Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation

Ryan Babbush, Jarrod McClean, Dave Wecker, Alán Aspuru-Guzik, and Nathan Wiebe

Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation

Ryan Babbush,[1] Jarrod McClean,[2] Dave Wecker,[1] Alán Aspuru-Guzik,[2] and Nathan Wiebe[1]

[1]*Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052*
[2]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138*

(Dated: January 20, 2015)

Although the simulation of quantum chemistry is one of the most anticipated applications of quantum computing, the scaling of known upper bounds on the complexity of these algorithms is daunting. Prior work has bounded errors due to Trotterization in terms of the norm of the error operator and analyzed scaling with respect to the number of spin orbitals. However, we find that these error bounds can be loose by up to sixteen orders of magnitude for some molecules. Furthermore, numerical results for small systems fail to reveal any clear correlation between ground state error and number of spin orbitals. We instead argue that chemical properties, such as the maximum nuclear charge in a molecule and the filling fraction of orbitals, can be decisive for determining the cost of a quantum simulation. Our analysis motivates several strategies to use classical processing to further reduce the required Trotter step size and to estimate the necessary number of steps, without requiring additional quantum resources. Finally, we demonstrate improved methods for state preparation techniques which are asymptotically superior to proposals in the simulation literature.

I. INTRODUCTION

The idea that the simulation of quantum systems would be efficient on a quantum computer dates back to Feynman's original work on quantum mechanical computers [1]. Almost a decade after Abrams and Lloyd [2] demonstrated a scalable scheme for the quantum simulation of fermions, Aspuru-Guzik *et al.* [3] proposed that these techniques could be used to efficiently determine the ground state energy of molecular Hamiltonians, solving what chemists refer to as the electronic structure problem. Since then, a great deal of work has focused on specific strategies for the quantum simulation of quantum chemistry. While most of these approaches are based on a second quantized representation of the problem making use of both phase estimation and Trotterization [3–13], recently some have proposed alternative schemes such as the quantum variational eigensolver [14], an adiabatic algorithm [15] and an oracular approach based on a 1-sparse decomposition of the configuration interaction Hamiltonian [16]. In fact, quantum chemistry is such a popular application that toy problems in chemistry have been solved on a variety of experimental quantum information processors which include quantum optical systems [14, 17], nuclear magnetic resonance [18, 19] and solid-state Nitrogen-vacancy center systems [20].

Recently, a series of papers [10–13] has provided improved analytical and empirical bounds on the resources required to simulate classically intractable benchmarks using a quantum computer. While the initial findings in [10] were pessimistic, improvements in both bounds and algorithms introduced in [11] and [12] have reduced these estimates by more than thirteen orders of magnitude for simulations of Ferredoxin. The primary contribution of [13] was to point out that in the limit of large molecules, the use of a local basis can substantially reduce asymptotic complexity of these algorithms. In this paper we build on the findings of [10–13] to offer new perspectives regarding the scaling of the second quantized, Trotter-

ized, phase estimation algorithm for quantum chemistry. In particular, we question a basic assumption implicit in all of these works: that the Trotter error explicitly depends on the number of spin orbitals being simulated.

Instead, we argue that chemical properties such as the filling fraction of electrons in a given basis, the particular choice of orbital basis and the nuclear potential play a more significant role in determining the Trotter error than does the number of spin orbitals for small molecules. We support these arguments with numerical analysis based on the explicit computation of the Trotter error operator derived in [12]. Additionally, we show that classically tractable approximations to the ground state wavefunction can be used to efficiently estimate the Trotter error expected in a particular ground state simulation. This result is of significant practical importance because without a procedure for estimating the Trotter error, one must rely on analytical error bounds which (as we show) tend to overestimate the ground state error by many orders of magnitude. Finally, we show asymptotically improved circuits for state preparation based on these classical ansatz states.

A. The electronic structure problem

The electronic structure problem is to estimate the energy of electrons interacting in a fixed nuclear potential to within an additive error of ϵ. This Hamiltonian may be written as,

$$H = -\sum_i \frac{\nabla_{r_i}^2}{2} - \sum_{i,j} \frac{Z_i}{|R_i - r_j|} + \sum_{i,j>i} \frac{1}{|r_i - r_j|} \quad (1)$$

where we have used atomic units, $\{R_i\}$ denotes nuclear coordinates, $\{r_i\}$ electronic coordinates, and $\{Z_i\}$ nuclear charge. Often times, the utility of these energies is to provide Born-Oppenheimer surfaces for molecular modeling at finite temperatures. Usually, chemists are

interested in obtaining free energy landscapes which provide mechanistic insight into chemical events of significant practical importance such as drug binding, catalysis and material properties. These free energy landscapes must be extremely accurate as chemical rates are exponentially sensitive to changes in free energy. Under typical laboratory conditions of room temperature and atmospheric pressure, "chemical accuracy" is required which sets ϵ to the order of 10^{-3} hartree [21] where 1 hartree is $\frac{\hbar^2}{m_e e^2 a_0^2}$ and m_e, e and a_0 denote the mass of an electron, charge of an electron and Bohr radius, respectively.

We represent the electronic structure Hamiltonian in second quantization [21] as this requires significantly fewer qubits than approaches using the first quantized Hamiltonian [22, 23],

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s \qquad (2)$$

in which creation and annihilation operators act on a basis of orthogonal spin orbitals, $\{\varphi_i\}$ and the one-electron and two-electron integrals are

$$h_{pq} = \int d\sigma \ \varphi_p^*(\sigma) \left(-\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \varphi_q(\sigma) \quad (3)$$

$$h_{pqrs} = \int d\sigma_1 \ d\sigma_2 \ \frac{\varphi_p^*(\sigma_1)\varphi_q^*(\sigma_2)\varphi_s(\sigma_1)\varphi_r(\sigma_2)}{|r_1 - r_2|} \qquad (4)$$

where σ_i contains spatial and spin degrees of freedom for the electrons. The operators $a_p^\dagger$ and a_r obey the fermionic anti-commutation relations

$$\{a_p^\dagger, a_r\} = \delta_{p,r}, \qquad \{a_p^\dagger, a_r^\dagger\} = \{a_p, a_r\} = 0. \quad (5)$$

In principle, the number of spin orbitals used to represent a molecule is not a property of the molecule. However, the quantum chemistry community has certain conventions (based on periodic trends) for the number of spin orbitals that should be used for each atom in the period table, depending on the desired level accuracy in the calculation. In a minimal basis, first period atoms receive two spin orbitals, second period atoms receive 10 spin orbitals and third period atoms receive 18 spin orbitals. The reasoning behind this scheme is that the most important orbitals are those which have a principal quantum number less than or equal to that of the highest occupied orbital according to Hund's rules.

In addition to choosing a spatial basis, one must choose an orbital basis that associates orthogonal spatial functions constructed from the spatial basis, with the second quantized sites. Throughout this paper we investigate three such orbital basis sets: the "local basis" is the set of orthogonal atomic orbitals discussed in [13], the "canonical basis" is the Hartree-Fock molecular orbitals, and the "natural basis" is that which diagonalizes the one-electron density matrices associated with the exact

ground state[1]. It is worth pointing out that the canonical orbitals are the natural orbitals of a Hartree-Fock calculation using a single determinant.

From Eq. (2), we see that the number of terms in the Hamiltonian scales as $\Theta\left(N^4\right)$[2]. However, McClean et al. [13] recently pointed out that the basis functions decay super-exponentially with distance in a local basis. This means that the integrals in Eq. (3) and Eq. (4) will be negligibly small for many of the orbitals which in turn allows the number of terms in the Hamiltonian to be truncated to $\tilde{\mathcal{O}}\left(N^2\right)$ or $\tilde{\mathcal{O}}\left(N\right)$ depending on the size and geometry of the molecule. All of the particular benchmarks studied in this paper involve less than four atoms and so we consider the number of non-negligible terms in the Hamiltonian to scale as $\Theta(N^4)$, even in a local basis.

B. Quantum simulation of quantum chemistry

The electronic structure problem is classically intractable to current methods even after discretizing the Hilbert space. This intractability can be understood as a consequence of the exponential size of the Hilbert space for the second quantized Hamiltonian. Similarly, existing methods such as configuration interaction, require consideration of a number of electronic configuration states that increases exponentially as the approximation becomes more exact. Quantum simulation offers a way to circumvent these challenges by directly mapping the chemical system onto a set of qubits that can be manipulated using a quantum computer. The particular problem that we focus on is the problem of computing the ground state energy of the system. Other important physical quantities such as dipole moments can be found by evaluating their expectation value with respect to the prepared state. The simulation problem that we consider is as follows.

Problem: *Assume that the user is provided with a classical database containing h_{pq} and h_{pqrs} for a molecule with N spin orbitals and a blackbox state preparation algorithm that prepares an approximation $|\tilde{0}\rangle$ to the ground state $|0\rangle$ such that $|\langle\tilde{0}|0\rangle|^2 \in \Omega\left(poly\left(N^{-1}\right)\right)$. Design a quantum circuit that uses these elements to estimate the ground state energy of Eq. (2) within additive error ϵ using a minimal expected number of gates and qubits.*

---

[1] The natural basis can be well approximated without performing an exact calculation by repeating truncated configuration interaction calculations from reference states defined using the natural orbitals associated with a previous solution.

[2] We use the typical computer science convention that $f \in \Theta(g)$, for any functions f and g, if f is asymptotically upper and lower bounded by a multiple of g, $\mathcal{O}$ indicates an asymptotic upper bound, $\tilde{\mathcal{O}}$ indicates an asymptotic upper bound up to polylogarithmic factors, Ω indicates the asymptotic lower bound and $f \in o(g)$ implies $f/g \to 0$ in the asymptotic limit.

Most proposals for quantum computer simulation of chemical systems use similar strategies to solve this problem. The first step involves translating the basis of the second quantized Hamiltonian to that of the quantum computer. The standard way to do this is to use the occupation number basis in which individual quits encode the occupation of a spin orbital. For example, the state $|00011\rangle$ would refer to an electronic state where the first two spin orbitals are occupied and the remaining three spin orbitals are unoccupied.

Although representing states is trivial, representing the Hamiltonian is not. The reason is that, although it may seem that the creation and annihilation operators $a_i^\dagger$ and a_i are translated to $(X_i - iY_i)/2$ and $(X_i + iY_i)/2$ respectively, the resulting operators do not obey the anti-commutation relations in Eq. (5). This problem is addressed by using either the Jordan-Wigner transformation [3, 24] or the Bravyi-Kitaev transformation [9, 25] to modify these operators to have the correct anti-commutation relations. Importantly, the operators that result from using either of these representations are tensor products of Pauli operations. While the number of such terms in the transformed Hamiltonian scales as $\mathcal{O}(N^4)$ using both approaches, the locality (i.e. many-body order) of these terms scales as $\mathcal{O}(N)$ under the Jordan-Wigner transformation and $\mathcal{O}(\log N)$ under the Bravyi-Kitaev transformation [9].

Since exponentials of a polynomial number of Pauli operators are known to be efficiently simulatable, $e^{-iHt}|\tilde{\psi}\rangle$ can be implemented using a polynomial number of gates using a quantum computer. There are many different approaches that can be used to achieve this and the majority of these rely on Trotter decompositions, which we will discuss in more detail later. However, each of these methods solves a dynamical simulation problem and does not directly solve the ground state energy estimation problem. The phase estimation algorithm provides the connection needed to relate the eigenvalue estimation problem to a dynamical simulation problem.

