinyi Zhu, Bryan O’Gorman, Yiheng Qiu, Steven R. White, Ryan Babbush, and Lin Lin, Discontinuous galerkin discretization for quantum simulation of chemistry, *New J. Phys.* **22**, 093015 (2020).
- [37] Mario Motta, Tanvi P. Gujarati, Julia E. Rice, Ashutosh Kumar, Conner Masteran, Joseph A. Latone, Eunseok Lee, Edward F. Valeev, and Tyler Y. Takeshita, Quantum simulation of electronic structure with a transcorrelated hamiltonian: Improved accuracy with a smaller footprint on the quantum computer, *Phys. Chem. Chem. Phys.* **22**, 24270 (2020).
- [38] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer, Elucidating reaction mechanisms on quantum computers, *Proc. Natl. Acad. Sci.* **114**, 7555 (2017).
- [39] Zhendong Li, Junhao Li, Nikesh S. Dattani, C. J. Umrigar, and Garnet Kin-Lic Chan, The electronic complexity of the ground-state of the FeMo cofactor of nitrogenase as relevant to quantum simulations, *J. Chem. Phys.* **150**, 024302 (2019).
- [40] Ivan Kassal, Stephen P. Jordan, Peter J. Love, Masoud Mohseni, and Alan Aspuru-Guzik, Polynomial-time quantum algorithm for the simulation of chemical dynamics, *Proc. Natl. Acad. Sci.* **105**, 18681 (2008).
- [41] In the first paper to discuss error-correcting quantum algorithms for chemistry, Jones *et al.* [111] estimate the number of Toffoli gates required to implement subroutines (Trotter steps) of the Kassal *et al.* [40] algorithm. However, their work stops short of the error analysis required to predict the constant factors associated with performing fixed precision quantum simulation. For example, they do not estimate how many Trotter steps would be required to achieve target precision.
- [42] Ryan Babbush, Dominic W. Berry, Jarrod R. McClean, and Hartmut Neven, Quantum simulation of chemistry with sublinear scaling in basis size, *Npj Quantum Inf.* **5**, 92 (2019).
- [43] John C. Polanyi, Concepts in reaction dynamics, *Acc. Chem. Res.* **5**, 161 (1972).
- [44] Max Born and Robert Oppenheimer, Zur quantentheorie der molekeln, *Annalen der Physik* **389**, 457 (1927).
- [45] David R. Yarkony, Nonadiabatic quantum chemistry-, past, present, and future, *Chem. Rev.* **112**, 481 (2012).
- [46] Laurie J. Butler, Chemical reaction dynamics beyond the Born-Oppenheimer approximation, *Annu. Rev. Phys. Chem.* **49**, 125 (1998).
- [47] Basile F. E. Curchod and Todd J. Martínez, Ab initio nonadiabatic quantum molecular dynamics, *Chem. Rev.* **118**, 3305 (2018).
- [48] Xiaolei Zhu and David R. Yarkony, Non-adiabaticity: The importance of conical intersections, *Mol. Phys.* **114**, 1983 (2016).
- [49] R. Car and M. Parrinello, Unified Approach for Molecular Dynamics and Density-Functional Theory, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [50] Dominik Marx and Jürg Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, U.K., 2009).
- [51] Jürg Hutter, Car–Parrinello molecular dynamics, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 604 (2012).

- [52] Andreas Grüneis, James J. Shepherd, Ali Alavi, David P. Tew, and George H. Booth, Explicitly correlated plane waves: Accelerating convergence in periodic wavefunction expansions, *J. Chem. Phys.* **139**, 84112 (2013).
- [53] Christof Hättig, Wim Klopper, Andreas Köhn, and David P. Tew, Explicitly correlated electrons in molecules, *Chem. Rev.* **112**, 4 (2011).
- [54] Judith Harl and Georg Kresse, Cohesive energy curves for noble gas solids calculated by adiabatic connection fluctuation-dissipation theory, *Phys. Rev. B* **77**, 45136 (2008).
- [55] James J. Shepherd, Andreas Grüneis, George H. Booth, Georg Kresse, and Ali Alavi, Convergence of many-body wave-function expansions using a plane-wave basis: From homogeneous electron gas to solid state systems, *Phys. Rev. B* **86**, 35111 (2012).
- [56] Trygve Helgaker, Wim Klopper, Henrik Koch, and Jozef Noga, Basis-set convergence of correlated calculations on water, *J. Chem. Phys.* **106**, 9639 (1998).
- [57] Wim Klopper, Limiting values for Møller-Plesset second-order correlation energies of polyatomic systems: A benchmark study on Ne, HF, H₂O, N₂, and He...He, *J. Chem. Phys.* **102**, 6168 (1995).
