Improving Quantum Algorithms for Quantum Chemistry

M. B. Hastings,[1,2] Dave Wecker,[2] Bela Bauer,[1] and Matthias Troyer[3]

[1] *Station Q, Microsoft Research, Santa Barbara, CA 93106-6105, USA*
[2] *Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052, USA*
[3] *Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland*

We present several improvements to the standard Trotter-Suzuki based algorithms used in the simulation of quantum chemistry on a quantum computer. First, we modify how Jordan-Wigner transformations are implemented to reduce their cost from linear or logarithmic in the number of orbitals to a constant. Our modification does not require additional ancilla qubits. Then, we demonstrate how many operations can be parallelized, leading to a further linear decrease in the parallel depth of the circuit, at the cost of a small constant factor increase in number of qubits required. Thirdly, we modify the term order in the Trotter-Suzuki decomposition, significantly reducing the error at given Trotter-Suzuki timestep. A final improvement modifies the Hamiltonian to reduce errors introduced by the non-zero Trotter-Suzuki timestep. All of these techniques are validated using numerical simulation and detailed gate counts are given for realistic molecules.

One of the most natural applications of a quantum computer is to simulate quantum mechanics, as suggested by Feynman[1]. There has been much work on constructing quantum algorithms to simulate various problems in quantum mechanics, ranging from problems in condensed matter physics such as the Hubbard model to quantum field theory[2] and quantum chemistry[3–16].

A recent large-scale study of quantum chemistry on a quantum computer[17] gave accurate gate counts for some of the standard circuits in the literature when applied to problems in quantum chemistry. Unfortunately, this study found that even modest molecules require enormously long simulation time. The reason for this is simple: in quantum chemistry, we consider a basis with N spin orbitals. The Hamiltonian considered takes the form

$$H = \sum_{pq} t_{pq} c_p^\dagger c_q + \frac{1}{2} \sum_{pqrs} V_{pqrs} c_p^\dagger c_q^\dagger c_r c_s \tag{1}$$

for a system of interacting electrons. Many of the two-body interaction terms V_{pqrs} are non-zero and so there are roughly N^4 separate, non-commuting terms in the Hamiltonian. The time requirements become large even for roughly 100 spin orbitals; since roughly 50-70 spin orbitals can already be simulated on a classical computer using exact or approximate techniques[18–22], clearly improvements are needed in the quantum algorithms to make them potentially useful.

Typical quantum algorithms to simulate this system need to implement unitary time evolution under the Hamiltonian (1). Efficient ways to implement this evolution have thus been the object of intense research efforts[23–30]. Other approaches that do not rely on time evolution have been proposed[31,32]. For the time evolution, a Trotter-Suzuki approach[33,34] is most common, where the evolution $\exp(iH\delta_t)$ for a small time step δ_t (controlled by an additional ancilla qubit used to perform the phase estimation) is implemented through a sequence of $O(N^4)$ unitary transformations $\exp(iA\delta_t)$ where A is some term in Eq. (1). Standard implementations[10] have an additional factor of N overhead in gate count to implement the Jordan-Wigner transformation, which encodes the anticommutation relations of the fermionic degrees of freedom, giving a time complexity $O(N^5)$. Further, as N increases, the Trotter time step must become smaller to obtain a fixed, given accuracy, further worsening the performance of the algorithm. All these effects combine to give poor scaling with N.

In this paper, we significantly alleviate these problems. One technique is a modified circuit which reduces the gate count overhead for Jordan-Wigner strings to a *constant* without requiring additional qubits. This improves both on the linear overhead of Ref. 10 and on the Bravyi-Kitaev scheme which has only a logarithmic overhead[35]. Further, we show that with this modified circuit, many of the operations can be parallelized, leaving the total gate count unchanged but reducing the parallel circuit depth. All these improvements significantly reduce the gate count and parallel circuit depth both asymptotically and for small molecules. Further, we modify the Trotter-Suzuki decomposition in two ways. First, we modify the term order to take into account special properties of a Hartree-Fock basis in quantum chemistry. Second, we modify the Hamiltonian that we study in a way that corrects for having a nonzero Trotter step. These improvements allow us to obtain accurate results at much larger Trotter step than one might obtain otherwise.

These improvements then take two forms. One type of improvement involves modifying the circuits to perform the same simulation in a more effective way. The other type involves modifying the simulation done to obtain more accurate answers at larger Trotter step. Below, we detail all these improvements. We explicitly check gate counts for various real molecules, using LIQ$Ui|\rangle$, a quantum simulator developed at Microsoft Research[36], so that we could obtain precise numbers for the improvements over the original circuits, rather than just asymptotic estimates.

To start, let us fix some terminology. We say that a term in Eq. (1) is an H_{pq} term if it is of the form $t_{pq}c_p^\dagger c_q$ for $p \neq q$, while it is an H_{pp} term if $p = q$. A term is an H_{pqrs} term if it is of the form $\frac{1}{2}V_{pqrs}c_p^\dagger c_q^\dagger c_r c_s$ for p, q, r, s all distinct. An H_{pqqp} term is of the form $\frac{1}{2}V_{pqqp}c_p^\dagger c_q^\dagger c_q c_p$ while an H_{pqqr} term is of the form $\frac{1}{2}V_{pqqr}c_p^\dagger c_q^\dagger c_q c_r$, again for p, q, r all distinct. Note that the H_{pp} and H_{pqqp} terms are diagonal. Since $c_p c_q = -c_q c_p$ and similarly $c_p^\dagger c_q^\dagger = -c_q^\dagger c_p^\dagger$, we will rewrite any term $\frac{1}{2}V_{pqpq}c_p^\dagger c_q^\dagger c_p c_q$ as an H_{pqqp} term and we will rewrite any term $\frac{1}{2}V_{pqrq}c_p^\dagger c_q^\dagger c_r c_q$ as an H_{pqqr} term, while for an an H_{pqrs} term we will assume that that $p < q$ and $r < s$.

For the molecules we studied, the terms with the largest magnitude are the H_{pp} terms. The H_{pqqp} terms are typically next in magnitude, then the H_{pq} terms, next the H_{pqqr} terms, and finally the H_{pqrs} terms. This ordering is not strict. Further, there is an important sum rule relating the H_{pqqr} and H_{pq} terms. After a basis transformation of the single-particle states to find a Hartree-Fock ground state, one has

$$t_{pq} + \sum_r V_{prrq}n_r = 0, \tag{2}$$

where $n_r = 0, 1$ is the occupation number in the Hartree-Fock state.

1. CANCELLING JORDAN-WIGNER STRINGS

We work in a second quantized basis, using one qubit per spin orbital. In the case of an H_{pq} term with $p < q$, the desired controlled unitary is

$$\exp(i(\sigma_p^X \sigma_q^X + \sigma_p^Y \sigma_q^Y)(\sigma_{p+1}^Z \sigma_{p+2}^Z ... \sigma_{q-1}^Z)\theta), \tag{3}$$

where the product of Zs implements the Jordan-Wigner string and where θ depends upon the coefficient t_{pq} and on δ_t. In the case of an H_{pqrs} term with $p < q < r < s$, there are several possible controlled unitaries, of the form

$$\exp(i(\sigma_p^X \sigma_q^X \sigma_r^X \sigma_s^X)(\sigma_{p+1}^Z ... \sigma_{q-1}^Z)(\sigma_{r+1}^Z ... \sigma_{s-1}^Z)\theta) \tag{4}$$

with Jordan-Wigner strings from $p+1$ to $q-1$ and from $r+1$ to $s-1$. In fact, for each p, q, r, s, it may be necessary to implement several of these terms, with $\sigma_p^X \sigma_q^X \sigma_r^X \sigma_s^X$ replaced by other Pauli matrices, such as $\sigma_p^X \sigma_q^X \sigma_r^Y \sigma_s^Y$, etc..., in every case having an even number of σ^X operators and an even number of σ^Y operators. We refer to these different choices of σ^X, σ^Y as different *subterms*.