The quantum phase estimation algorithm (PEA) uses a quantum computer to efficiently estimate energies from the phases $\{\theta_n(t)\}$ accumulated during time evolution under a propagator $U_H(t)$ associated with the Hamiltonian of interest H; i.e.

$$e^{iHt}|n\rangle = U_H(t)|\psi_n\rangle = e^{i\theta_n(t)}|\psi_n\rangle \tag{6}$$
$$\theta_n(t) = (E_n t) \mod 2\pi \tag{7}$$

where $\{|\psi_n\rangle\}$ and $\{E_n\}$ represent eigenstates and eigenvalues of H. If we initialize a quantum register in a state $|\tilde{\psi}_0\rangle$ then time evolution under a static Hamiltonian produces the superposition,

$$U_H(t)|\tilde{\psi}_0\rangle = \left(\sum_{n=0}^{2^N-1} e^{i\theta_n(t)}|\psi_n\rangle\langle\psi_n|\right)|\tilde{\psi}_0\rangle. \tag{8}$$

Measuring the phase of this superposition projects the system to state $|\psi_0\rangle$ with probability $|\langle\psi_0|\tilde{\psi}_0\rangle|^2$. Thus,
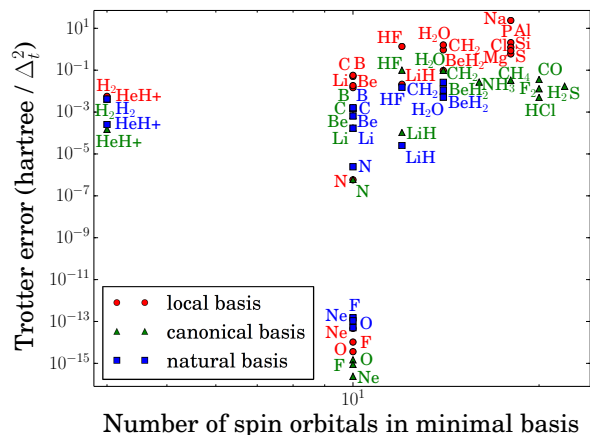

FIG. 1. Spin-orbitals versus ground state Trotter error for various molecular benchmarks in three different basis sets. Despite analytical predictions to the contrary (in prior works), it would appear that no clear relation holds between the Trotter error induced on the ground state and the number of spin orbitals for these benchmarks.

under the assumptions of our problem, at most a polynomial number of repetitions of the phase estimation algorithm will be needed to find the ground state energy.

There are obviously two contributions to the cost of solving the electronic structure problem via quantum computing: (a) the overlap $|\langle\psi_0|\tilde{\psi}_0\rangle|^2$ and (b) the cost of simulating the dynamics of the system. Since the overlap is independent of the simulation method used (to second-order in perturbation theory) most work on the topic has focused on reducing the latter cost. We discuss both of these issues in the following.

Our main focus is on Trotter-Suzuki based methods, which involve a discretization of the time evolution known as Trotterization. Trotterization approximates $U_H(t)$ as a series of time steps known as "Trotter slices" during which only one of the Hamiltonian terms is actually active. A Trotter series containing μ Trotter slices is said to have a "Trotter number" of μ and the error in this approximation, which arises from non-commutativity of the Hamiltonian terms, vanishes as $\mu \to \infty$. For a fixed order Trotter-Suzuki formula, each Trotter slice contains a number of gates that is proportional to the number of terms in the Hamiltonian, m. The value of m depends on basis and molecular size and its scaling with N ranges from $\tilde{\mathcal{O}}(N) - \tilde{\mathcal{O}}(N^4)$. Since the the total complexity of the quantum simulation circuit for chemistry is $\tilde{\mathcal{O}}(m\mu)$, understanding how both of these terms scales is vital for determining whether quantum chemistry will be viable on small scale quantum computers.

The big question that several recent papers have attempted to address is: "how does μ scale with N?" Indeed, this issue is central to the optimizations introduced in many of these simulation methods. Given the importance of this issue in the literature, the data in Figure 1 may come as a complete surprise. We see there that

for modestly small molecules, the error in the second-order Trotter-Suzuki formula does not have a clear functional dependence on N. This is especially surprising for cases of canonical and natural orbitals where there is little evidence of even an increasing trend in the error as a function of N. This lack of monotonicity is particularly striking for the atoms N, O, F, Ne which show negligibly small Trotter errors. In fact, for these molecules (along with others such as Helium Hydride and Lithium Hydride) $\mu = 1$ or $\mu = 2$ is sufficient to achieve chemical accuracy despite the fact that their Hamiltonians contain hundreds of non-commuting terms.

In order to understand why the Trotter error deviates so strongly from prior expectations, we analyze a leading order perturbative expression for the error in the second-order Trotter formula. The insights gained from this analysis raise an interesting point: although there is not a strong correlation between N and the Trotter error, other chemical properties play a decisive role in the Trotter error. This forces us to reconsider how we conceptualize the scaling of quantum chemistry simulation relative to prior results in quantum simulation, e.g. [2, 3, 11–13, 23, 26–29].

II. ANALYSIS OF TROTTER ERROR OPERATOR

The second-order Trotter-Suzuki decomposition allows us to approximate the propagator as a series of unitaries corresponding to the individual Hamiltonian terms. In particular, the second-order[3] Trotter formula gives us,

$$U_H^{\text{TS}}\left(\Delta_t\right) \equiv \prod_{\alpha=0}^{m-1} U_{m-\alpha}\left(\frac{\Delta_t}{2}\right) \prod_{\alpha=1}^{m} U_\alpha\left(\frac{\Delta_t}{2}\right) \quad (9)$$

where,

$$U_\alpha\left(\frac{\Delta_t}{2}\right) = e^{-iH_\alpha \Delta_t/2}. \quad (10)$$

The second-order formula applies each unitary twice with the second half of the Trotter series in reverse order of the first half to cancel out error terms in the ground state energy that would arise at first-order in Δ_t. We use this to make the approximation, valid for sufficiently small values of Δ_t, that

$$U = e^{iHt} \approx \left(U^{\text{TS}}\left(\Delta_t\right)\right)^\mu, \quad \Delta_t = t/\mu. \quad (11)$$

Poulin *et al.* [12] focus on bounding the error in this approximation with the Baker-Campbell-Hausdorff (BCH)

formula,

$$\log\left(e^X e^Y\right) = X + Y + \frac{1}{2}[X, Y] \quad (12)$$
$$+ \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + ...$$

By recursively applying Eq. (12) to Eq. (9), the error operator may be written as $V = \sum_{j=1}^{\infty} V^{(j)}$. The leading order term in this expansion is,

$$V^{(1)} = -\frac{\Delta_t^2}{12} \sum_{\alpha \leq \beta} \sum_{\beta} \sum_{\gamma < \beta} \left[H_\alpha\left(1 - \frac{\delta_{\alpha,\beta}}{2}\right), [H_\beta, H_\gamma]\right] \quad (13)$$

with errors on the order of $\mathcal{O}\left(\Delta_t^4\right)$.

The leading order shift in the energy of the ith eigenstate is given by non-degenerate perturbation theory as,

$$\Delta E_i = \langle\psi_i| V^{(1)} |\psi_i\rangle + \mathcal{O}\left(\Delta_t^4\right) \quad (14)$$

where $H |\psi_i\rangle = E_i |\psi_i\rangle$. Solving the electronic structure problem requires fixed precision in the energy, i.e. $\Delta E = \mathcal{O}(1)$. This suggests that we must shrink the time step for larger problem instances in order to offset any increase in Trotter error. In order to make the leading order shift in the energy eigenvalue at most δ it suffices to take

$$\mu = \mathcal{O}\left(t\sqrt{\frac{1}{\delta}\left\langle\sum_{\alpha \leq \beta}\sum_{\beta}\sum_{\gamma < \beta}\left[H_\alpha\left(1 - \frac{\delta_{\alpha,\beta}}{2}\right), [H_\beta, H_\gamma]\right]\right\rangle}\right). \quad (15)$$

Higher-order Trotter-Suzuki algorithms can be used to reduce the scaling of μ; however they require a number of gates that scales exponentially with the order of the Trotter formula. This means that for many problems with modest error tolerances, the second-order Trotter formula Eq. (9) yields the most efficient results. Although a similar expression based on degenerate perturbation theory must be used for molecules near disassociation, in most practical cases Eq. (15) will accurately predict the required Trotter number in the limit of small δ.

In practice, it is difficult to determine precisely how this error scales with problem size for real molecules. By inspection of Eq. (13), a loose bound of $\mu = \mathcal{O}\left(N^5\right)$ is obtained [12]. This bound is obtained by recognizing that the double commutator sum in Eq. (13) contains $\mathcal{O}(N^{12})$ terms but only $\mathcal{O}(N^{10})$ such terms are non–zero. In some cases, such as large molecules represented in a local orbital basis, many of these interactions can be neglected and the actual scaling of μ needed to achieve chemical accuracy may be closer to $\mu = \tilde{\mathcal{O}}\left(N^3\right)$ or $\mu = \tilde{\mathcal{O}}\left(N^{3/2}\right)$.

All of these scalings follow from worst case assumptions about the error and liberal application of the triangle inequality. Such arguments are not sufficient to explain the data in Figure 1 which does not show a clear dependence of μ on N. We therefore focus in the remainder on two quantities: (a) the error in the ground state energy

---

[3] Note that in work that focuses on high-order Trotter-Suzuki formulas Eq. (9) is often called the first-order Trotter Suzuki formula because it is the lowest iteration order in Suzuki's iterative construction of high-order splitting formulas.

and (b) the operator norm of the Trotter error operator. While (a) is the best measure of the error in quantum chemistry simulation, we also focus on (b) because it upper bounds (a) and because it can be well approximated without diagonalizing the Hamiltonian.

In the numerics that follow we construct error operators by explicitly computing all $\mathcal{O}(N^{10})$ nonzero terms in Eq. (13). Once all the terms in the error operator are constructed, we simplify the resulting expression by normal-ordering the result. Here normal-ordering refers to a sorting process where any chain of creation and annihilation operators that result from Eq. (13) are reordered such that creation operators always occur at the leftmost part of the chain. This reordering is done by using the anti-commutation relations in Eq. (5). For example, $a_2 a_1 a_1^\dagger a_3^\dagger = a_1^\dagger a_3^\dagger a_1 a_2 - a_3^\dagger a_2$. These normal-ordered terms are then grouped, allowing their actions on computational basis states to be efficiently computed.

The Trotter scheme we investigate does not use the coalescing strategies introduced in [12], which would surely lead to even more error cancellation. We use a minimal spatial basis (STO-6G). The Trotter series is ordered in the "interleaving" scheme introduced in [11] and PQRS terms are ordered lexicographically. All molecular integrals in this work were calculated at equilibrium configurations using the GAMESS electronic structure package [30, 31]. While computing the error operator is efficient, evaluating the error operator on an eigenstate of the Hamiltonian cannot be performed in polynomial time on a classical computer. Due to the expensive nature of these calculations, we limit our investigation to benchmarks containing less than twenty spin orbitals. We study the scaling of the norm of the Trotter error operator as this quantity is the focus of analytical bounds introduced in [10] and [12]. Though the bounds in [12] are based on a upper bound for the operator norm of the error operator, here we use the exact value of $\|V^{(1)}\|$.

