qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor

Ho Lun Tang,1 Edwin Barnes,1 Harper R. Grimsley,2 Nicholas J. Mayhall,2 and Sophia E. Economou1

1Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA
2Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA

(Dated: November 26, 2019)

Quantum simulation, one of the most promising applications of a quantum computer, is currently being explored intensely on existing hardware using the variational quantum eigensolver. The feasibility and performance of this algorithm depend critically on the form of the wavefunction ansatz. Recently, an algorithm termed ADAPT-VQE was introduced to build system-adapted ansätze with substantially fewer variational parameters compared to other approaches. However, deep state preparation circuits remain a challenge. Here, we present a hardware-efficient variant of this algorithm, which we term qubit-ADAPT. Through numerical simulations on H₂, LiH and H₂, we show that qubit-ADAPT reduces the circuit depth by an order of magnitude while maintaining the same accuracy as the original ADAPT-VQE. A key result of our approach is that the additional measurement overhead of qubit-ADAPT compared to fixed-ansatz variational algorithms scales only linearly with the number of qubits. Our work provides a crucial step forward in running algorithms on near-term quantum devices.

I. INTRODUCTION

Finding the ground state of a many-body interacting electronic Hamiltonian is one of the most important problems in modern quantum chemistry and physics. As the dimension of the Hamiltonian scales exponentially with the number of particles, classical computers can only handle problems with small numbers of electrons. Many classical computational techniques were developed to approximate the ground state. These techniques can be classified into two types. One is the mean field approach, such as Hartree-Fock (HF) and density functional theory [1, 2], which fails to describe strongly correlated systems. Alternatively, Full Configuration Interaction (FCI), which retains all the Slater determinants and is hence exact, comes with an exponential computational cost. A radically different approach is Feynman’s proposal to study quantum systems using quantum computers [3].

The long-term method for simulating chemical systems with quantum computers is the quantum phase estimation algorithm (PEA) [2], which shows an exponential speedup over classical algorithms [4, 5]. Since the number of gates, i.e. unitary operations, involved in this algorithm is very large, it requires a long coherent evolution which can only be realized in fault-tolerant quantum computers [6]. These scalable, error-correcting devices may take decades to realize experimentally. In the meantime, the community is exploring algorithms that can be applied to existing and near-term processors, namely noisy intermediate-scale quantum (NISQ) devices [7].

A promising algorithm for NISQ hardware is the variational quantum eigensolver (VQE) [8, 9]. VQE is a hybrid method that combines classical computational power with a quantum processor. The idea is based on the variational principle in quantum physics. The VQE algorithm constructs the wavefunction ansatz by applying gates on a quantum device and estimates the average energy by measuring the Hamiltonian on that device. The energy minimization is done by a classical computer, so quantum resources are only used for the classically intractable parts (state preparation and energy evaluation) of the calculation [10]. Compared to PEA, VQE has much more modest requirements on the coherence times of the quantum processor, so it has already been realized on NISQ devices, such as superconducting qubits [11, 12], photons [8] and trapped ions [13, 14]. The accuracy of VQE is highly dependent on the explicit form of the wavefunction ansatz.

One of the commonly used ansätze for VQE is unitary coupled cluster singles and doubles (UCCSD) [15, 16]. Stemming from the coupled cluster theory commonly used in chemistry, this unitary version is more suitable for quantum circuit implementation. The ‘singles and doubles’ in UCCSD means only single and double excitation operators are included in the ansatz, each carrying its own variational parameter. The drawback of UCCSD is that including all the singles and doubles operators leads to a (potentially unnecessarily) deep circuit and a large number of parameters to optimize. To reduce this complexity, several alternatives have been proposed which attempt to keep only the most important operators [17–23]. UCCSD is also generally not exact and suffers from ambiguities in operator ordering upon factorization into a product of exponentiated operators (Trotterization), which is a necessary step in converting the ansatz to a state preparation circuit [24, 25]. Another approach to building the ansatz is to use the most accessible gates in the quantum device, alternating single-qubit gates and two-qubit gates layer by layer. This is referred to as a hardware-efficient ansatz [12]. Although this approach lessens the demands on the quantum processor, the Hilbert space spanned by the wavefunction ansatz may be too large and could result in difficulties in parameter optimization [26]. This problem was addressed by constructing particle-conserving entangling gates in-
stead of ordinary two-qubit gates \cite{27,28} and symmetry-preserving circuits \cite{29}. While these fixed-ansatz approaches can be applied to any problem and can reduce the number of variational parameters and circuit depths, further optimization may still be possible by tailoring the ansatz to a given simulation problem.