Since the H_{pqrs} terms are the most numerous and thus dominate the computation time, we focus on speeding up the implementation of these terms; while the techniques can be straightforwardly used to also speed up H_{pq} or H_{prrq} terms, we prefer to execute those terms in a different order for reasons explained later. The standard circuit[10] to implement this is shown in Figure 1. The lines represent qubits, with the top line labelled "Phase" being used to control the application of the term. We show only a single subterm here; the circuit will be replicated several times with different subterms. Our notation in writing circuits is that two qubit gates with open circles denote CNOTs, with the open circle representing the target of the CNOT. Boxes containing an H represent Hadamard gates, which transform perform a basis change between the eigenbasis of σ^X and σ^Z operators and thus effective interchange these operators, i.e. $H\sigma^Z H = \sigma^X$. Y and $Y^\dagger$ perform the analogous operation between σ^Y and σ^Z operators, i.e. $Y^\dagger \sigma^Y Y = \sigma^Z$. We thus refer to these as basis changes, and they are given by

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad Y = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{5}$$

We notice immediately that the length of the Jordan-Wigner string may be proportional to N, leading to a linear overhead. However, as we now show, it is possible to replace this with a constant overhead. Consider the circuit of Figure 2. The Jordan-Wigner string now appears outside all the basis changes, giving a constant factor improvement since it appears only once per term, rather than once per subterm. The equivalence of the circuits Figure 1 and Figure 2 can be worked out after some linear algebra. A brief explanation is as follows. Consider the subterm consisting of $\sigma_p^X \sigma_q^X \sigma_r^X \sigma_s^X$. If the product of the Zs appearing in the Jordan-Wigner string is equal to $+1$, then the controlled-Z has no effect and the gate implements the transformation $\exp(i\sigma_p^X \sigma_q^X \sigma_r^X \sigma_s^X \theta)$. However, if the product is equal to -1, then the controlled-Z applies the operation σ_s^Z. This σ_s^Z anti-commutes with σ_s^X, so that instead the gate implements $\exp(-i\sigma_p^X \sigma_q^X \sigma_r^X \sigma_s^X \theta)$, agreeing with Eq. (4).

At this point, we have only achieved an improvement by a constant factor. However, a further improvement appears when we implement several of these terms in succession. For the Trotter-Suzuki decomposition, we have to implement
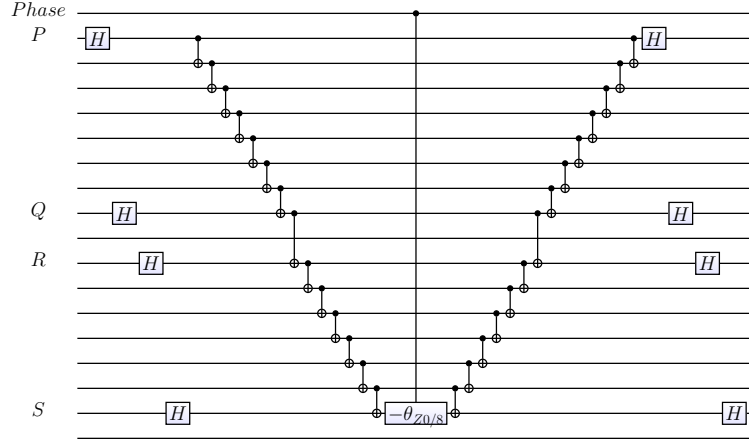
FIG. 1. Standard circuit for H_{pqrs}[10]. H represents Hadamard gate. Up to 8 copies of this circuit appear in succession, with Hadamard gates replaced by other basis change gates to execute different subterms.
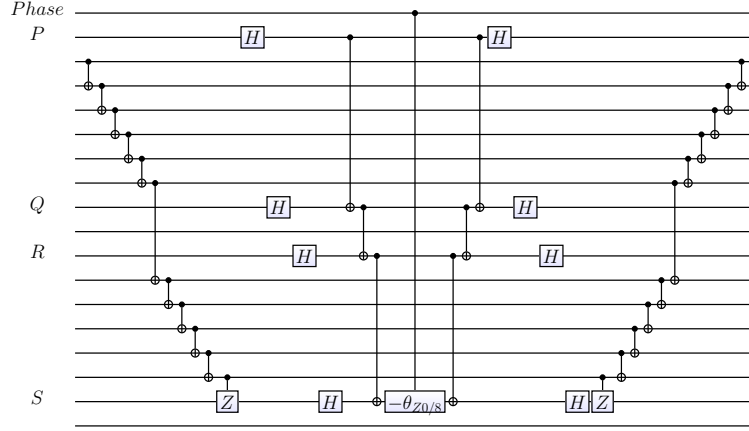

FIG. 2. Alternative circuit for H_{pqrs}. The ... indicate several different subterms. The two-qubit gate with a box labelled Z is a controlled-Z gate, executing σ^Z on the target if the first qubit is 1 and identity otherwise.

every single possible term H_{pqrs}. Let us order these terms lexicographically, meaning that we implement a given H_{pqrs} term, then increment s and implement the next H_{pqrs} term; when $s = N$, we then increment r and reduce s to its minimum value, and so on. In this case, the Jordan-Wigner strings largely cancel between successive terms; if we follow a term with given p, q, r, s with a term with $p, q, r, s + 1$, then all but $O(1)$ CNOTs can be cancelled between the end of one term and the start of the next, using the fact that the square of a CNOT gate is equal to the identity. If we follow the term with given p, q, r, s with a term other than $p, q, r, s + 1$, then more of the string of CNOTs may be left uncancelled; see Figures 3,4 for an illustration.

There is a string of uncancelled CNOTs when we increment r once $s = N$. However, since this occurs only every $O(N)$ steps, it leads to only subleading overhead. We can, however, cancel even these string by using the modified circuit shown in Figure 5 where we show several successive H_{pqrs} terms. Here we show a circuit that uses an additional ancilla (initialized to $|0\rangle$ at the start of the computation) to track the total parity of electrons on orbitals $p + 1, ..., q - 1, r + 1, ..., s - 1$. Then, it is possible to modify the lexicographic order by first increasing s to N, then incrementing r by 1 and then decrementing s from N, alternately incrementing and decrementing s for each choice of p, q, r. This particular modification of the lexicographic order is *not* shown; rather we show a particular ordering that works well for parallelization as discussed later.