A. Comparison of norm of error operator and ground state error

An important question to ask is, "how does the error in the simulated ground state energy compare to that predicted by the norm of the error operator?" This is important for two reasons. The first reason is that there can be substantial cancellation in the sum implicit in Eq. (14). This effect is also discussed in [12]. The second reason is that the ground state may only have limited overlap with the eigenstates of the error operator that have large eigenvalues. We will discuss these two effects in detail later, but for now it suffices to ask how substantial the differences between the two measures are.

Figure 2 shows that substantial differences exist between the computed Trotter error and the norm of the error operator. In particular, for O, F and Ne these discrepancies can be as large as sixteen orders of magnitude. Other molecules, such as H_2O and HF differ by
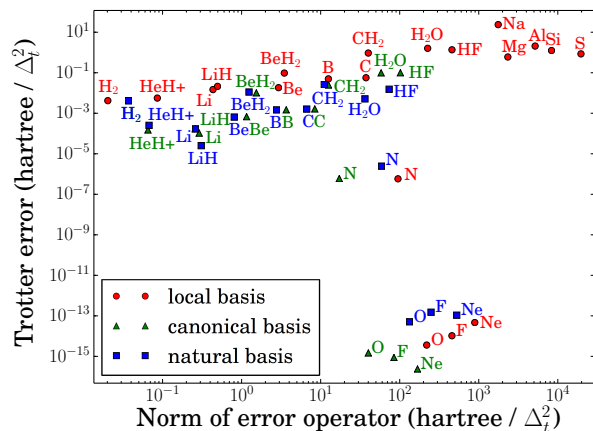

FIG. 2. A comparison between the norm of the error operator and the error induced in the ground state. Notice that in many cases the basis of natural orbitals have the lowest Trotter error (especially for examples with large Trotter error).

only two orders of magnitude. This shows that existing estimates of the error can ludicrously overestimate the error in Trotter Suzuki formulas if the properties of the ground state are not also taken into account. Similar comparable results have also been observed for random many-body Hamiltonians [32].

To see this, let us consider Ne. By the convention for second-period atoms, Ne is given 10 spin orbitals in a minimal basis but it also has 10 electrons. This means that all of its spin orbitals will be occupied, i.e. $|\psi_0\rangle = |1\rangle^{\otimes 10}$. If we consider the action of a single normal-ordered term from Eq. (14), $\alpha\, a_{p_1}^\dagger \cdots a_{p_5}^\dagger a_{q_1} \cdots a_{q_5}$, then we see that $\langle\psi_0|\, \alpha\, a_{p_1}^\dagger \cdots a_{p_5}^\dagger a_{q_1} \cdots a_{q_5}\, |\psi_0\rangle = 0$ unless $\{p_1, \ldots, p_5\} = \{q_1, \ldots, q_5\}$ up to permutations. Thus, the vast majority of the terms present in the error operator will evaluate to zero, irrespective of the magnitude of their coefficients. A similar argument can be made for F and O except that the ground state will no longer precisely be the Hartree-Fock state and instead will be a linear combination of computational basis states. Nonetheless, it is easy to see that the vast majority of these expectation values will be zero for these highly constrained systems. We therefore expect from this argument that molecules that have spin orbitals that are nearly fully occupied will have abnormally low error compared to molecules that are half filled where the dimension of the space is maximal for a given number of basis functions. This not only justifies the shockingly small error in N, O, F, and Ne but also explains why only considering the norm of the error operator obscures this trend.

For most benchmarks there is still evidence of correlation between the norm of the error operator and the Trotter error. This means that trends in the norm of the error operator are often reflected in the simulation error. As we have seen, the properties of the molecules in question can change the nature of this relationship.

Basis	Type	Orbitals	error / norm
STO-6G	local	4	0.2063
3-21G	local	8	0.0568
6-31G	local	8	0.0592
6-31++G	local	12	0.0328
STO-6G	canonical	4	0.1131
3-21G	canonical	8	0.0231
6-31G	canonical	8	0.0242
6-31++G	canonical	12	0.0108
STO-6G	natural	4	0.1131
3-21G	natural	8	0.0472
6-31G	natural	8	0.0547
6-31++G	natural	12	0.0194

TABLE I. Ratio of ground state error to error operator norm for molecular hydrogen in various basis sets.

B. Dependence on basis

In addition to showing that Trotter error in the ground state is usually substantially less than the error operator norm, Figure 2 suggests that the error is also basis dependent. While previous works have focused on the local and canonical basis sets, this figure suggests that using natural orbitals can often lower Trotter error by several orders of magnitude relative to a local orbital basis.

Furthermore, we argue that the discrepancy between error norm and ground state error increases with the number of spin orbitals to such an extent that the former should not be used to make arguments about the asymptotic scaling of the latter. One can always add more spin orbitals to a molecular Hamiltonian but given a reasonable orbital basis, the ground state and physically meaningful excited states will have increasingly limited occupancy in high energy orbitals. In this context, the energy of an orbital is understood to mean the energy of a single electron occupying that orbital in the absence of other electrons (appropriate for atomic orbitals) or in the presence of the average density of all other electrons (appropriate for the canonical orbitals). Additionally, the natural orbital basis is known to have the property that states with an odd number of excitations from ground state reference often have negligible overlap with the exact ground state [33].

While the error operator will inevitably contain many terms involving excitations to and from these high energy spin orbitals, eigenstates of physical interest (e.g. the ground state) are superpositions of configurations which have a limited number of excitations. Accordingly, terms involving combinations of high energy orbitals are not expected to significantly contribute to the error induced in relevant eigenstates despite increasing the norm of the error operator. This principle is demonstrated in Table I which shows the ratio between ground state error and error norm for molecular hydrogen in various basis sets.
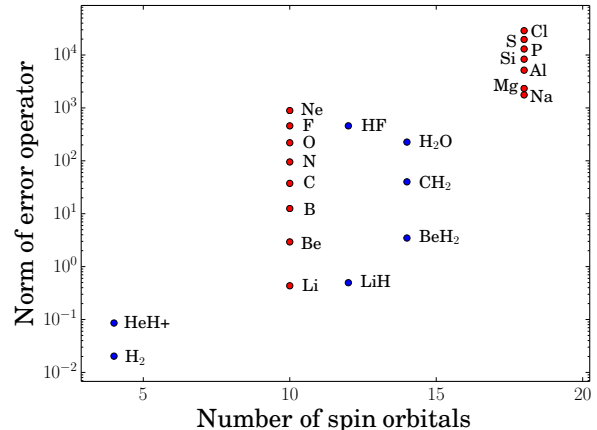

FIG. 3. We correlate the number of spin orbitals with the norm of the error operator in the local basis. The semblance of a positive slope appears to be a symptom of increasing nuclear charge as the number of spin orbitals increase. Red dots are atoms and blue dots are molecules.

C. Dependence on nuclear charge

Figure 3 indicates that Trotter error norm correlates especially well with the maximum nuclear charge, as further demonstrated in Figure 4. The local basis is formed from the set of orthogonal atomic orbitals which are obtained for molecules using Löwdin symmetric orthogonalization on the original non-orthogonal local Gaussian orbitals [13]. These Gaussian basis functions are constructed as approximations to eigenfunctions of Hydrogen-like systems, with some fitting adjustments. As such, we can determine the scaling behavior by considering the eigenfunctions of Hydrogen-like systems which are simple enough to permit analytical determination of how each term in the Hamiltonian will scale with nuclear charge. We begin by writing the eigenfunctions of a single electron in the potential of a point charge Z in a convenient way,

$$\psi_{n\ell m}(\rho,\theta,\phi) = \tag{16}$$
$$\sqrt{\left(\frac{2Z}{n}\right)^3 \frac{(n-\ell-1)!}{2n(n+\ell)!}} e^{-\frac{\rho}{n}} \left(\frac{2\rho}{n}\right)^\ell L_{n-\ell-1}^{2\ell+1}\left(\frac{2\rho}{n}\right) Y_\ell^m(\theta,\phi)$$

where $\rho = rZ$, $L_{n-\ell-1}^{2\ell+1}\left(\frac{2\rho}{n}\right)$ is a generalized Laguerre polynomial of degree $n-\ell-1$, and $Y_\ell^m(\theta,\phi)$ is a spherical harmonic of degree ℓ and order m. With the convention,

$$\varphi_p(\sigma_i) = \psi_p(\rho_i,\theta_i,\phi_i)\chi(s_i) \propto Z^{3/2} \tag{17}$$

$$d\sigma_i = \frac{\rho_i^2 \, d\rho_i}{Z^3} \sin(\theta_i) \, d\theta_i \, d\phi_i \, ds_i \propto Z^{-3} \tag{18}$$

$$\nabla^2 = Z^2 \left(\frac{\partial^2}{\partial\rho^2} + \frac{2}{\rho}\frac{\partial}{\partial\rho}\right) + \frac{Z^2}{\rho^2 \sin^2(\theta)}\frac{\partial^2}{\partial\phi^2}$$

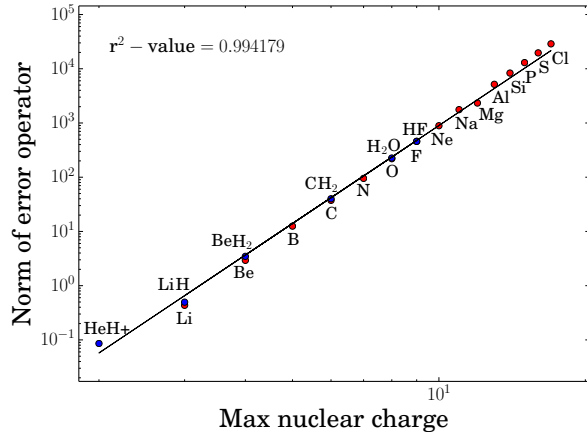$$+ \frac{Z^2}{\rho^2 \sin^2(\theta)}\frac{\partial^2}{\partial\phi^2} \propto Z^2 \tag{19}$$

FIG. 4. The norm of the error operator appears extremely well correlated with the maximum nuclear charge in a molecule when using a local basis of atomic orbitals. The black line is the line of best fit for a Z_{max}^6 scaling.
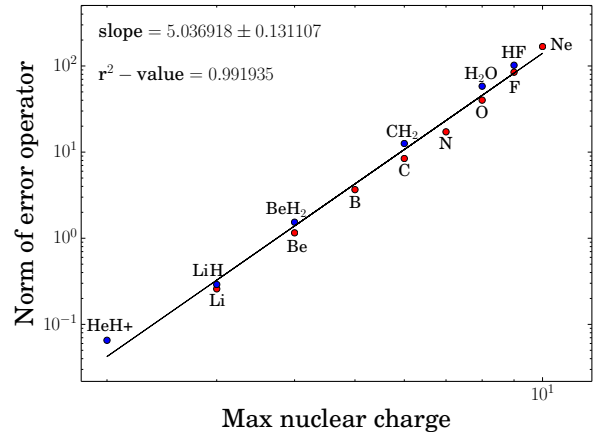