- [58] Asger Halkier, Trygve Helgaker, Poul Jørgensen, Wim Klopper, Henrik Koch, Jeppe Olsen, and Angela K. Wilson, Basis-set convergence in correlated calculations on Ne, N₂, and H₂, *Chem. Phys. Lett.* **286**, 243 (1998).
- [59] Tosio Kato, On the eigenfunctions of many-particle systems in quantum mechanics, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [60] Richard M. Martin, *Electronic Structure* (Cambridge University Press, Cambridge, U.K., 2004).
- [61] Here and throughout the paper we use $\tilde{O}(\cdot)$ to indicate an asymptotic upper bound suppressing polylogarithmic factors and $o(1)$ to represent a positive number that approaches zero as some parameter grows.
- [62] Conventionally, “chemical accuracy” is defined as having precision in the energy that is better than 1 kcal/mol, or equivalently, 0.0016 Hartree. An error in the energy on the order of this quantity corresponds to an error in the predicted rates of chemical reactions by roughly an order of magnitude at room temperature [118].
- [63] Sam McArdle, Earl Campbell, and Yuan Su, Exploiting fermion number in factorized decompositions of the electronic structure Hamiltonian, [arXiv:2107.07238](https://arxiv.org/abs/2107.07238) (2021).
- [64] Ryan Babbush, Dominic W. Berry, Yuval R. Sanders, Ian D. Kivlichan, Artur Scherer, Annie Y. Wei, Peter J. Love, and Alan Aspuru-Guzik, Exponentially more precise quantum simulation of fermions in the configuration interaction representation, *Quantum Sci. Technol.* **3**, 015006 (2018).
- [65] The first quantized plane-wave algorithm of Ref. [42] has a Clifford complexity matching the Toffoli complexity whereas the second quantized arbitrary basis algorithm of Ref. [27] has Clifford complexity that is worse than the Toffoli complexity by a factor of $\tilde{O}(N)$. Hence, in terms of overall gate complexity the scaling to the continuum limit is actually $\tilde{O}(1/\epsilon^4)$ versus $\tilde{O}(1/\epsilon^{1/3})$ —a difference by a power of 12!
- [66] Alexei Y. Kitaev, Quantum measurements and the Abelian Stabilizer Problem, [arXiv:quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026) (1995).
- [67] Daniel S. Abrams and Seth Lloyd, Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors, *Phys. Rev. Lett.* **83**, 5162 (1999).
- [68] Andrew M. Childs and Nathan Wiebe, Hamiltonian simulation using linear combinations of unitary operations, *Quantum Inf. Comput.* **12**, 901 (2012).
- [69] Mario Szegedy, in *45th Annual IEEE Symposium on Foundations of Computer Science* (IEEE, 2004), p. 32.
- [70] Guang Hao Low and Isaac L. Chuang, Optimal Hamiltonian Simulation by Quantum Signal Processing, *Phys. Rev. Lett.* **118**, 010501 (2017).
- [71] Mária Kieferová, Artur Scherer, and Dominic W. Berry, Simulating the dynamics of time-dependent hamiltonians with a truncated dyson series, *Phys. Rev. A* **99**, 042314 (2019).
- [72] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Phys. Rev. A* **86**, 32324 (2012).
- [73] Guang Hao Low, Vadym Kliuchnikov, and Luke Schaeffer, Trading T-gates for dirty qubits in state preparation and unitary synthesis, [arXiv:1812.00954](https://arxiv.org/abs/1812.00954) (2018).
- [74] Yuval R. Sanders, Dominic W. Berry, Pedro C. S. Costa, Louis W. Tessler, Nathan Wiebe, Craig Gidney, Hartmut Neven, and Ryan Babbush, Compilation of fault-tolerant quantum heuristics for combinatorial optimization, *PRX Quantum* **1**, 020312 (2020).
- [75] Yuval R. Sanders, Guang Hao Low, Artur Scherer, and Dominic W. Berry, Black-Box Quantum State Preparation Without Arithmetic, *Phys. Rev. Lett.* **122**, 020502 (2019).
- [76] Roger A. Horn and Charles R. Johnson, *Matrix Analysis* (Cambridge university press, Cambridge, England, 2012).
- [77] Dominic W. Berry, Richard Cleve, and Sevag Gharibian, Gate-efficient discrete simulations of continuous-time quantum query algorithms, *Quantum Inf. Comput.* **14**, 0001 (2014).
- [78] Sheng-Tzong Cheng and Chun-Yen Wang, Quantum switching and quantum merge sorting, *IEEE Trans. Circuits Systems I: Regular Papers* **53**, 316 (2006).
- [79] Robert Beals, Stephen Brierley, Oliver Gray, Aram W. Harrow, Samuel Kutin, Noah Linden, Dan Shepherd, and Mark Stather, Efficient distributed quantum computing, *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **469**, 20120686 (2013).