A further advantage of the ancilla circuit appears already without parallelization, in that it allows additional term cancellations. Note that for many molecules studied, not all H_{pqrs} terms are non-zero. Hence, it will often not be the case that given a particular p, q, r, s that there is a non-zero term $p, q, r, s + 1$. This at first seems to limit the number of cancellations as we often have uncancelled strings. However, in studying parallelization, we found that the
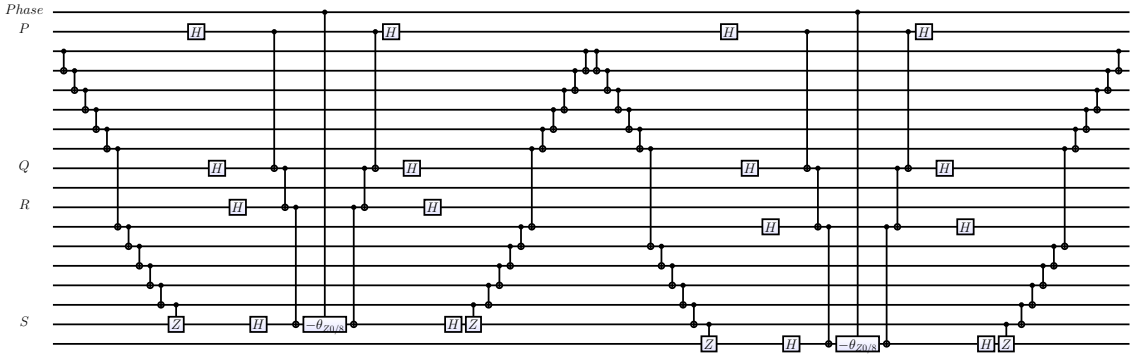
FIG. 3. Alternative circuit for H_{pqrs} shown for two successive choices of p, q, r, s. In this case, both r and s changed so that some of the CNOT cannot be cancelled.
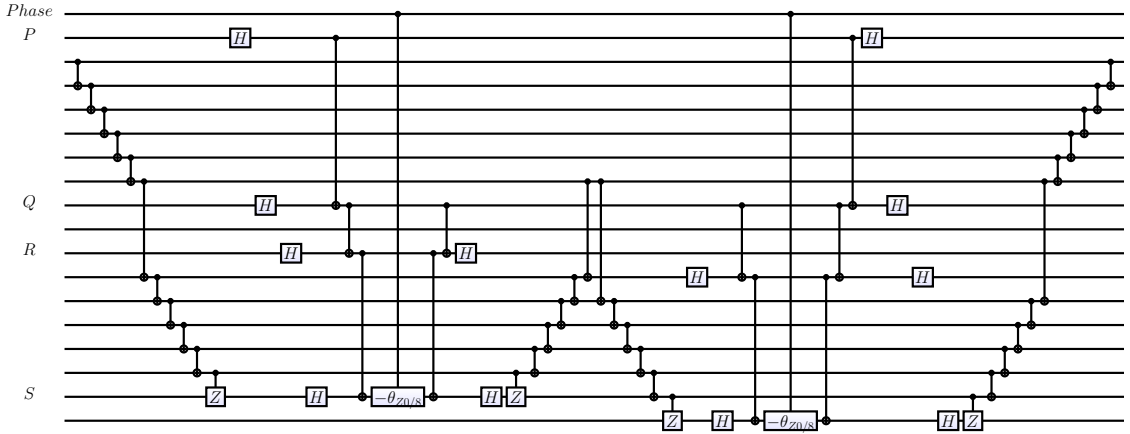

FIG. 4. Equivalent circuit to Figure 3, with redundancies cancelled.

following serendipitous cancellation often was possible: we often encountered a string of qubits acting with CNOTs on a given ancilla, with the qubits in that string in the order $b_1, b_2, ..., b_n$, where b_i labels a qubit. This string was then followed by some other string of qubits $c_1, c_2, ...$. If $c_1 = b_n, c_2 = n_{n-1}$, and so on, then we can straightforwardly cancel successive CNOTs. However, note that since all the CNOTs in both strings have the same target ancilla, they all commute with each other, and hence we can cancel *any* CNOT in the first string against *any* CNOT in the second string, so long as they have the same source qubit (i.e., $b_i = c_j$ for some i, j). See Figure 5,6.

The basic idea of all these techniques is simple: the Jordan-Wigner string computes the parity of a given set of qubits. If we then compute the parity of a set of qubits which is almost the same (for example, incrementing s to $s + 1$, adding a single qubit to that set), then much of the work of computation is already done and we do not need to do it again. In fact, similar improvements can be obtained using the original circuit of Figure 1 if, rather than choosing a specific choice of p, q, r, s and then executing all needed subterms, we instead pick a given subterm and then execute all p, q, r, s terms in lexicographic order using that subterm; however, this method does not allow the parallelization technique discussed next.

2. PARALLELIZATION

In previous work on simulation of quantum chemistry using quantum computers[17], it was found that the main limitation for practical applications will likely not be the number of coherent qubits available, but the number of gates that can be executed coherently. Therefore, approaches that reduce the gate count at the cost of a few additional qubits may be highly useful. We now introduce such an approach, which allows significant parallelization at the cost of using additional ancilla qubits.

Assume that we can execute gates in parallel that act on distinct qubits. There are many simple ways to parallelize
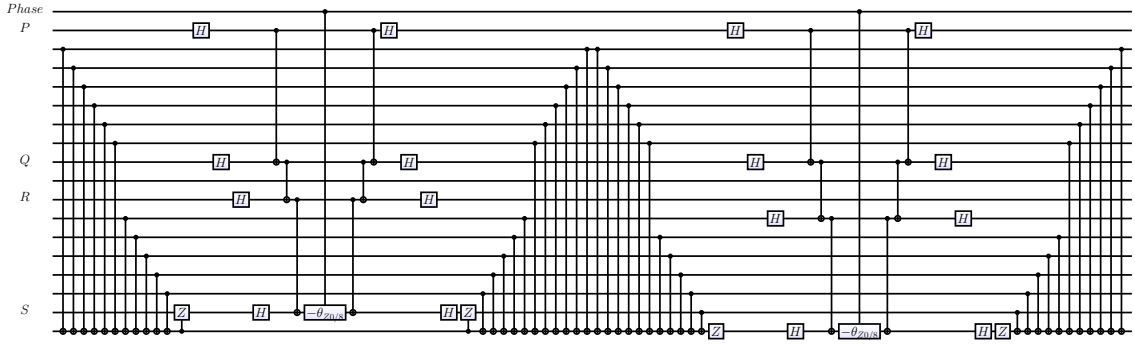
FIG. 5. Circuit with ancilla for H_{pqrs} shown for two successive choices of p, q, r, s.
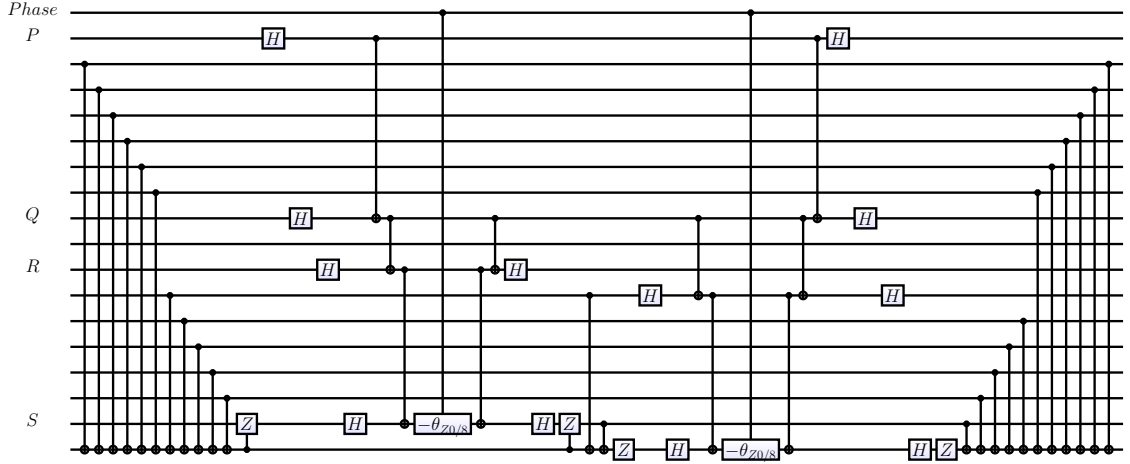