FIG. 5. The norm of the error operator appears also well correlated with the maximum nuclear charge in a molecule when using the canonical basis of molecular orbitals. The black line is a least squares fit to the data which is roughly consistent with a Z_{max}^5 scaling.

where $\chi(s_i)$ is a spin assignment and σ represents all degrees of freedom for an electron, we rewrite Eq. (3) and Eq. (4) in terms of ρ, assuming a single nuclei,

$$h_{pq} = \int d\sigma \; \varphi_p^*(\sigma) \left(-\frac{\nabla^2}{2} - \frac{Z^2}{\rho} \right) \varphi_q(\sigma) \qquad (20)$$

$$h_{pqrs} = \int d\sigma_1 \, d\sigma_2 \frac{\varphi_p(\sigma_1) \, \varphi_q(\sigma_2) \, \varphi_s(\sigma_1) \, \varphi_r(\sigma_2)}{|\rho_1 - \rho_2|/Z}. \qquad (21)$$

For both integrals, factors of Z from the differential volume elements $d\sigma$ cancel with factors of Z from the spin orbitals φ and we find that,

$$|h_{pq}| = \Theta\left(Z^2\right) \qquad (22)$$

$$|h_{pqrs}| = \Theta\left(Z\right). \qquad (23)$$

Thus, it is clear that we can upper bound the scaling of individual Hamiltonian terms with nuclear charge as $\mathcal{O}\left(Z_{max}^2\right)$. While this result is rigorous only when the orbital basis is the basis of true atomic orbitals, we expect qualitatively similar behavior in other bases. Assuming the h_{pq} terms dominates the error in the Trotter formula then Eq. (13) implies that the Trotter error should scale as $\mathcal{O}(Z_{max}^6)$. This scaling is qualitatively consistent with the empirical scaling in Figure 4 which fits Z_{max}^6 scaling to the norm of the error operator with an r^2-value of 0.994. Comparable results to this scaling have also been observed in diffusion Monte Carlo algorithms [34, 35].

These results imply that if an atomic basis is used then the error in the second-order Trotter-Suzuki formula scales at most as

$$\|V^{(1)}\| \in \mathcal{O}\left(N^4 Z_{max}^6 + N^{10} Z_{max}^3\right). \qquad (24)$$

This result is a direct consequence of bounds on the Trotter-Suzuki error in [11] and the observation that double commutators of the one- and two-body terms produce at most N^4 and N^{10} terms respectively. This implies that the computational complexity of performing the simulation on an arbitrary state, given fixed error tolerance of chemical accuracy, is $\mathcal{O}(N^4(N^2 Z_{max}^3 + N^5 Z_{max}^{3/2}))$. However, our numerical results are consistent with an $\mathcal{O}(N^4 Z_{max}^3)$ which suggests that this scaling may be loose. It also important to note that the gate depth can be further reduced by using interleaving and nesting as per [11], which is significant when the algorithm is implemented on systems where quantum operations can be executed in parallel. It is also worth noting that the one-body terms dominate the two-body terms in every numerical example that we considered. Larger molecules with more h_{pqrs} terms may lead to Trotter errors that scale as $\mathcal{O}(Z_{max}^3)$ rather than $\mathcal{O}(Z_{max}^6)$. More extensive numerical results may be needed to determine the conditions under which the two-body terms asymptotically dominate the one-body terms (if such conditions exist).

Figure 5 shows that these error estimates are pessimistic for the molecules considered when using the canonical basis. While the error norm is still strongly correlated to nuclear charge, unlike the scaling in the local basis, the fit to a Z_{max}^6 scaling is less convincing. Instead, the data empirically seems to follow a Z_{max}^5 scaling. Intuitively, this is easy to envision because the molecular orbitals are inherently delocalized and thus it is natural to expect that the maximum nuclear charge should make less of an impact in this basis. We also see no evidence of explicit scaling with N over this range in Z_{max}. It is interesting to note that although the number of non-negligible integrals in a local orbital basis can be quadratically or quartically smaller than the size of an untruncated canonical molecular orbital basis, the scaling with Z_{max} seems to be better by a linear factor. This suggests interesting trade-offs between the two methods and hints that neither is intrinsically superior for quantum simulation.
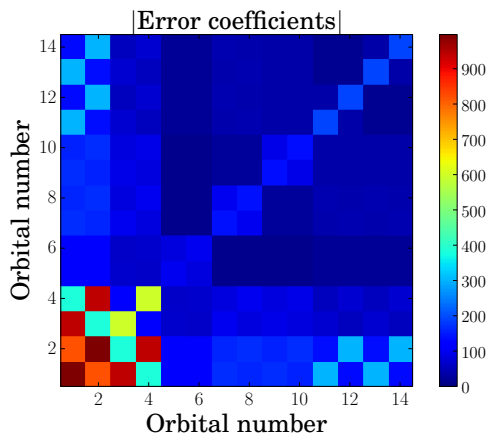
FIG. 6. This plot shows the coefficients of normal-ordered terms in the error operator for water in a local basis as a function of the orbitals on which they act. The coefficients of the error terms are binned according to the orbitals involved in the term. This plot shows the marginal distribution of the magnitudes of those terms.
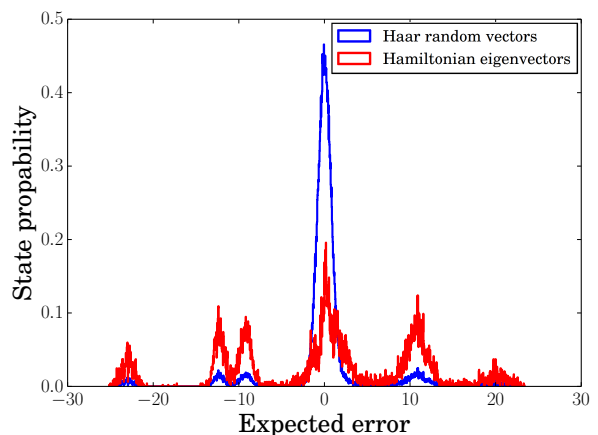

FIG. 7. This plot shows the distribution of expectation values of the error operator for water in the local basis over its eigenstates and Haar random vectors. We see that the random vectors lead to substantially less error, on average, than do the Hamiltonian eigenstates. The Haar distribution of errors has a standard deviation of 4.82 while the Hamiltonian error distribution has standard deviation of 10.68.

D. Dependence on orbital structure

The terms that appear in the error operator include interactions between every orbital in the basis set. This begs the question of whether terms in the Hamiltonian that involve particular orbitals have larger contribution to the error. In order to assess this, we compute the error operator for a number of different molecules and normal-order the resultant operator. We then sum the magnitudes of every remaining term that either create or annihilate an electron in each of the orbitals. An example of this is provided in Figure 6, which shows the marginal coefficient magnitudes of all terms in the error operator (after normal-ordering) in terms of two spin orbitals they contain. Appendix B shows similar analysis for other molecules in other basis sets. As we can see, terms which involve the inner shell electrons dominate the norm of the error operator in the local basis.

We see from such figures that the inner orbitals, especially the single particle terms which are on the diagonal of the plot above, have a substantial impact on the Trotter error. This is not surprising as the inner atomic orbitals interact very strongly with nuclei so the single particle integrals are likely to be much larger than the interaction integrals for these orbitals. Interestingly, although the valence shell electrons are often the most important for determining the chemical properties of a molecule, the inner orbitals are the ones that affect the error most significantly. This suggests that pseudo-potentials, which allow the core electrons to be treated as effectively "frozen", may provide a way to reduce the Trotter error in some circumstances. We leave this as an open question for future work.

E. Dependence on structure of eigenstates

Due to the substantial discrepancy between error induced on the ground state and operator norm, we might ask the following question: given the error operators for real molecules, what is the distribution of errors that would be induced on a random ensemble of vectors? This question is important as the answer will help us to identify the source of the observed error cancellation. We consider the ensemble of Haar random vectors which form a unitarily invariant ensemble of vectors with uniformly distributed complex elements. Unitary invariance ensures that the ensemble has uniform distribution in an arbitrary complete, orthonormal basis such as the eigen-basis of the error operator.

Denoting vectors from the random ensemble as $|v\rangle$ and eigenvectors of the error operator as $|k\rangle$ with eigenvalue λ_k, we are interested in analyzing properties of the following distribution of expected errors given by,

$$\Delta E(v) = \sum_k \lambda_k |\langle v | k \rangle|^2. \tag{25}$$

First, note that $\sum_k \lambda_k = 0$. This is because if $C = [A, B] = \sum_j \lambda_k |k\rangle\langle k|$ then

$$\sum_k \lambda_k = \text{Tr}(C) = \text{Tr}(AB) - \text{Tr}(BA) = 0, \tag{26}$$

from the cyclic property of the trace. Since $V^{(1)}$ is the sum of such operators, it follows that its trace is also zero. This implies that the Haar-expectation value of the error, over all possible random states, is

$$\mathbb{E}_H(\Delta E(v)) = \sum_k \lambda_k \mathbb{E}_H |\langle v | k \rangle|^2 = \frac{1}{2^N} \sum_k \lambda_k = 0. \tag{27}$$

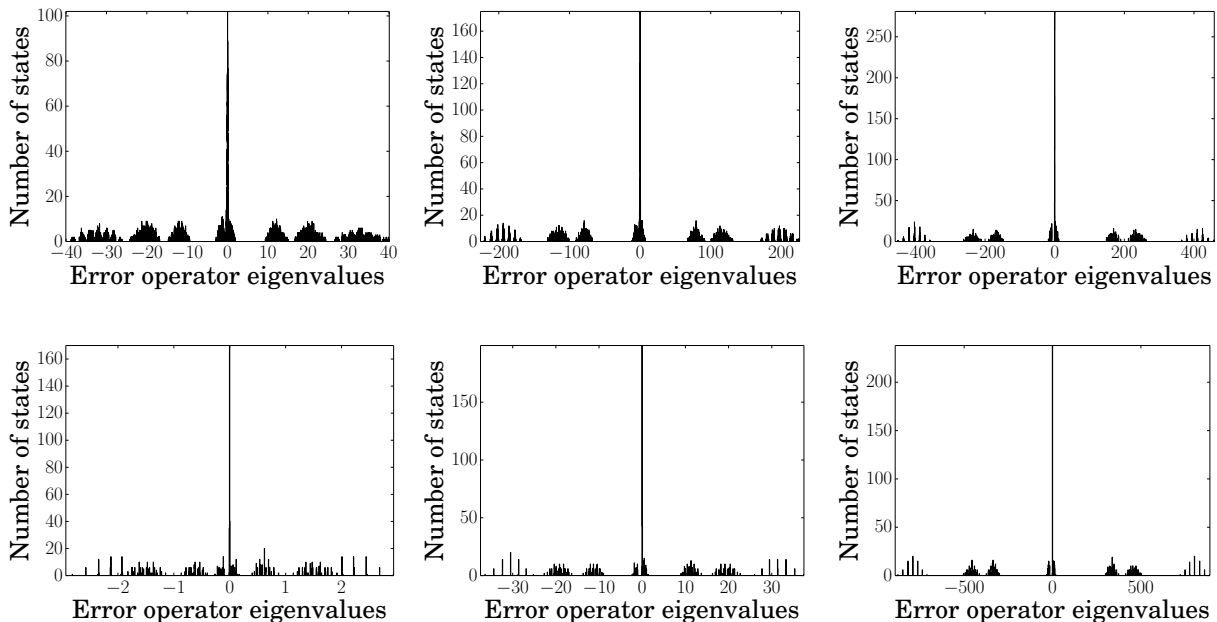FIG. 8. These are histograms of the eigenspecta of the error operators for various molecular and atomic benchmarks in the local basis. Proceeding clockwise from the top left, the molecules are water, Hydrogen Fluoride, Methylene, atomic Beryllium, atomic Carbon and atomic Neon. Error operators for all of our benchmarks have surprisingly similiar eigenspectra, regardless of the orbital basis. The source of this striking similarity and the reason for the particular structure is unknown.