Recently, a new algorithm that provides a systematic method to build an ansatz dynamically was introduced. This algorithm, termed Adaptive Derivative Assembled Pseudo-Trotter (ADAPT) VQE \cite{30}, employs a predetermined pool of operators from which the ansatz is dynamically constructed. The ansatz is grown iteratively, such that at each step, the operator that affects the energy the most is added to the ansatz. Using fermionic operators as a pool, Ref. \cite{30} demonstrated that ADAPT-VQE substantially outperforms UCCSD, in terms of both number of variational parameters and accuracy. This result demonstrates the promise of the ADAPT algorithm. However, due to the gate overhead of the fermion-to-spin mapping, the operators considered in Ref. \cite{30} translate to a fairly large number of quantum gates, and, therefore, while the number of parameters is very low, the circuit depth may still be impractically large, limiting the applicability of ADAPT-VQE to NISQ devices. In addition, it is not clear how the operator pool should be chosen in general or how many operators it should contain.

In this paper, we address these issues and enhance the practicality of ADAPT-VQE by considering a more hardware-efficient pool of operators. We term this algorithm qubit-ADAPT, in contrast with the implementation in Ref. \cite{30}, which we refer to as fermionic-ADAPT in this paper. Through classical simulations of several different molecules, we demonstrate that compared to fermionic-ADAPT, qubit-ADAPT reduces the circuit depth by an order of magnitude while maintaining the same accuracy. Moreover, we introduce a pool completeness criterion that determines whether a given pool will generate an exact ADAPT ansatz. We provide evidence that the minimal number of pool operators that satisfy this condition grows linearly with the number of qubits. This is much smaller than the quartic scaling originally assumed in fermionic-ADAPT, and it implies that the additional measurement overhead of ADAPT-VQE remains modest for larger systems (increasing only linearly over conventional, fixed-ansatz VQEs). Our results pave the way toward both practical and accurate VQE algorithms on NISQ devices.

II. RESULTS

The ADAPT ansatz is grown by one operator \( \hat{\tau}_i = -\hat{\tau}_i^\dagger \) at each iteration, and after the \( n \)-th iteration is given by

\[
|\psi^{ADAPT}(\vec{\theta})\rangle = e^{\theta_n \tau_n} \cdots e^{\theta_2 \tau_2} e^{\theta_1 \tau_1} |\psi^{HF}\rangle,
\]

where \( |\psi^{HF}\rangle \) is the HF state. These operators are selected from an operator pool defined upfront. At each iteration, the operator which induces the maximum change to the energy is selected. This energy response is represented by the gradient of the energy with respect to the corresponding parameter, i.e.

\[
\frac{\partial}{\partial \theta_i} (E) = \langle \psi | \hat{H}, \tau_i | \psi \rangle,
\]

which can be measured on the quantum device. The ansatz keeps growing until the norm of the gradient vector,

\[
|\vec{g}| = \sqrt{\sum_i \left( \frac{\partial}{\partial \theta_i} \langle E \rangle \right)^2},
\]

is zero or, in practice, smaller than a chosen threshold, \( \epsilon \). Compared to ordinary VQE, ADAPT-VQE requires additional measurements to obtain the gradient at each iteration; the number of these measurements is roughly equal to the size of the pool times the number of terms in the Hamiltonian \cite{31}, although the number of measurements needed to compute the mean energy can be decreased using recently developed techniques \cite{32,33}.