FIG. 6. Equivalent circuit to Figure 5, with redundancies cancelled. Compare to Figure 4 where, without ancillas, not as much of the CNOT string could be cancelled.

the Trotter-Suzuki step. For example, we can clearly execute two H_{pqrs} terms with p, q, r, s and p', q', r', s' at the same time if $p < q < r < s < p' < q' < r' < s'$, since the unitaries in the different gates act on different qubits. However, this can only lead to a constant factor parallelization improvement.

Using the circuit with additional ancilla qubits allows much better parallelization, improving parallel depth by a factor of $\Theta(N)$ compared to the serial gate count (thus giving a total $\Theta(N^2)$ improvement compared to the original). Consider two H_{pqrs} terms. Given any choice of p, q, r, s and p', q', r', s' for which the sites p', q', r', s' intersect an *even* number of sites in the Jordan-Wigner string of the p, q, r, s term, the p', q', r', s' term does not change the parity and hence can be moved through the Jordan-Wigner string. We refer to this as *nesting*, since we can execute terms in parallel when one sits inside another (for example, when $p < p' < q' < q$ and $r < r' < s' < s$), as illustrated in Figure 7. For a specific illustration, Figure 9 shows the circuit of Figure 8 after nesting and CNOT cancellation; both of these equivalent circuits show a few terms in succession from a real molecule. Different orderings are possible and we have not fully explored the optimum ones, but up to $\Theta(N)$ terms can be executed in parallel this way with an appropriate term ordering.

3. IMPLEMENTATION ON LIQ$Ui|\rangle$

To determine gate counts, the circuit for a single Jordan-Wigner step was specified by giving a sequence of p, q, r, s and then the various circuits described above were used to implement each such term; then, LIQ$Ui|\rangle$ automatically looked for cancellation of successive gate using rules such as the square of a CNOT or a Hadamard being equal to the identity and also automatically looked for possible nestings. LIQ$Ui|\rangle$ also used the commutation rule discussed at the end of the above section to look for possible cancellations between CNOTs appearing in different orders.
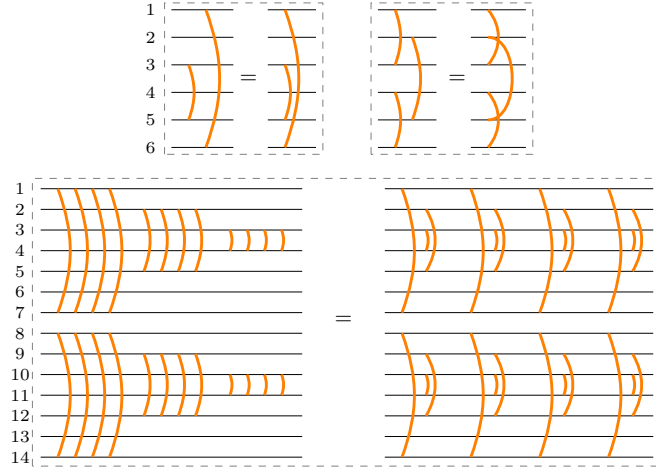
FIG. 7. Schematic illustration of nesting procedure. *Top left panel:* Nesting the terms h_{16} and h_{35}. *Top right panel:* Nesting the terms h_{1436} and h_{25}. This is allowed because the lines cross an even number of times. *Bottom panel:* Nesting of three h_{pqrs} terms, each with several subterms. This is a schematic representation of the example shown below in Figs. 8 and 9.

The results are summarized in Figure 10 for a sequence of simple molecules, showing the improvement over the original circuit, as a function of the number of spins orbitals, for a number of real molecules. The H_{pq} and H_{pqrs} terms were obtained using PyQuante[37], an open source quantum chemistry program. While the number of H_{pqrs} terms scales as N^4, many of them are zero due to symmetries of these simple molecules; we expect that the gains would be even larger for larger and less-symmetric molecules when all H_{pqrs} terms were non-zero because when many of the terms are zero, the string cancellation is only partial.

For all cases of gate counts shown, the count shown is the parallel depth, where LIQ$Ui|\rangle$ used the rule that it can execute gates in parallel which act on distinct qubits. Further, we allowed the execution of gates in parallel even if more than one of them is controlled by the phase estimation ancilla, so long as all the other qubits in the gates are distinct; that is, we allow the phase estimation ancilla to control more than one rotation at a time. To justify this last choice, suppose that we had *not* allowed the execution of such controlled rotations in parallel; we will now show that at the cost of a small number of additional ancilla qubits, we could still obtain exactly the same parallel depth as in the case where we allow the phase estimation ancilla to control multiple rotations. To do this, immediately before performing any of the Trotter-Suzuki evolution steps, we would copy (in the S^z basis) the state of the phase estimation ancilla onto some number of additional ancilla qubits. Then, we would use these additional ancillas to control the rotations so that every controlled rotation was controlled by a distinct qubit and the gates did not overlap. Finally, at the end of all the Trotter-Suzuki evolution steps we would apply CNOT gates from the phase estimation ancilla onto the additional ancillas to restore them all to the $|0\rangle$ state, obtaining the desired unitary evolution.

The $\Theta(N)$ gains due to removing Jordan-Wigner strings and parallelization do lead to only modest improvement for these size molecules, but they are indeed seen to increase with N. The $\Theta(N^2)$ improvement is not obvious in these figures, but we have found that the ratios in gate count (when going from the Whitfield et. al. circuit to the ancillas, or when adding nesting to the ancillas) can be approximately fit to a linear function of N, $a + b * N$ with $b << a$, suggesting that at larger N the relative gains will indeed become more significant.