This shows that there is no inherent bias that arises from Trotterization towards either overestimating or underestimating the true expectation value.

This result does not represent the *typical error* that we expect to see in a simulation. We also need to find the Haar variance of the expected error to estimate the typical variation of simulation errors about the mean. It is then easy to see that the Haar variance is

$$\mathbb{V}_H(\sum_k \lambda_k |\langle v|k\rangle|^2) = \sum_k \lambda_k^2 \mathbb{V}_H(|\langle v|k\rangle|^2). \qquad (28)$$

In Appendix A, we derive the Haar variance of the squared projection,

$$\mathbb{V}_H\left(|\langle v|k\rangle|^2\right) = \frac{2}{2^N(2^N+1)} - \frac{1}{2^{2N}}, \qquad (29)$$

where N is the number of spin orbitals. Combining Eq. (29) and Eq. (28) and using Chebyshev's inequality, we see that with high probability over $|v\rangle$

$$|\langle v| V^{(1)} |v\rangle| \in O\left(\frac{\sqrt{\sum_k \lambda_k^2}}{2^N}\right). \qquad (30)$$

Eq. (30), surprisingly, shows that a concentration of measure argument causes the expectation of the Trotter error to be asymptotically zero if (a) $|v\rangle$ is typical of a Haar random state, (b) $\sum_k \lambda_k^2 \in o(2^{2N})$ and (c) $|v\rangle$ is chosen independently of the $|k\rangle$.

We do not expect a concentration of measure argument like this to hold for actual quantum simulations because it would imply that the Trotter errors *in eigenvalue estimation* shrink rapidly with system size for physically reasonable distributions of λ_k. Thus, it is natural to expect that one or both of assumptions (a) and (c) are not reasonable for eigenvalue estimation.

In Figure 7, we show the expected errors according to Eq. (25) over an ensemble of Haar random vectors as well as the expected errors over the eigenstates of the Hamiltonian for water. The results clearly show that the errors observed in this chemical example are much greater than we would expect from Haar random states. Furthermore, we see little evidence of concentration of measure of the errors about zero for the case where $|v\rangle$ is an eigenvector of H; whereas the Haar random $|v\rangle$ lead to results that are much more concentrated about zero error. This suggests that the discrepancies between the norm of the error operator and the ground state error cannot be explained by a simple randomization argument as the actual errors observed are much worse than would be otherwise expected.

Eq. (25) shows that the expected error is the convolution of the functions λ_k and $|\langle v|k\rangle|^2$. Thus, we expect the distribution of errors to resemble the underlying distribution of eigenvalues of $V^{(1)}$. This intuition can easily be seen by comparing Figure 7 to the eigenspectrum of the water error operator in Figure 8a. As expected, the distribution of errors for the random ensemble (Figure 7) resembles the error operator eigenspectrum (Fig-

ure 8a) with concentration about the mean (as anticipated by Eq. (30)). Also, it is interesting to note that the eigenspectra of the error operators for various molecules and atoms studied in this paper bear a remarkable degree of similarity and appear extremely structured as Figure 8 demonstrates. Additionally, every example has a sharp peak in its spectrum about zero error. This suggests that much of the rift between the norm of the error operator in Figure 2 may be due to the large number of eigenvectors with near-zero eigenvalue.

III. IMPROVED SIMULATION METHODS INSPIRED BY CLASSICAL APPROACHES TO QUANTUM CHEMISTRY

Given the large disparity between error operator norm and error induced on the exact ground state, any efficient method which allows one to approximate the error induced on the ground state (which implies an estimate for the number of Trotter steps needed) would be of critical importance for anyone wishing to actually run a quantum chemistry simulation on a quantum computer. A natural way to address this problem is to directly evaluate the error over a mesh in position and fit the data to a power law. This process can be made efficient using the SWAP test, as proposed by Wiebe et al [36]. A major drawback of this approach is that it requires roughly twice the qubits that the basic simulation used and also the variance in the estimate returned by the SWAP test can be prohibitively large. In this section, we propose an alternative method that estimates the error in the ground state energy by evaluating the error operator on a classical ansatz for the ground state numerically. This method also allows the contribution to the error in the quantum simulation from the Trotter error to be subtracted off of the final estimate, improving the accuracy of the simulation without requiring additional quantum operations.

Perhaps the most well-known classical algorithm for solving the electronic structure problem is a mean-field approach known as the Hartree-Fock method [21]. In this scheme, single particle molecular orbitals are obtained using a self-consistent variational procedure in which each particle is made to interact with the average density of the other particles. The output of this calculation provides molecular orbitals which, together with a spin assignment, are used to approximate the n-particle wavefunction as an anti-symmetric product of the orbitals (known by chemists as a Slater determinant).

Unfortunately, the Hartree-Fock method is incapable of approximating dynamic electron correlation and is known to overestimate energies by an amount that is typically well above the threshold of chemical accuracy. To correct for this problem, one can expand the wavefunction in a basis of multiple Slater determinants and variationally solve for the coefficients which minimize the electronic energy. In general, there are $M = \binom{N}{n}$ valid configurations for n electrons arranged into N spin or-

bitals. The ground state wavefunction in Eq. (1) may be represented as a linear combinations of these arrangements,

$$|\Psi\rangle = \sum_{i=1}^{M} a_i |i\rangle. \tag{31}$$

The energies may be solved for variationally,

$$E = \min_{\{a_i\}} \frac{\langle\Psi| H |\Psi\rangle}{\langle\Psi|\Psi\rangle} \quad \rightarrow \quad H^{\text{CI}} |\Psi\rangle = E |\Psi\rangle \tag{32}$$

where $H^{\text{CI}}_{ij} = \langle i| H |j\rangle$. In chemistry this method is known as full configuration interaction (FCI).

FCI is strongly believed to be classically intractable because M scales combinatorially with N and n. Accordingly, a common classical approach is to truncate the expansion in Eq. (31) to include only configurations that represent a fixed number of excitations from a reference configuration. Though this work and recent work [13] discuss using different orbital basis choices, usually the reference is taken to be the Hartree-Fock state (this orbital basis is known in chemistry literature as the "canonical basis"). This approach defines a hierarchy of methods referred to as truncated configuration interaction (CI) which approach exactness as the number of excitations is increased to the FCI space spanned by $N - n$ excitations. Fixing the maximum number of excitations at k, combinatorics suggests that the number of basis functions in truncated CI scales as $\Theta\left(\binom{N-n}{k}\binom{n}{k}\right)$. Truncation to the level of single and double excitations is referred to as configuration interaction singles, doubles (CISD) and is used for several purposes in this paper. Finally, we note that the accuracy of truncated CI is extremely sensitive to the quality of the reference state and it is therefore difficult to determine when these methods are expected to approximate the ground state energy within even a fixed multiplicative error.

Since the error operator can be efficiently computed and normal-ordered in second quantized form, we suggest evaluating the expectation value of this operator on a classical ansatz for the ground state. In particular, we focus on the use of the configuration interaction ansatz. Figure 9 illustrates the utility of this idea by showing the discrepancy between actual error and the error from evaluation of the error operator using a classical ansatz. Figure 10 shows the extent to which the effective error is reduced using a classical ansatz.

Apart from estimating errors, CISD states may also be of use in coalescing schemes [12] which use the Hartree-Fock approximation to determine whether a term in the Hamiltonian can be executed less frequently without significantly impacting the quality of the simulation. This process can substantially reduce the costs of simulating molecules with many small, but non–negligible, h_{pqrs} terms but may fail if the Hartree-Fock approximation breaks down. In such cases, the use of CISD states may
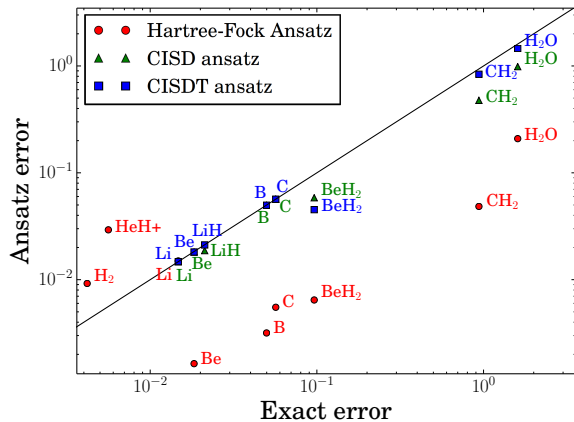
FIG. 9. Magnitude of the Trotter error in the exact ground state against the magnitude of the error induced on a classical ansatz for the ground state. Truncated CI computations are only performed when inexact; e.g., we have not computed HF using CISD because the calculation is exact in STO-6G.
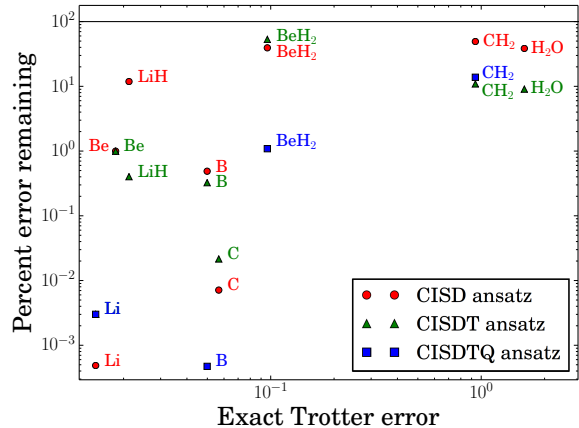

FIG. 10. Magnitude of the Trotter error induced in the exact ground state against percentage of the error that remains after subtracting the ansatz error from the exact error. This plot is intended to indicate the reduction in effective error when using a classical ansatz estimate. A black line is drawn at one-hundred percent remaining. In all benchmarks, using these ansatzs reduces effective error. Note that the quadruple calculation is so accurate for Be and LiH that the effective error appears to be exact to within double precision.

lead to superior coalescing schemes at the price of requiring more classical computing time to find the coalescing schedule.

Though the Hartree-Fock ansatz is usually not accurate enough to reduce error by an order of magnitude, the use of a truncated CI ansatz often exhibits enough accuracy to very substantially reduce effective error. While we focus on the CI ansatz to provide proof-of-principle, we believe that more intelligent truncation schemes can substantially increase ansatz accuracy without additional computational cost. For instance, the use of multi-reference methods has been shown to greatly improve the quality of the classical solution in many cases, especially near molecular dissociation limits where the exact electronic states become nearly degenerate [21].