- [80] Michael Codish, Luís Cruz-Filipe, Michael Frank, and Peter Schneider-Kamp, in *2014 IEEE 26th International Conference on Tools with Artificial Intelligence* (2014), p. 186.
- [81] A more efficient procedure for modeling nonperiodic systems using plane waves is to use a truncated Coulomb operator with a slightly larger cell size [98–100] or to subtract terms from the molecular multipole [119] or Martyna-Tuckerman corrections [120]. This can likely be incorporated into our approach in a straightforward fashion but would require minor modifications to the potential block encodings which we leave for future work.
- [82] John Preskill, Quantum computing and the entanglement frontier, [arXiv:1203.5813](https://arxiv.org/abs/1203.5813) (2012).

- [83] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven, Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595 (2018).
- [84] Many of the prior works we compare to, including Kim *et al.* [94], Kivlichan *et al.* [24], and Babbush *et al.* [20], report the number of T gates required rather than the number of Toffoli gates required. In order to facilitate a clear comparison we assume that the cost of a Toffoli gate is equal to the cost of 2 T gates. This equivalence is justified by the surface-code space-time volume required for a Toffoli gate compared to a T gate using the catalyzed T gate distillation scheme of Ref. [121].
- [85] Christopher J. Stein and Markus Reiher, AUTOCAS: A program for fully automated multiconfigurational calculations, *J. Comput. Chem.* **40**, 2216 (2019).
- [86] Elvira R. Sayfutyarova, Qiming Sun, Garnet Kin-Lic Chan, and Gerald Knizia, Automated construction of molecular active spaces from atomic valence orbitals, *J. Chem. Theory Comput.* **13**, 4063 (2017).
- [87] Sebastian Keller, Katharina Boguslawski, Tomasz Janowski, Markus Reiher, and Peter Pulay, Selection of active spaces for multiconfigurational wavefunctions, *J. Chem. Phys.* **142**, 244104 (2015).
- [88] Jie J. Bao and Donald G. Truhlar, Automatic active space selection for calculating electronic excitation energies based on high-spin unrestricted hartree-fock orbitals, *J. Chem. Theory Comput.* **15**, 5308 (2019).
- [89] Matthew R. Hermes, Riddish Pandharkar, and Laura Gagliardi, Variational localized active space self-consistent field method, *J. Chem. Theory Comput.* **16**, 4923 (2020).
- [90] Sergio Tosoni, Christian Tuma, Joachim Sauer, Bartolomeo Civalleri, and Piero Ugliengo, A comparison between plane wave and gaussian-type orbital basis sets for hydrogen bonded systems: Formic acid as a test case, *J. Chem. Phys.* **127**, 154102 (2007).
- [91] George H. Booth, Theodoros Tsatsoulis, Garnet Kin-Lic Chan, and Andreas Grüneis, From plane waves to local gaussians for the simulation of correlated periodic systems, *J. Chem. Phys.* **145**, 84111 (2016).
- [92] Leonard Kleinman and D. M. Bylander, Efficacious Form for Model Pseudopotentials, *Phys. Rev. Lett.* **48**, 1425 (1982).
- [93] Andrea Dal Corso, Pseudopotentials periodic table: From H to Pu, *Computational Materials Science* **95**, 337 (2014).
- [94] Isaac H. Kim, Eunseok Lee, Ye-Hua Liu, Sam Pallister, William Pol, and Sam Roberts, Fault-tolerant resource estimate for quantum chemical simulations: Case study on Li-ion battery electrolyte molecules, [arXiv:2104.10653](https://arxiv.org/abs/2104.10653) (2021).
- [95] V. E. Elfving, B. W. Broer, M. Webber, J. Gavartin, M. D. Halls, K. P. Lorton, and A. Bochevarov, How will quantum computers provide an industrially relevant computational advantage in quantum chemistry? [arXiv:2009.12472](https://arxiv.org/abs/2009.12472) (2020).
- [96] László Füsti-Molnar and Peter Pulay, Accurate molecular integrals and energies using combined plane wave and Gaussian basis sets in molecular electronic structure theory, *J. Chem. Phys.* **116**, 7795 (2002).
- [97] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp, in *Quantum Computation and Information*, edited by Vitaly I Voloshin, Samuel J. Lomonaco, and Howard E. Brandt (American Mathematical Society, Washington D.C., 2002), Chap. 3, p. 53.
- [98] Carlo A. Rozzi, Daniele Varsano, Andrea Marini, Eberhard K. U. Gross, and Angel Rubio, Exact coulomb cutoff technique for supercell calculations, *Phys. Rev. B* **73**, 205119 (2006).
- [99] Ravishankar Sundararaman and T. A. Arias, Regularization of the coulomb singularity in exact exchange by wigner-seitz truncated interactions: Towards chemical accuracy in nontrivial systems, *Phys. Rev. B* **87**, 165122 (2013).