At the same time, some seemingly minor improvements led to a large constant factor improvement in gate depth. One was to change the ordering of the qubits, from the order of Ref. 17 where the qubits corresponded to (in order) the up and down spin states on the first orbital, then the up and down spin states and the second orbital, and so on, to a different order where first all up spin states were encoded and then all down spin states; this new order is referred to as "HalfUp" in Figure 10. Another improvement is to change the sequence of subterms from that in Ref. 10 to a sequence which *minimized* the number of ways in which the basis changed at each step. Thus, rather than executing in the order

$$\sigma^X\sigma^X\sigma^X\sigma^X, \sigma^Y\sigma^Y\sigma^Y\sigma^Y, \sigma^X\sigma^X\sigma^Y\sigma^Y, \sigma^Y\sigma^Y\sigma^X\sigma^X$$

we change to

$$\sigma^X\sigma^X\sigma^X\sigma^X, \sigma^X\sigma^X\sigma^Y\sigma^Y, \sigma^Y\sigma^Y\sigma^Y\sigma^Y, \sigma^Y\sigma^Y\sigma^X\sigma^X,$$

allowing further cancellation of the basis change gates since two basis changes cancel at each step. This modified sequence is referred to as "Optimized PQRS sub-terms" in Figure 10.
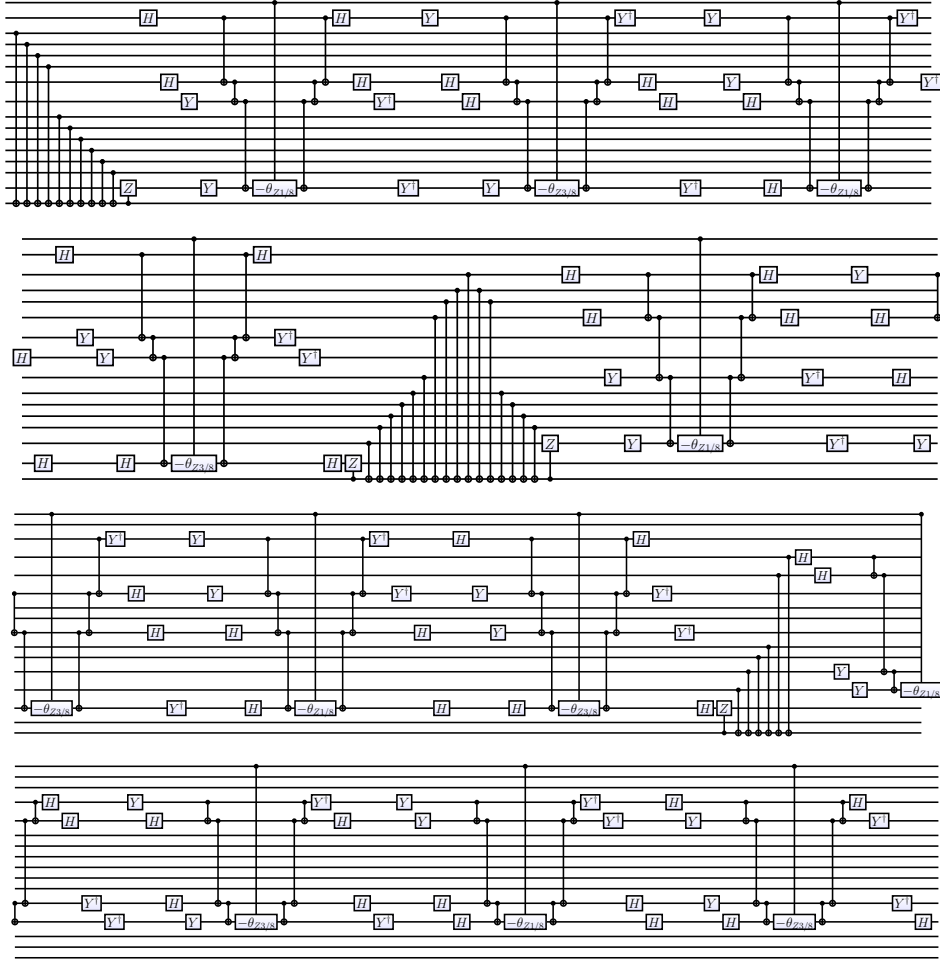
FIG. 8. Example of a sequence of unnested H_{pqrs} terms.

All simulations were implemented using a first order Trotter-Suzuki decomposition. Similar improvements occur using a second order Trotter-Suzuki decomposition, but in fact in this case the first order Trotter-Suzuki gives the same error scaling as a second order one, as shown in App. A.

4. TROTTER-SUZUKI TERM ORDERING

There are many possible orderings of the terms in the Trotter-Suzuki decomposition (lexicographic, by magnitude, etc...) and certain term orderings can greatly reduce the error at a given Trotter step δ_t. In the standard Hartree-Fock approach to a quantum chemistry problem, a variational state is constructed by finding a basis of single-particle states and then occupying certain orbitals, while others are empty. A variational optimization is then performed over the single-particle states and occupation vectors. Let $n_r = 0, 1$ be the occupancy of the r-th orbital in the lowest-energy solution. Then, in the Hartree-Fock basis for all $p \neq q$,[38]

$$t_{pq} + \frac{1}{2} \sum_r V_{prrq} n_r = 0. \tag{6}$$

That is, the "effective" hopping (t_{pq} corrected by interaction terms) becomes diagonal in this basis.

One finds that for many molecules, the Hartree-Fock solution is a reasonable approximation, with the occupancies of the Hartree-Fock orbitals in the full ground state obtained by FCI being close to 0 or 1. Hence, the above equation holds to reasonable approximation in expectation. This suggests the following "Interleaved Term Order":
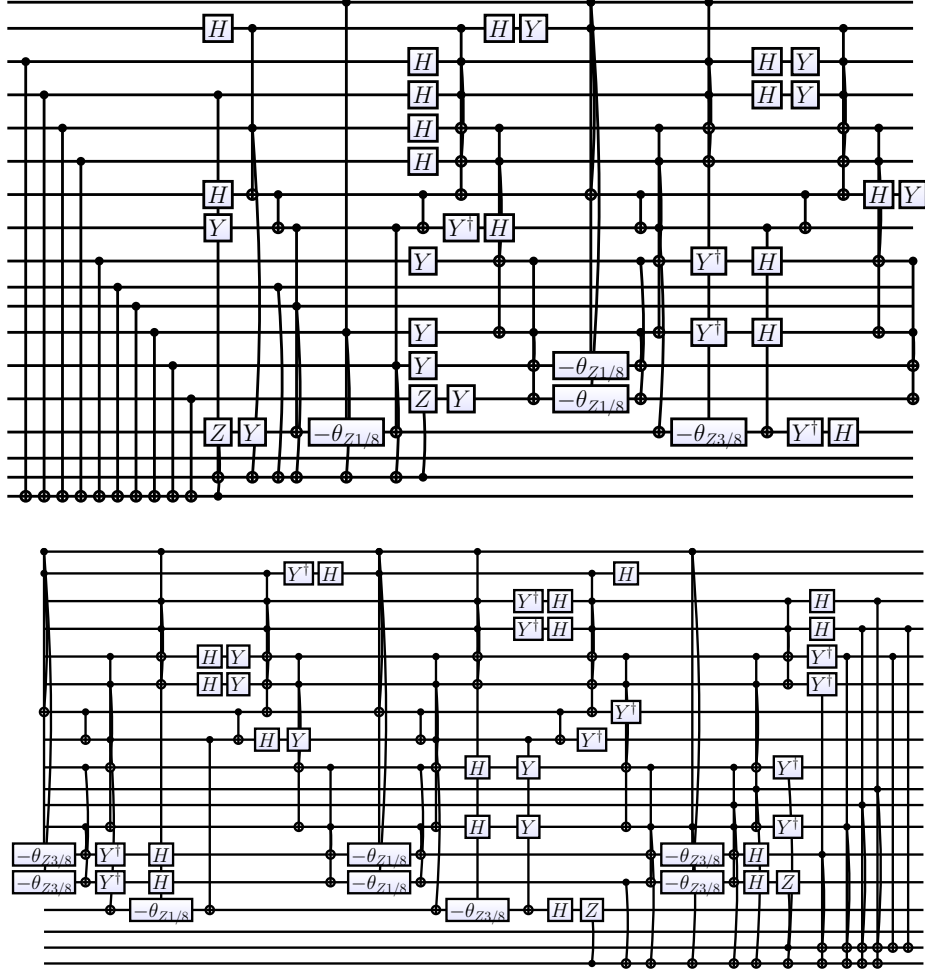
1. Execute all H_{pp} and H_{pqqp} terms.