The idea of using a classical ansatz to reduce the effective error in a quantum calculation is useful for two reasons. The first reason is that the error in a quantum simulation can usually be reduced by approximating the error with a classical ansatz at the CISD level of theory or greater, as demonstrated in Figure 10. The second (and perhaps more important) reason this technique is useful is that it gives a realistic a priori estimate of the error to expect in the quantum simulation (expected to be correct to at least an order of magnitude) which provides a methodology for selecting the number of Trotter steps required to obtain a desired precision. Finally, we point out that while the error operator might be computationally costly to compute (albeit, efficient in the polynomial-time scaling sense), Monte Carlo methods could be used to tractably sample the error operator expectation values with a classical ansatz.

A. Circuit for state preparation based on CI ansatz

In contrast to the Hartree-Fock states, CISD states are not computational basis states. Instead they are a linear combination of quantum states that are formed by single and double excitations away from a reference state which is often taken to be the Hartree-Fock state. Although the CISD state can be efficiently computed for a given electronic structure problem, preparing the state on a quantum computer is non-trivial. Here we present a method based on state-of-the-art multi-qubit synthesis methods to prepare the CISD state. Previous work has considered preparing this state using single qubit rotations and CNOT gates [37–39]. Such gate sets are unrealistic for fault tolerant quantum computing so we discuss the problem of compiling the state preparation circuit into Clifford and T gates. In the following analysis we will take the cost of the circuit to be given by the number of T gates because these gates are the most expensive gates to implement fault tolerantly in error correcting codes such as the surface code.

Let us begin by assuming the initial state for the quantum simulation (i.e. the state we wish to prepare) is of the form

$$|\psi\rangle = \sum_{k=1}^{D} \alpha_k |j_k\rangle, \qquad (33)$$

where j_k is a sequence of computational basis vectors that spans the space that state has support over and D is the dimension of that space.

It is unrealistic to assume that the state $|\psi\rangle$ will be exactly preparable using gates from the Clifford $+ T$ gate

library. Instead, the initial state will typically have to be approximated using these circuit elements. For years the Solovay-Kitaev algorithm provided the best known method for solving this approximation problem, but recently more advanced methods based on number theoretic results have provided much more efficient ways of performing this decomposition [40–42].

Therefore the problem of finding the best sequence of Clifford and T gates to approximate a multi-qubit unitary reduces to the following problem

1. Find integers x_0, x_1, y_0, y_1 such that

$$U_{p,q} \approx \tilde{U}_{p,q} = \frac{x_0 + x_1\sqrt{2} + iy_0 + iy_1\sqrt{2}}{\sqrt{2}^m}.$$

and $\tilde{U}$ is a unitary that can be exactly synthesized using elements from the gate library.

2. Find a sequence of Clifford and T gates that exactly implements $\tilde{U}$.

Note that because we are interested in preparing a state, not implementing a multi-qubit unitary, only the first column of U needs to be approximated. In particular, the first column of $\tilde{U}$ should approximate $|\psi\rangle$ to within a fixed error tolerance δ.

Before proceeding it is necessary to briefly review number theoretic approaches to multi-qubit circuit synthesis using Clifford and T gates. The key insight behind this strategy is that the unitary matrices that can be prepared with such circuits take on a very special form. The form can easily be seen from the Hadamard and T gates,

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad T = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1+i}{\sqrt{2}} \end{bmatrix}. \qquad (34)$$

It is then clear that any unitary matrix formed by a sequence of H and T gates will consist of matrix elements that are of the form

$$\tilde{U}_{i,j} = \frac{x_0 + x_1\sqrt{2} + iy_0 + iy_1\sqrt{2}}{\sqrt{2}^m}, \qquad (35)$$

for integer x_0, x_1, y_0, y_1. Since the remainder of the gate set consists of CNOT gates and Pauli gates which have (complex) integer valued matrix elements, it is then clear that every unitary that can be formed by the gate library also has matrix elements whose denominators are powers of $\sqrt{2}$ and whose numerators are in the ring of Gaussian integers $\mathbb{Z}[1/\sqrt{2}, i]$.

Just like ordinary fractions, these fractions also can be reduced. This notion of reducing a fraction manifests itself as the least denominator exponent k. In order to understand this concept concretely, it is necessary to introduce some terminology. Let $\omega = e^{i\pi/4}$ and

$$\mathbb{Z}[\omega] = \{a\omega^3 + b\omega^2 + c\omega + d | a, b, c, d \in \mathbb{Z}\}. \qquad (36)$$

Similarly, if we let $\mathbb{D} = \{a2^{-b} | a, b \in \mathbb{Z}\}$ denote the ring of dyadic fractions then we can express the ring $\mathbb{Z}[1/\sqrt{2}, i]$ as

$$\mathbb{D}[\omega] = \{a\omega^3 + b\omega^2 + c\omega + d | a, b, c, d \in \mathbb{D}\}. \qquad (37)$$

Then for every $t \in \mathbb{D}[\omega]$ there is a notion of a *least denominator exponent* that describes the fraction in Eq. (35) and uses the smallest value of m possible while requiring that x_0, x_1, y_0, y_1 are integer. Or more formally, the least denominator exponent, k, is the smallest non-negative integer such that $t\sqrt{2}^k \in \mathbb{Z}[\omega]$.

The smallest denominator exponent measures the precision in the approximation $U \approx \tilde{U}$ because Eq. (35) allows arbitrary complex numbers to be represented with zero error in the limit as $k \to \infty$. This means that the value of k used in the rounding process of the first column of U is a key property for characterizing the complexity of the state preparation. In fact, the problem of bounding the error in this approximation problem as a function of k has already been solved by Kliuchnikov [43]:

$$\|(U - \tilde{U})|0\rangle\| \le 2(D+2)2^{-4k} + 2\sqrt{2(D+2)}2^{-2k}, \qquad (38)$$

where D is the number of nonzero components of the state $|\psi\rangle = U|0\rangle$. As a technical point, the dimension of $\tilde{U}$ is at most $D+2$ rather than D because the first column of U must have *at least two zero-valued components* in order to guarantee that a solution exists to the Diophantine equation for $\tilde{U}$. This requires enlarging the Hilbert space dimension by two in the worst case scenario, which may require adding at most an additional qubit. However, the CISD state vector will likely have many zero valued components so this extra qubit will often not be needed in practice.

Using Eq. (38) we see that the state preparation error can be made less than δ by choosing

$$k = \left\lceil \frac{1}{4} \left[1 + \log_2 \left(\frac{D+2}{(\sqrt{1+\delta}-1)^2} \right) \right] \right\rceil. \qquad (39)$$

This means that if D is polynomial in n then $k \in O(\log(n/\delta))$.

Once the unitary $\tilde{U}$ has been found then the task of decomposing the unitary into fundamental operations remains a non-trivial problem. This problem is addressed by Giles and Selinger in [44]. The idea behind this approach is to decompose $\tilde{U}$ into a series of two level unitary operations. These two level unitary operations are then implemented using a Clifford circuit and a series of controlled operations to map each two level subspace to a single qubit. This process involves first identifying pairs of levels that can be simplified and then performing circuits of the form $H^w T^x H^y T^z$ to the two level subspace such that the denominator exponent is systematically reduced. Once the least denominator exponent is reduced to 0 then the subspace either takes the form $[\omega^p, 0]^T$ or $[0, \omega^p]^T$ for integer p. Thus, the inverse of the state preparation circuit can be found (up to a global phase) by

performing this reduction process iteratively of the D dimensional initial state until only one nonzero component remains and then mapping this component to $|0\rangle$ using a Clifford circuit and a multiply controlled not gate.

At most k reduction steps are needed to reduce each two level subspace and there are at most $(D+2)-1$ subspaces that must be looped through. Therefore, there are at most $k(D+1)$ reduction steps taken. Each reduction step consists of applying at most two H gates and two T^x gates to each subspace, as well as a multiply controlled not gate to map the final state to one proportional to $|0\rangle$. Hence, in order to assess the cost of the algorithm we need to compute the costs of each of these gates.

Let us imagine that we need to perform a gate on the subspace span$(|j\rangle , |k\rangle)$. We want to map this to span$(|2^n - 1\rangle , |2^n - 2\rangle)$ so that the gate can be applied to the last qubit. By performing a sequence of $\mathcal{O}(n)$ X gates, we can map

$$\text{span}(|j\rangle , |k\rangle) \to \text{span}(|j \oplus k \oplus 2^n - 1\rangle , |2^n - 1\rangle),$$

where $\oplus$ is bitwise exclusive or. There are two cases that we need to consider. If $j \oplus k = 1 \mod 2$ then the least significant bit of $j \oplus k \oplus 2^n - 1$ is 0. This means that the state $|j \oplus k \oplus 2^n - 1\rangle$ can be mapped to $|2^n - 2\rangle$ using a sequence of $n - 1$ zero-controlled not gates while not affecting $|2^n - 1\rangle$. Otherwise, if $j \oplus k = 1 \mod 2$ then we can reduce this case by finding the least significant bit where j and k differ and swap that bit with the least significant bit. Since $|2^n - 1\rangle$ is an eigenstate of the swap operator, the swap does not affect that vector. Hence in either case we can perform the subspace mapping using $\mathcal{O}(n)$ Clifford operations.

In order to apply the H and T gates required by the synthesis algorithm on the correct qubits, we need to implement controlled variants of these circuits. There are many constructions for these controlled gates [45–47]. Here we anticipate that the cost of state preparation for the CISD state will be sub-dominant to the cost of the simulation. This means that minimizing the number of qubits needed is an important design goal. Let us define $\Lambda_m(G)$ to be the m-controlled version of the gate G. Then the gate $\Lambda_m(H)$ can be implemented using two $\Lambda_{n-1}(X)$ gates, a $\Lambda_1(H)$ gate and an ancilla qubit,
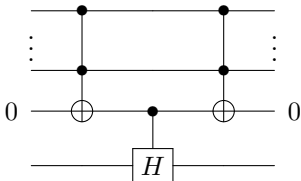
Controlled T^q gates can be performed similarly,
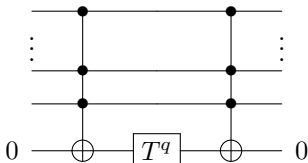
The resulting circuits can be further optimized by noting that many of the Toffoli gates needed to perform the reductions of the least denominator exponent are redundant. In particular we can express the simplified reduction circuit as,
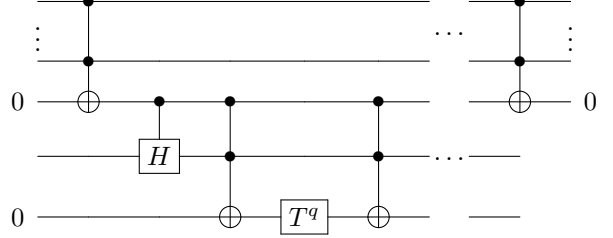
The gate $\Lambda_1(H)$ requires 2 T-gates [44], and the Toffoli gates can be implemented, up to an irrelevant phase, using 4 T-gates [46, 47] and an ancilla qubit. The entire process requires at most $N + 4$ qubits, which is typically less memory than is required for the quantum simulation and eigenvalue estimation phases of the algorithm. This means that the additional four qubits required for the state preparation algorithm will not impact the memory requirements of the overall simulation algorithm.