In order to attain a compact circuit for the ADAPT ansatz, we want to minimize both the number of parameters and circuit depth. While the first requirement can be satisfied by the general structure of the ADAPT-VQE algorithm, the second one is not guaranteed and depends on the chosen pool. Using fermionic operators, a parameter-efficient operator pool can be constructed from spin-adapted single excitation operators:

\[
\hat{\tau}_1 \propto |\uparrow\rangle_a \langle \uparrow|_b + |\downarrow\rangle_a \langle \downarrow|_b - h.c.,
\]

and double excitation operators:

\[
\hat{\tau}_{2,r} \propto |T,1\rangle_{pq} \langle T,1|_{rs} + |T,-1\rangle_{pq} \langle T,-1|_{rs}
+ |T,0\rangle_{pq} \langle T,0|_{rs} - h.c.
+ |S,0\rangle_{pq} \langle S,0|_{rs} - h.c.
\]

where \( a, b, p, q, r, s \) are spatial orbitals and \( T, S \) refer to triplets and singlets formed by \( p, q \) or \( r, s \). To implement ADAPT-VQE on qubits, we have to map these fermionic operators to Pauli operators, which has the consequence that a parameter-efficient pool is most likely not gate-efficient. In this paper, we employ the Jordan-Wigner (JW) mapping, i.e., \( \hat{a}_i \rightarrow \prod_{j=0}^{i-1} Z_j (X_i - i Y_i) \).

A. Circuit depth estimate for fermionic-ADAPT

In the JW mapping, a double excitation operator may contain more than one fermionic operator \( a_i^\dagger a_j^\dagger a_k a_l - h.c. \), which is transformed into at most 8 Pauli strings. A product of four fermionic operators gives 16 terms in total, but each symmetric term is cancelled by its Hermitian conjugate \cite{34}. These excitation operators conserve \( S^2 \), \( S_z \), and particle number by summing a number of
FIG. 1: Ground state energy difference between ADAPT and the exact (FCI) result for (a,d) $H_4$, (b,e) LiH, and (c,f) $H_6$ molecules. Results are shown for qubit-ADAPT (blue) and fermionic-ADAPT (orange). Panels (a)-(c) show the energy difference as a function of ADAPT iterations (which is equal to the number of variational parameters), while (d)-(f) show it as a function of the number of CNOTs in the corresponding state preparation circuit. Although qubit-ADAPT requires more variational parameters, it entails significantly fewer CNOTs compared to fermionic-ADAPT.

Pauli strings together, which results in a high gate count per excitation operator.

We can make a rough estimate of the number of CNOT gates involved in one operator from the fermionic pool. We consider the generalized singles-doubles excitation fermionic pool. ‘Generalized’ means we do not restrict to excitations between occupied and unoccupied orbitals, but allow all combinations of orbitals. To obtain the gate count, we perform first-order Trotterization of each unitary, i.e., $e^{i\theta \tau_i} \approx \prod_j e^{i\theta P_j}$, where $P_j$ are the Pauli strings appearing in $\tau_i$ after the JW mapping. For simplicity, we assume only double excitation operators are picked by the algorithm, and we do not perform any transpilation and cancellation on the resulting circuit, which makes this a conservative estimate. The number of CNOTs needed for a single Pauli string $P_i$ with length $q$ is $2(q-1)$ [35]. Here, the length $q$ is the number of non-identity Pauli operators in the string. The average number of CNOTs involved in a spin-adapted double excitation operator is approximately $N_{Pauli}(6 + 2N_Z)N_{spin}$, where $N_{Pauli}$ is the average number of Pauli strings in a doubles operator $a_p^\dagger a_q^\dagger a_r a_s - h.c.$, $N_Z$ is the average number of Pauli $Z$’s in a Pauli string due to the anticommutation relation of fermionic operators, and $N_{spin}$ is the average number of doubles operators summed in the spin-adapted grouping in Eq. (5). These quantities are given by

$$N_{Pauli} = \frac{8k(5k - 7)}{k(5k - 1) - 8},$$  \hspace{1cm} (6)$$

$$N_Z = \frac{12k^3 - 30k^2 + 34k - 20}{3(5k^2 - k - 8)},$$  \hspace{1cm} (7)$$

$$N_{spin} = \frac{5k^2 - k - 8}{k^2 + k},$$  \hspace{1cm} (8)$$

where $k$ is the number of spatial orbitals. For large $k$, the number of CNOTs in a spin-adapted doubles operator is approximately $64k$.