FIG. 9. Circuit that implements the same unitary as Figure 8, after nesting to reduce parallel depth.

2. For each p, q **do**

      a. Do H_{pq} and all H_{prrq} terms.

3. Do all H_{pqrs} terms.

The terms in **1** commute with each other, as do those in **2a**, and thus can be executed in any order. Since the terms in **3** are the most numerous by far, we execute them in an order to minimize parallel depth as described above.

Unfortunately, we do not yet have a theory that allows us to understand the scaling of the error for large molecules, but we empirically find consistent large reductions in Trotter error using this ordering, largely because the H_{pq} terms are much larger than the others. Figure 11 shows the reduction in Trotter error for H_2O. Similar though slightly smaller, improvements were seen for other molecules; one notable exception was for N_2, where the Hartree-Fock basis found by PyQuante poorly captured the occupied orbitals in the real molecule. We expect that working in the correct single-particle basis would lead to a large gain there as well.

In general, this approach relies on working in a basis in which Eq. (6) holds, and thus it is most applicable to quantum chemistry; for problems such as the Hubbard model, it is possible to find such a basis of course (using a basis of Fourier modes), but this leads to a large overhead by increasing the number of terms in the Hamiltonian. On a speculative note, it may be possible in the Hubbard model to use a wavelet basis which would preserve approximate locality while still approximately satisfying Eq. (6).
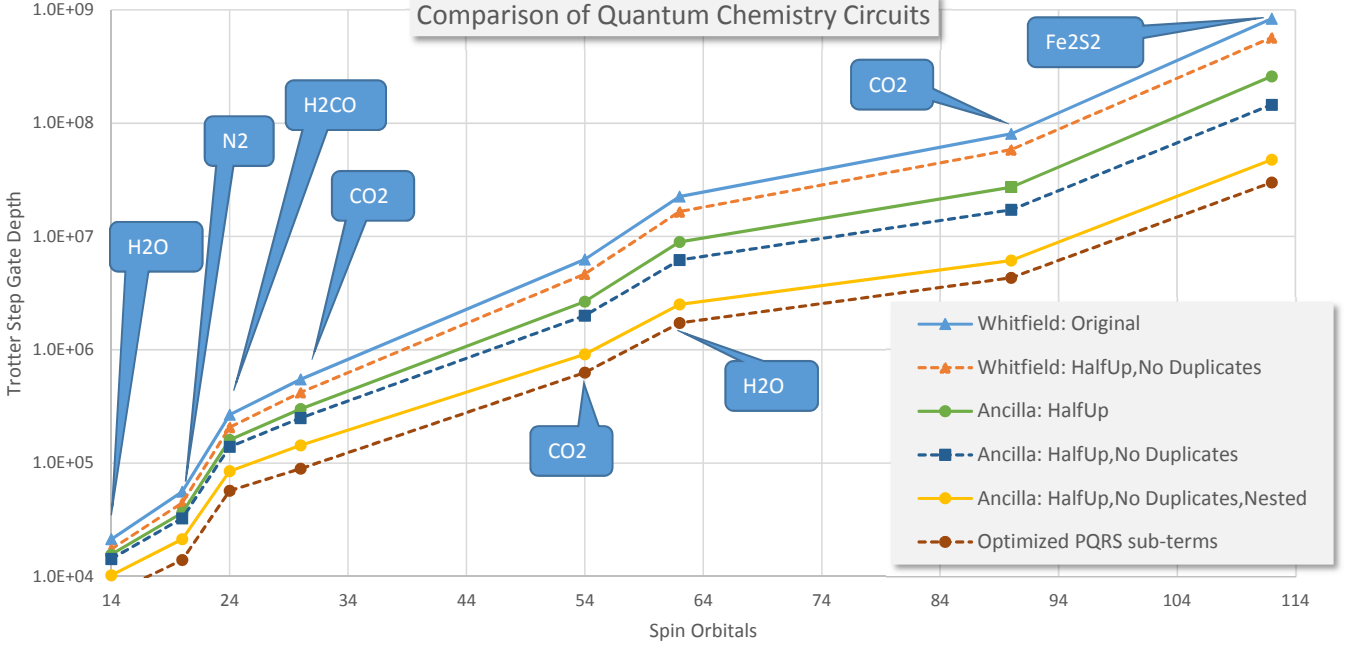
FIG. 10. Gate counts for various circuits, as a function of number of spin orbitals, for various molecules. Horizontal axis indicates number of spin orbitals; each point is labelled by the corresponding molecule and certain molecules appear multiple times as the same molecule was studied with different basis sets. Vertical axis denotes gate depth. The lines marked Whitfield denote the original Whitfield et. al. circuit[10]. "HalfUp" denotes a change in ordering of the qubits (see text). Ancilla denotes the use of the ancilla circuit. "No Duplicates" means that LIQ$Ui|\rangle$ searched for gates that could be cancelled, such as pairs of CNOTs that square to the identity. "Nested" refers to the use of nesting. The bottom line, marked "Optimized PQRS sub-terms" is "Ancilla: HalfUp, No Duplicates, Nested" with a modified sequence of subterms, described in text.

5. CORRECTED HAMILTONIAN

The final improvement leads to only a minor improvement (see Figure 11), but the idea may be useful more generally. For a given δ_t, rather than simulating Hamiltonian H, we simulate a different Hamiltonian, in which the coefficients t_{pq} and V_{pqrs} are modified in a way that goes to zero as $\delta_t \to 0$.

To define the modification, we separate the Hamiltonian into diagonal (H_{pp} and H_{pqqp}) terms, and off-diagonal terms. We make an approximation in the spirit of Hartree-Fock, defining an effective diagonal term $\omega_p = t_{pp} + (1/2)\sum_q V_{pqqp}$. We then calculate for a Hamiltonian $H = H_{\rm diag} + H_{\rm off-diag}$ the effect of the off-diagonal term to second order in perturbation theory. The shift in energy E_i of eigenstate i is given by $\sum_{j \neq i}(E_i - E_j)^{-1}|\langle\psi_j|H_{\rm off-diag}|\psi_j\rangle|^2$. However, using a second-order Trotter-Suzuki expression, the energy shift is instead $\sum_{j \neq i}(E_i - E_j)^{-1}f(\delta_t(E_i - E_j))^{-2}|\langle\psi_j|H_{\rm off-diag}|\psi_j\rangle|^2$, where

$$f(x) = \sqrt{\frac{2(1 - \cos(x))}{\sin(x)}}. \tag{7}$$

To compensate, we make the modification

$$V_{pqrs} \leftarrow V_{pqrs}f(\delta_t(\omega_p + \omega_q - \omega_r - \omega_s)), \tag{8}$$

and $t_{pq} \leftarrow t_{pq}f(\delta_t(\omega_p - \omega_q))$.