For the present problem, the CISD state is in $\mathbb{C}^{2^{N+1}}$ (recall one additional qubit is needed to ensure a solution to the norm equations for synthesis). This means that we also need to consider the cost of implementing $\Lambda_N(X)$ gates. Although highly time-efficient constructions for the multiply controlled circuits can be made using the circuits of [48], they require a large number of qubits. In order to ensure that the space complexity of state preparation does not dominate the algorithm, we use the less time-efficient construction of Barenco et al [49] to compile the $\Lambda_N(X)$ gates. Using Corollary 7.4 from [45] and the $\Lambda_2(iX)$ gate from [46, 47] to implement the Toffoli gate, the cost of implementing such circuits is at most

$$T_{\text{count}}\big(\Lambda_N(iX)\big) \le 32(N - 3). \tag{40}$$

At most $N + 4$ qubits, where $N \ge 5$, are needed to implement these gates [45].

The reduction of each of the two dimensional subspaces requires two steps. First, the application of the $\Lambda_N(X)$ gates to mark the subspace and a sequence of k controlled operations to reduce the denominator exponent of that subspace. This reduction process requires k steps, each of which involves at most two $\Lambda_1(H)$ and two $\Lambda_2(T^q)$. Including the cost of the two $\Lambda_N(X)$ gates, the total cost of the reduction is at most $22k + 64(N - 3)$ T-gates. At most $D+1$ reduction steps are required in this process as well as potentially a swap of the final state into the $|0^{N+1}\rangle$ state (which can be performed using Clifford operations). Thus the overall T–count for this process is

$$(22k + 64(N - 3))(D + 1). \tag{41}$$

This also is the T–count for for preparing the CISD state from the $|0^{N+1}\rangle$ state because the necessary circuit can

be found by taking the Hermitian conjugate of the resultant gate sequence. Thus using Eq. (39) the total number of non-Clifford operations required in the state preparation scales at most as

$$\mathcal{O}\left(D \log\left(\frac{D}{\delta}\right) + ND\right) \qquad (42)$$

If the approach of Wang et al [38], coupled with recent methods for decomposing single qubit rotations into T gates, is used to prepare the CISD state then the resultant T-count scales at most as $\tilde{\mathcal{O}}(2^{n_e} N^{n_e}/n_e!) \log(1/\delta)$. If $n_e \approx N/2$ then this method is inefficient, whereas ours is not since $D \in \mathcal{O}(N^4)$ for CISD states. If $n_e \leq 3$ then the method of Wang et al does provide superior scaling as N increases, though cases where $n_e \leq 3$ and N is large may be rare. In contrast, the method of Ortiz et al [37] requires $\tilde{\mathcal{O}}(D^2 N^2 \log(1/\delta))$ gates, which is nearly quadratically slower than our method.

As a final point, the cost of the state preparation algorithm is $\mathcal{O}(N^5)$ in worst case scenarios. This can be comparable to, or greater than, the cost of quantum simulation in the limit of large N. This means that using a naïve CISD approximation in cases with half filling may seriously degrade the performance of the algorithm. This means that in order to see the performance advantages promised by recent algorithms, which have scaling near $\mathcal{O}(N^4)$, sophisticated state preparation methods are needed in cases where the Hartree-Fock state has poor overlap with the FCI ground state.

IV. CONCLUSION

Our work calls into question the basic assumption that the error in Trotter-Suzuki based methods for simulating quantum chemistry is *explicitly* a function of the number of spin orbitals used to represent the system. We find through numerical evidence that such errors do not seem to be directly related to the number of spin orbitals in the system for small molecules. We observe this lack of correlation for a variety of orbital bases including local, canonical and natural orbitals. Instead, we see that chemical features such as the maximum nuclear charge is a strong indicator of the complexity of a simulation. We argue that the errors should scale as $\mathcal{O}(Z_{\max}^6)$ for an atomic orbital basis, which is in close agreement with the scaling observed numerically. We also observe that some atoms, such as Oxygen, Fluorine and Neon, have vanishingly small Trotter errors despite available error bounds predicting large Trotter errors for these molecules. We show that this discrepancy can be understood as a consequence of the large filling fraction for these molecules. This suggests that chemical features of a molecule may be much better predictors of the number of Trotter steps needed in a simulation than the number of spin orbitals assigned to the molecule.

We further analyze the errors and see that the discrepancy between the observed Trotter error and the norm of the error operator does not arise from random cancellation. Indeed, the errors observed are much greater than what would be expected if the ground state were a Haar random state that was chosen independently from the eigenvectors of the error operator. Furthermore, we observe that the distribution of eigenvalues of the error operator is highly structured and has many near-zero eigenvalues, which likely is the cause of the orders of magnitude separation between the Trotter error and the norm of the error operator.

We also use the error operator to improve quantum simulation methods by providing a computationally efficient algorithm for estimating the error in a simulation. This leads to two applications: (a) compensating for Trotter error in a quantum simulation by subtracting the prediction off the result and (b) predicting the number of Trotter steps needed in a simulation. Finally, we provide a quantum algorithm for preparing CISD states that is polynomially more efficient than existing methods and may provide a viable alternative to adiabatic state preparation in cases where the Hartree-Fock approximation to the ground state leads to poor success probability.

There are several natural avenues of inquiry that this work reveals. First, although this work shows strong numerical evidence for small molecules we do not have sufficient evidence to state conclusively that the error in the Trotter-Suzuki formula is independent of N in the asymptotic limit. Larger numerical experiments may be needed to shed more light on the scaling of Trotter–Suzuki errors in this regime. Secondly, Ferredoxin is often suggested as a strong candidate for quantum chemistry simulation but Fe_2S_2 has large nuclear charges which make it a challenging molecule from the perspective of simulation. This suggests that there may be other large organic molecules with smaller nuclear charges that may be even more natural targets for quantum simulation. Finally, although our work has suggested that the number of spin orbitals in a molecule may not uniquely characterize the cost of a quantum chemistry simulation, it does not provide a simple criteria for determining which molecules are easy or hard to simulate. Finding molecular features, beyond the maximum nuclear charge and the filling fraction, that can be used to predict the relative difficulty of simulation would not only constitute an important step forward for quantum chemistry simulations but would also be an important contribution to quantum chemistry as a whole.

V. ACKNOWLEDGEMENTS

[1] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[2] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 4 (1997).
[3] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science (80-.). **309**, 20 (2006).
[4] G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **64**, 022319 (2001).
[5] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann, Phys. Chem. Chem. Phys. **10**, 5388 (2008).
[6] H. Wang, S. Ashhab, and F. Nori, Phys. Rev. A - At. Mol. Opt. Phys. **79** (2009).
[7] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Mol. Phys. **2**, 106 (2010), arXiv:1001.3855.
[8] N. Cody Jones, J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. V. Meter, A. Aspuru-Guzik, and Y. Yamamoto, New J. Phys. **14**, 115023(1 (2012).
[9] J. T. Seeley, M. J. Richard, and P. J. Love, J. Chem. Phys. **137** (2012), 10.1063/1.4768229.
[10] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Phys. Rev. A **90**, 022305 (2014).
[11] M. Hastings, D. Wecker, B. Bauer, and M. Troyer, e-print arXiv:1403.1539 (2014).
[12] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer, arXiv e-print: 1406.4920 , 13 (2014), arXiv:1406.4920.
[13] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, e-print arXiv: 1407.7863 (2014).
[14] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. OBrien, Nat. Commun. **5** (2014), 10.1038/ncomms5213.
[15] R. Babbush, P. J. Love, and A. Aspuru-Guzik, Sci. Rep. **4**, 6603 (2014).
[16] B. Toloui and P. J. Love, e-print arXiv:1312.2579 (2013).
[17] B. P. Lanyon, J. D. Whitfield, G. G. Gillet, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Nat. Chem. **2**, 20 (2009).
[18] J. Du, N. Xu, X. Peng, P. Wang, S. Wu, and D. Lu, Phys. Rev. Lett. **104**, 030502 (2010).
[19] D. Lu, N. Xu, R. Xu, H. Chen, J. Gong, X. Peng, and J. Du, Phys. Rev. Lett. **107**, 020501 (2011).
[20] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, and J. Wrachtrup, e-print arXiv: 1405.2696 (2014), arXiv:1405.2696.
[21] T. Helgaker, J. Olsen, and P. Jorgensen, *Phys. Today* (Wiley, 2013) p. 944.
[22] C. Zalka, Fortschritte der Phys. **46**, 877 (1998).
[23] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Proc. Natl. Acad. Sci. **105**, 18681 (2008).
[24] R. Somma, G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **65**, 17 (2002).
[25] S. Bravyi and A. Kitaev, Ann. Phys. (N. Y). **298**, 18 (2000).
[26] D. Aharonov and A. Ta-Shma, in *Proc. thirty-fifth ACM Symp. Theory Comput. - STOC '03* (ACM Press, New York, New York, USA, 2003) p. 20.
[27] A. M. Childs and N. Wiebe, J. Math. Phys. **54**, 062202 (2013).
[28] D. W. Berry and A. M. Childs, Quantum Inf. Comput. **12**, 29 (2012).
[29] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, arXiv:1312.1414 (2013).
[30] M. Schmidt, K. Baldridge, J. Boatz, S. Elbert, M. Gordon, J. Jensen, S. Koeski, N. Matsunaga, K. Nguyen, S. Su, T. Windus, M. Dupuis, and J. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[31] M. Gordon and M. Schmidt, in *Theory Appl. Comput. Chem. first forty years*, edited by C. Dykstra, G. Frenking, K. Kim, and G. Scuseria (Elsevier, Amersterdam, 2005) pp. 1167–1189.
[32] S. Raeisi, N. Wiebe, and B. C. Sanders, New Journal of Physics **14**, 103017 (2012).
[33] C. L. Benavides-Riveros, J. M. Garcia-Bondia, and M. Springborg, arxiv e-print: 1409.6435 (2014).
[34] D. M. Ceperley, J. Stat. Phys. **43**, 815 (1986).
[35] B. L. Hammond, P. J. Reynolds, and W. A. Lester, J. Chem. Phys. **87** (1987), 10.1063/1.453345.
[36] N. Wiebe, D. W. Berry, P. Høyer, and B. C. Sanders, Journal of Physics A: Mathematical and Theoretical **44**, 445308 (2011).
[37] G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme, Physical Review A **64**, 022319 (2001).
[38] H. Wang, S. Ashhab, and F. Nori, Physical Review A **79**, 042335 (2009).
[39] M. H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, Sci. Rep. **4**, 9 (2014), arXiv:1307.4326.
[40] V. Kliuchnikov, D. Maslov, and M. Mosca, Quantum Information & Computation **13**, 607 (2013).
[41] N. J. Ross and P. Selinger, arXiv preprint arXiv:1403.2975 (2014).
[42] A. Bocharov, M. Roetteler, and K. M. Svore, arXiv preprint arXiv:1409.3552 (2014).
[43] V. Kliuchnikov, arXiv preprint arXiv:1306.3200 (2013).
[44] B. Giles and P. Selinger, Physical Review A **87**, 032332 (2013).
[45] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Physical Review A **52**, 3457 (1995).
[46] C. Jones, Phys. Rev. A **87**, 022328 (2013).
[47] P. Selinger, Phys. Rev. A **87**, 042302 (2013).
[48] C. Jones, Physical Review A **87**, 022328 (2013).
[49] A. Barenco, C. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995).
[50] P. Hayden, D. W. Leung, and A. Winter, Communications in Mathematical Physics **265**, 95 (2006).