- [100] Sohrab Ismail-Beigi, Truncation of periodic image interactions for confined systems, *Phys. Rev. B* **73**, 233103 (2006).
- [101] Daniel A. Lidar and Haobin Wang, Calculating the thermal rate constant with exponential speedup on a quantum computer, *Phys. Rev. E* **59**, 2429 (1999).
- [102] Ryan Babbush, Peter J. Love, and Alan Aspuru-Guzik, Adiabatic quantum simulation of quantum chemistry, *Sci. Rep.* **4**, 6603 (2014).
- [103] While this style of encoding second quantized operators is by far the most common, it is not necessary. One can also perform second quantized simulations in a fashion that requires $\mathcal{O}[(\eta + L) \log N]$ space rather than $\mathcal{O}[(\eta + L)N]$ space. For example, the work of Ref. [64] performs a quantum simulation of the electronic structure Hamiltonian in a fixed particle-number manifold of the second quantized Hamiltonian in a fashion that requires $\mathcal{O}(\eta \log N)$ qubits rather than $\mathcal{O}(N)$ qubits.
- [104] R. D. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Simulating physical phenomena by quantum networks, *Phys. Rev. A* **65**, 17 (2002).
- [105] F. Verstraete and J. I. Cirac, Mapping local hamiltonians of fermions to local hamiltonians of spins, *J. Stat. Mech.: Theory Exp.* **2005**, P09012 (2005).
- [106] Zhang Jiang, Jarrod McClean, Ryan Babbush, and Hartmut Neven, Majorana Loop Stabilizer Codes for Error Mitigation in Fermionic Quantum Simulations, *Phys. Rev. Appl.* **12**, 064041 (2019).
- [107] Craig Gidney, Halving the cost of quantum addition, *Quantum* **2**, 74 (2018).
- [108] Andrew M. Childs, Dmitri Maslov, Yunseong Nam, Neil J. Ross, and Yuan Su, Toward the first quantum simulation with quantum speedup, *Proc. Natl. Acad. Sci.* **115**, 9456 (2018).
- [109] Ian D. Kivlichan, Nathan Wiebe, Ryan Babbush, and Alan Aspuru-Guzik, Bounding the costs of quantum simulation of many-body physics in real space, *J. Phys. A: Math. Theor.* **50**, 305301 (2017).
- [110] John C. Light and Tucker Carrington Jr, Discrete-variable representations and their utilization, *Adv. Chem. Phys.* **114**, 263 (2000).
- [111] N. Cody Jones, James D. Whitfield, Peter L. McMahon, Man-Hong Yung, Rodney van Meter, Alan Aspuru-Guzik,

- and Yoshihisa Yamamoto, Faster quantum chemistry simulation on fault-tolerant quantum computers, *New J. Phys.* **14**, 115023 (2012).
- [112] Again, we emphasize that T is only approximately given by the expression involving the QFT. This relation is exact in the continuum limit. For finite-sized grids, it cannot be the case that the QFT completely diagonalizes the momentum operator.
- [113] Richard Cleve and John Watrous, in *Proceedings 41st Annual Symposium on Foundations of Computer Science* (IEEE, 2000), p. 526.
- [114] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma, in *STOC '14 Proceedings of the 46th Annual ACM Symposium on Theory of Computing* (2014), p. 283.
- [115] Mathias Soeken, Martin Roetteler, Nathan Wiebe, and Giovanni De Micheli, in *2017 54th ACM/EDAC/IEEE Design Automation Conference (DAC)* (IEEE, 2017), p. 1.
- [116] Thomas Häner, Martin Roetteler, and Krysta M. Svore, Optimizing Quantum Circuits for Arithmetic, [arXiv:1805.12445](https://arxiv.org/abs/1805.12445) (2018).
- [117] Bill Poirier and Jonathan Jerke, Full-Dimensional Schrödinger Wavefunction Calculations using Tensors and Quantum Computers: the Cartesian component-separated approach, [arXiv:2105.03787](https://arxiv.org/abs/2105.03787) (2021).
- [118] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, New York, 2000).
- [119] G. Makov and M. C. Payne, Periodic boundary conditions in ab initio calculations, *Phys. Rev. B* **51**, 4014 (1995).
- [120] Glenn J. Martyna and Mark E. Tuckerman, A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters, *J. Chem. Phys.* **110**, 2810 (1999).
- [121] Craig Gidney and Austin G. Fowler, Efficient magic state factories with a catalyzed $|CCZ\rangle$ to $2|T\rangle$ transformation, *Quantum* **3**, 135 (2019).

Correction: Headings for Sec. II C and Sec. II D contained typographical errors introduced during the production process and have been set right.