While only justified to second order, we find that this expression gives a slight consistent improvement in all cases. See the lines labelled "Diagonal Fix" in Figure 11. The function f is slightly larger than 1 for $x \neq 0$, so it increases the off-diagonal terms and goes to 1 as $\delta_t \to 0$. Heuristically, this can be justified as follows: for large $E_i - E_j$, the effect of the off-diagonal term should produce only a virtual transition lasting time $\sim (E_i - E_j)^{-1}$; however, the minimum time of the transition is δ_t, removing certain short-time transitions; hence we increase the magnitude of the term to compensate.
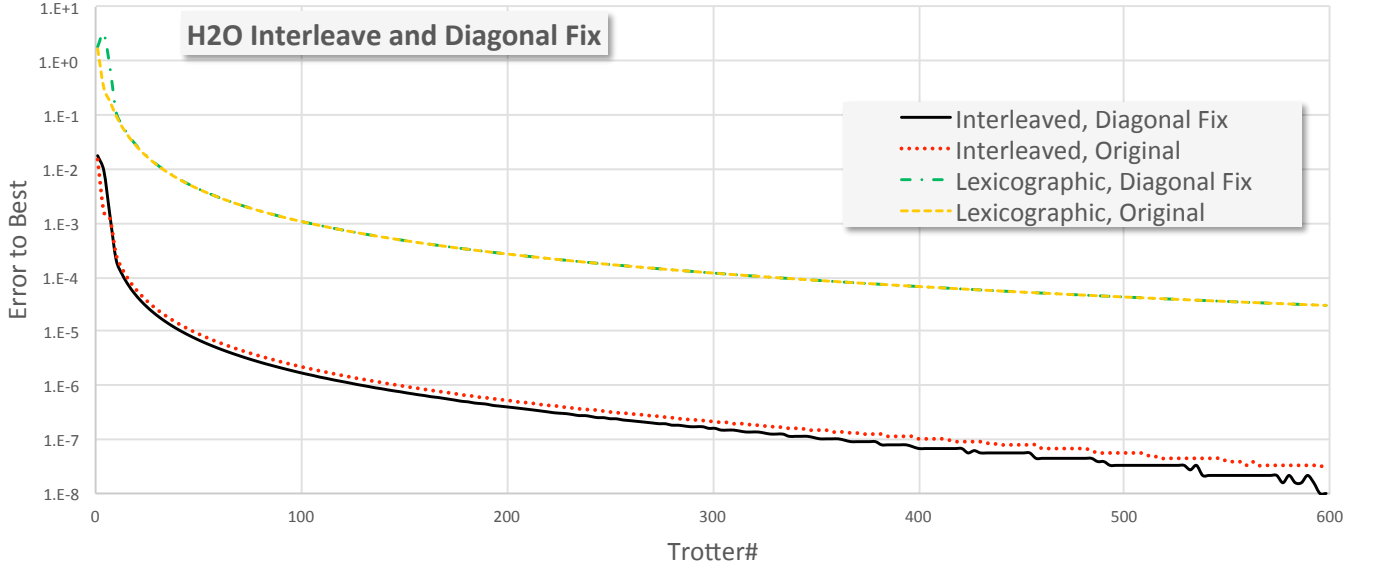
FIG. 11. Trotter error for H_2O, as a function of Trotter number n, for time step $\delta_t = 1/n$. Lexicographic (Lex) and Interleaved (Int) orders; Fix is explained in section 5. y-axis is base-10 logarithm of error in ground state measured in Hartrees.

This modification can be thought of as using a "renormalization group corrected Hamiltonian". Perhaps a more general approach can allow the addition of other terms to the Hamiltonian to compensate; however, we would like to avoid adding any terms involving six or more Fermi operators as they would be computationally expensive to implement.

6. DISCUSSION

We have described several improvements to the basic Trotter-Suzuki approach used in the quantum simulation of quantum chemistry. The cancellation of Jordan-Wigner strings and the nesting improve the N^9 scaling of the gate depth found in Ref. 17 to $O(N^7)$. Using approaches based on teleportation, it is possible to execute a sequence of measurements in parallel to implement in time $O(1)$ a CNOT string[39] and even to implement, in certain coding schemes, an arbitrary Clifford operation[40]; with these approaches the gate depth using previous circuits scales as N^8. Thus, even if all Clifford operations are assumed to be free (i.e., the cost of a Clifford operation is negligible compared to the cost of an arbitrary rotation), it is still worth using the circuits here as they enable additional parallelization of the rotations using nesting, reducing the time from N^8 scaling to N^7 scaling.

Assuming that the improvement from the interleaved term ordering continues to hold for larger molecules, we obtain a combined reduction in gate depth of two to three orders of magnitude for molecules such as Fe_2S_2 in a basis with 112 spin orbitals, with larger improvements arising if we insist upon smaller Trotter error (see Figure 11). Other schemes other than Trotter-Suzuki are certainly possible, with the two schemes with best theoretical scaling being Refs. 28 and 30. However, Trotter-Suzuki has advantages of allowing improved parallelism, especially when each term commutes with many of the other terms. Also, the possibility of choosing different orderings allows us to exploit the ability to choose a Hartree-Fock basis in which the errors due to larger terms partly cancel. Future work using LIQ$Ui|\rangle$ will allow accurate gate count comparisons using other schemes.

Appendix A: Error Scaling in First Order Trotter-Suzuki

Given a Hamiltonian H written as a sum of terms

$$H = \sum_{j=1}^{N_{\text{terms}}} H_j, \tag{A1}$$

and given the ability to construct small quantum circuits to implement $\exp(iH_j\delta_t)$ for any j and any δ_t, the standard first order Trotter-Suzuki approximation for $\exp(iH\delta_t)$ is

$$\exp(iH\delta_t) \approx \exp(iH_1\delta_t)\exp(iH_2\delta_t)...\exp(iH_{N_{\text{terms}}}\delta_t), \tag{A2}$$

while the second order Trotter-Suzuki approximation is

$$\exp(iH\delta_t) \tag{A3}$$
$$\approx \exp\left(iH_1\frac{\delta_t}{2}\right)\exp\left(iH_2\frac{\delta_t}{2}\right)...\exp\left(iH_{N_{\text{terms}}}\frac{\delta_t}{2}\right)$$
$$\exp\left(iH_{N_{terms}}\frac{\delta_t}{2}\right)...\exp\left(iH_2\frac{\delta_t}{2}\right)\exp\left(iH_1\frac{\delta_t}{2}\right).$$

As one may guess from the names, the first expression is exact to order δ_t, giving an error of order δ_t^2, while the second is exact to order δ_t^2, giving an error of order δ_t^3.