Appendix A: Computation of Haar Expectations

In order to determine whether the error cancellations observed for ground state quantum simulations arise because of properties of the eigenstates of the Hamiltonian, we need to determine whether these results would also be typical of random vectors. Here we provide a derivation, for completeness, of the Haar expectation value and variance of the $|\langle v|k\rangle|^2$.

In the following we will take k to be fixed and v to represent the Haar random variable. We will also use the convention that $\mathbb{E}_H$ denotes the expectation value of a quantity over a set of Haar random vectors, and $\mathbb{V}_H$ denotes the variance over the set. To be clear, $\mathbb{E}_H|\langle v|k\rangle|^2 = \int_{U \in \text{Haar}} |\langle 0| U^\dagger |k\rangle|^2 \mathrm{d}U$.

We wish to compute the variance,

$$\mathbb{V}_H(|\langle v|k\rangle|^2) = \mathbb{E}_H(|\langle v|k\rangle|^4) - \mathbb{E}_H(|\langle v|k\rangle|^2)^2, \quad (A1)$$

of the square of the overlap of an arbitrary Haar random vector, $|0\rangle$, with an eigenvector of an arbitrary Hermitian operator (in this case, the Trotter error operator), $|k\rangle$. We begin by stating the correspondence,

$$|v\rangle \langle v| = U |0\rangle \langle 0| U^\dagger \quad (A2)$$

where U is the unitary Gram matrix which affects a basis transformation into the error operator eigenbasis (for instance), and $|v\rangle$ represents $|0\rangle$ in the error operator eigenbasis. We are interested in the projection of this state onto an eigenvector of the error operator,

$$a_k = \langle k | v \rangle = \langle k| U |0\rangle \quad (A3)$$

$$|a_k|^2 = \langle k| U |0\rangle \langle 0| U^\dagger |k\rangle \quad (A4)$$

$$= \text{tr} \left[|k\rangle \langle k| U |0\rangle \langle 0| U^\dagger \right].$$

We compute this trace in two steps. From the unitary invariance of the Haar measure we have that

$$\int_{U(n)} \left[U |0\rangle \langle 0| U^\dagger \right] \mathrm{d}U = \frac{\mathbb{1}}{2^N} \quad (A5)$$

Therefore

$$\mathbb{E}_H(|a_k|^2) = \text{tr} \left[|k\rangle\langle k| \frac{\mathbb{1}}{2^n} \right] = \frac{1}{2^N}. \quad (A6)$$

Thus, $\mathbb{E}_H(|a_k|^2) = \frac{1}{2^N}$, and hence $\mathbb{E}_H(|a_k|^2)^2 = \frac{1}{2^{2N}}$.

Focusing on the remaining component of the variance,

$$|a_k|^4 = \langle k| U |0\rangle \langle 0| U^\dagger |k\rangle \langle k| U |0\rangle \langle 0| U^\dagger |k\rangle \quad (A7)$$

$$= \text{tr} \left[(|k\rangle \langle k|)^{\otimes 2} U^{\otimes 2} (|0\rangle \langle 0|)^{\otimes 2} U^{\dagger \otimes 2} \right].$$

To further evaluate the trace we follow the treatment in [50] which uses the spectral theorem to derive orthogonal projectors onto symmetric and antisymmetric subspaces. This begins by defining a flip operator, $\mathbb{F} \in \mathbb{C}^{2^{2N} \times 2^{2N}}$,

$$\mathbb{F} \left(|\psi\rangle \otimes |\varphi\rangle \right) = |\varphi\rangle \otimes |\psi\rangle. \quad (A8)$$

From this definition it is clear that

$$\mathbb{F} = \pi_{\text{sym}} - \pi_{\text{antisym}} \quad (A9)$$

$$\mathbb{1}^{\otimes 2} = \pi_{\text{sym}} + \pi_{\text{antisym}}. \quad (A10)$$

Thus,

$$\pi_{\text{sym}} = \frac{1}{2} \left(\mathbb{1}^{\otimes 2} + \mathbb{F} \right) \quad (A11)$$

$$\pi_{\text{antisym}} = \frac{1}{2} \left(\mathbb{1}^{\otimes 2} - \mathbb{F} \right). \quad (A12)$$

Since $\text{tr} \left[\mathbb{1}^{\otimes 2} \right] = 2^{2N}$ and $\text{tr}(\mathbb{F}) = 2^N$,

$$\text{tr} \left[\pi_{\text{sym}} \right] = \frac{2^N \left(2^N + 1 \right)}{2} \quad (A13)$$

$$\text{tr} \left[\pi_{\text{sym}} \right] = \frac{2^N \left(2^N - 1 \right)}{2}. \quad (A14)$$

Since $|0\rangle \langle 0| \otimes |0\rangle \langle 0|$ is entirely symmetric, it is straight forward to see from unitary invariance that

$$\int_{U(n)} \left[U^{\otimes 2} (|0\rangle \langle 0|)^{\otimes 2} U^{\dagger \otimes 2} \right] \mathrm{d}U = \frac{2}{2^N \left(2^N + 1 \right)} \pi_{\text{sym}}. \quad (A15)$$

Thus,

$$\mathbb{E}_H(|a_k|^4) = \frac{2}{2^N \left(2^N + 1 \right)} \text{tr} \left[\pi_{\text{sym}} (|k\rangle \langle k|)^{\otimes 2} \right]$$

$$= \frac{2}{2^N \left(2^N + 1 \right)}. \quad (A16)$$

Finally, we arrive at the variance of $|a_k|^2$,

$$\mathbb{V}_H(|a_k|^2) = \frac{2}{2^N \left(2^N + 1 \right)} - \frac{1}{2^{2N}}. \quad (A17)$$

This gives us the variance in the $|a_k|^2$ terms, which in turn allows us to find the deviation from the expected error for a quantum chemistry simulation. The key point here is that the standard deviation is on the order of the expectation value which means that we expect relatively large fluctuations in the probabilities that correspond to particular eigenvalues of the Trotter error operator. Hence we do not expect a concentration of measure result to hold in high dimensional spaces.

Appendix B: Contributions of orbitals to Trotter error operator

(a) Water, local basis, Hamiltonian coefficients.

(b) Water, local basis, error coefficients.

(c) Water, local basis, error contributions.

(d) Water, natural basis, Hamiltonian coefficients.

(e) Water, natural basis, error coefficients.

(f) Water, natural basis, error contributions.

(g) Beryllium hydride, local basis, Hamiltonian coefficients.

(h) Beryllium hydride, local basis, error coefficients.

(i) Beryllium hydride, local basis, error contributions.

(j) Beryllium hydride, natural basis, Hamiltonian coefficients.

(k) Beryllium hydride, natural basis, error coefficients.

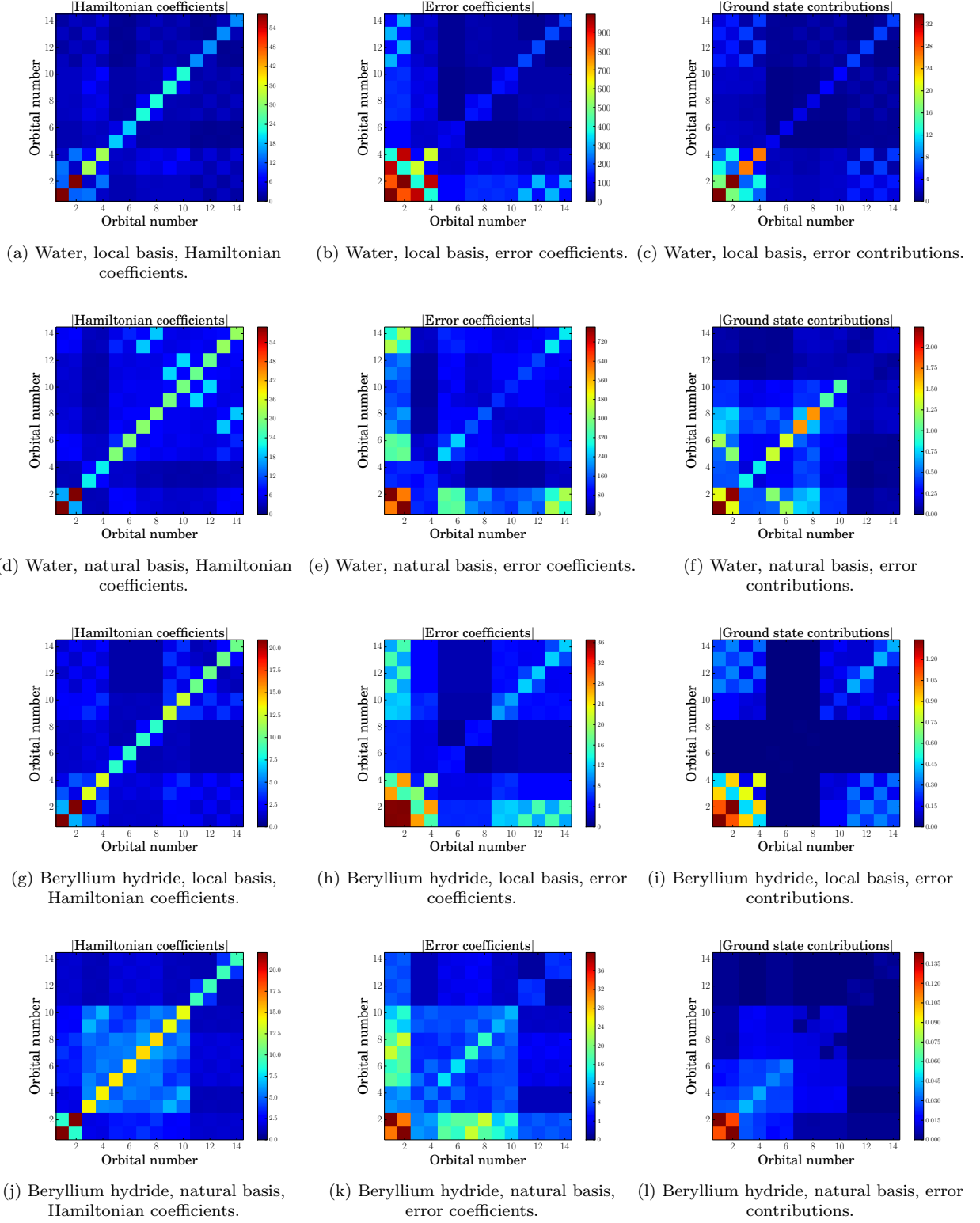(l) Beryllium hydride, natural basis, error contributions.

FIG. 11. These plots show the coefficients of normal-ordered terms in the Hamiltonian and error operator as well as expectation values of the error operator terms for the ground state. The terms are binned according to the orbitals involved in the term. This plot shows the marginal distribution of the magnitudes of these terms.