Ultimately, the goal is to approximate $\exp(iHt)$ for some large t. One does this by writing $\exp(iHt) = \exp(iHt/n)^n$ for some integer $n \geq 1$, and then using a Trotter-Suzuki approximation for $\exp(iH\delta_t)$ with $\delta_t = t/n$. Taking into account that the we have decomposed the evolution for time t into n smaller steps, and using the error estimates for each time step above, one would expect that the error from first order Trotter-Suzuki scales as $1/n$ and that from second order Trotter-Suzuki scales as $1/n^2$. In fact, in practice we have observed that both give errors *in the ground state energy* scaling proportional to $1/n^2$; that is, first order Trotter-Suzuki has *better* error scaling than expected.

In this appendix we explain why this occurs, showing that this better error scaling is to be expected assuming that the Hamiltonian is real and has a non-degenerate ground state, both of which are true for the Hamiltonians considered in quantum chemistry. If instead we had a non-zero magnetic field, the Hamiltonian would not be real, and such better error scaling would not be expected.

Before explaining this, we note that the result holds trivially if $N_{\text{terms}} = 2$, as then the first order and second order approximations to the unitary $\exp(iH\delta_t)$ have the same set of eigenvalues (to verify this, conjugate the first order approximation by the unitary $\exp(iH_2\frac{\delta_t}{2})$ and this gives the second order approximation). However, our result holds for arbitrary N_{terms}.

To show the result, an explicit computation using commutators gives

$$\exp(iH_1\delta_t)\exp(iH_2\delta_t)...\exp(iH_{N_{terms}}\delta_t) \tag{A4}$$
$$= \exp(i(H+\Delta)\delta_t + O(\delta_t)^3),$$

where

$$\Delta = i\sum_{j<k}[H_j, H_k]\frac{\delta_t}{2}. \tag{A5}$$

Hence, up to an error which is $O(\delta_t^3)$, the first order Trotter-Suzuki formula computes the evolution under a modified Hamiltonian, $H + \Delta$. Let E_0 be the ground state energy of H and let E_0' be the ground state energy of $H+\Delta$, so that the error in ground state energy using a first order Trotter-Suzuki formula is equal to $E_0' - E_0 + O(1/n^2)$. Assuming H indeed has a non-degenerate ground state, then $E_0' - E_0$ is given by first order perturbation theory, up to error $O(\delta_t^2)$. However, if H is real and has a non-degenerate ground state, then the ground state expectation value of Δ vanishes. Hence, $E_0' - E_0$ is $O(\delta_t^2)$. So, in this case, indeed first order Trotter-Suzuki gives the correct energy up to error $1/n^2$. To see that the ground state expectation value of Δ vanishes, let ψ_0 be the ground state wavefunction, which is real, while Δ is pure imaginary. Hence, the expectation value $\langle\psi_0|\Delta|\psi_0\rangle$ is pure imaginary; however, this expectation value is a real number, being the expectation value of a Hermitian operator, hence it equals zero.

It is interesting to observe that the non-degenerate ground state is necessary. Consider a single qubit with $H_1 = \sigma_x, H_2 = \sigma_z, H_3 = -\sigma_x, H_4 = -\sigma_z$ so that $H = 0$ and $E_0 = 0$. One may verify that first order Trotter-Suzuki gives an error proportional to $1/n$ in this case.

---

[1] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).
[2] S. P. Jordan, K. S. M. Lee, and J. Preskill, Science **336**, 1130 (2012).
[3] D. A. Lidar and H. Wang, Phys. Rev. E **59**, 2429 (1999).
[4] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **64**, 022319 (2001).

[5] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).

[6] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Proceedings of the National Academy of Sciences **105**, 18681 (2008).

[7] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann, Physical Chemistry Chemical Physics **10**, 5388 (2008).

[8] I. Kassal and A. Aspuru-Guzik, The Journal of Chemical Physics **131**, (2009).

[9] B. P. Lanyon *et al.*, Nat Chem **2**, 106 (2010).

[10] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Molecular Physics **109**, 735 (2011).

[11] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, Annual Review of Physical Chemistry **62**, 185 (2011).

[12] M.-H. Yung, J. D. Whitfield, S. Boixo, D. G. Tempel, and A. Aspuru-Guzik, Preprint (2012), arXiv:1203.1331.

[13] J. D. Whitfield, Preprint (2013), arXiv:1306.1147.

[14] M.-H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, Preprint (2013), arXiv:1307.4326.

[15] L. Lamata, A. Mezzacapo, J. Casanova, and E. Solano, Preprint (2013), arXiv:1312.2849.

[16] B. Toloui and P. J. Love, Preprint (2013), arXiv:1312.2579.

[17] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, arXiv:1312.1695.

[18] H. Nakano and T. Sakai, J. Phys. Soc. Jpn. **80**, 053704 (2011).

[19] A. M. Läuchli, J. Sudan, and E. S. Sørensen, Physical Review B **83**, 212401 (2011).

[20] S. Capponi, O. Derzhko, A. Honecker, A. M. Läuchli, and J. Richter, Phys. Rev. B **88**, 144416 (2013).

[21] Z. Gan and R. J. Harrison, in *Supercomputing, 2005. Proceedings of the ACM/IEEE SC 2005 Conference* (IEEE, 2005) pp. 22–22.

[22] Y. Kurashige, G. K.-L. Chan, and T. Yanai, Nat Chem **5**, 660 (2013).

[23] D. Berry, G. Ahokas, R. Cleve, and B. Sanders, Commun. Math. Phys. **270**, 359 (2007).

[24] A. M. Childs and R. Kothari, in *Theory of Quantum Computation, Communication, and Cryptography*, Lecture Notes in Computer Science, Vol. 6519, edited by W. Dam, V. Kendon, and S. Severini (Springer Berlin Heidelberg, 2011) pp. 94–103.

[25] A. Daskin and S. Kais, The Journal of Chemical Physics **134**, (2011).

[26] A. M. Childs and N. Wiebe, Quantum Information & Computation **12**, 901 (2012), arXiv:1202.5822.

[27] A. M. Childs and N. Wiebe, J. Math. Phys. **54**, 062202 (2013), arXiv:1211.4945.

[28] D. W. Berry and A. M. Childs, Quantum Information & Computation **12**, 29 (2012), arXiv:0910.4157.

[29] A. Daskin, A. Grama, G. Kollias, and S. Kais, The Journal of Chemical Physics **137**, (2012).

[30] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Preprint (2013), arXiv:1312.1414.

[31] R. Babbush, P. J. Love, and A. Aspuru-Guzik, Preprint (2013), arXiv:1311.3967.

[32] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Preprint (2013), arXiv:1304.3061.

[33] H. F. Trotter, Proc. Amer. Math. Soc. **10**, 545 (1959).

[34] M. Suzuki, Communications in Mathematical Physics **51**, 183 (1976).

[35] S. Bravyi and A. Kitaev, Ann. Phys. **298**, 210 (2002).

[36] D. Wecker and K. Svore, Preprint (2014), arXiv:1402.4467.

[37] Available at http://pyquante.sourceforge.net.

[38] To extremize the Hartree-Fock energy, Eq. (6) must hold for p occupied and q unoccupied. A further basis rotation which does not mix occupied and unoccupied orbitals makes it hold for all $p \neq q$.

[39] N. Cody Jones, J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. V. Meter, A. Aspuru-Guzik, and Y. Yamamoto, New Journal of Physics **14**, 115023 (2012).

[40] A. Fowler, arXiv:1210.4626.