**B. qubit-ADAPT framework**

Because the spin-adapted fermionic operators each introduce a large number of CNOTs into the state preparation circuit, we are motivated to construct a new pool
consisting of operators that involve fewer CNOTs. One way is to break down the spin-adapted fermionic operators after the JW mapping and choose the individual Pauli strings as the operator pool \( \hat{\tau} = \hat{P} = i \prod_{i} p_{i}, \) \( p_{i} \in \{X, Y, Z\} \). This more hardware-efficient choice effectively reduces \( \bar{N}_{\text{spin}} \) and \( \bar{N}_{\text{Pauli}} \) to 1, while it contains the same basic elements as the spin-adapted fermionic pool. An important property of this qubit pool is that it only contains Pauli strings with odd numbers of \( Y \)'s because the fermionic operators are real, hence \( \hat{P} = i \prod_{i} p_{i} \) has to be real. We refer to these as “odd” Pauli strings. The remaining “even” Pauli strings have no effect on the energy. This is because \( \hat{H} \) is symmetric (time-reversal symmetry is preserved), and so the expectation value of the commutator in Eq. (2) will vanish for real \( |\psi\rangle \) if \( \hat{\tau}_{i} \) is symmetric (i.e., even). We can therefore restrict \( \hat{\tau}_{i} \) to the odd Pauli strings, which are antisymmetric. Using such operators in the pool also ensures that the ansatz remains real throughout the qubit-ADAPT algorithm, which should be the case when \( \hat{H} \) is time-reversal symmetric. The length of these strings ranges from 2 to \( n \) due to the Pauli \( Z \) chain responsible for the fermionic anticommutation relation. To further reduce gate depth, we remove these Pauli \( Z \) chains from the operators. We find that these two pools (with and without \( Z \) chains) perform similarly. The pool without \( Z \) chains gives Pauli strings with maximum length 4. The size of this pool is much smaller than the full set of Pauli strings with maximum length 4, as we only pick operators that already appear in the fermionic pool, which are capable of transforming \( |\psi^{HF}\rangle \) to the ground state. We refer to this reduced pool as the “qubit pool”.

C. Numerical simulations

We compare the performance of the qubit pool to the fermionic pool in terms of the number of parameters and the number of CNOTs for different molecules, \( H_{4} \), LiH and \( H_{6} \) (Fig. 1). \( H_{4} \) has 4 electrons in 8 spin orbitals, and the calculations are done at bond distance \( r = 1.5\) Å. LiH has 4 electrons in 12 spin orbitals and the calculations are done at bond distance \( r = 2\) Å. \( H_{6} \) has 6 electrons in 12 spin orbitals and the calculations are done at bond distance \( r = 1.5\) Å. The bond distances are chosen to ensure that correlation effects are substantial.

It is evident from Fig. 1 that in the case of the qubit pool, more operators and parameters are used compared to the fermionic pool. On the other hand, the number of CNOTs is reduced significantly, by about an order of magnitude in the case of \( H_{4} \). Switching from the fermionic pool to the qubit pool increases the number of parameters in the ansatz, which is the price for compressing the circuit depth. However, the increase in parameter number only increases the required classical computational power during the classical optimization, while the decrease in circuit depth reduces the demands on the quantum processor. For NISQ devices, the number of CNOTs that can be implemented within the coherence time is very limited, so qubit-ADAPT’s ability to divert more of the computational cost away from the quantum processor and onto the classical optimizer should be advantageous overall.

To evaluate the effectiveness of using the gradient as a means to grow the ADAPT ansatz, we compare the performance of qubit-ADAPT with random operator orderings drawn from the same qubit pool. In Fig. 2, we see that qubit-ADAPT always converges much faster than the random orderings for both \( H_{4} \) and LiH. When random orderings are used for \( H_{4} \), convergence to the ground state requires 46-78 parameters, compared to only 30 parameters for qubit-ADAPT. In the case of LiH, the random orderings require more than three times as many parameters as qubit-ADAPT to converge. We can only provide a lower bound on the number of parameters needed to converge the random ordering results in Fig. 2 (b) due to the long computational times needed in this case. These findings suggest that the role of the gradient selection in qubit-ADAPT is crucial for larger problems.
FIG. 3: qubit-ADAPT energy error of $H_4$ using the full qubit pool, and randomly chosen 1/4, 1/16 and 1/32 pools are plotted against the number of iterations. For the 1/4 and 1/16 pools, the algorithm can converge for almost every run. For the 1/32 pool, most of the runs can only reach an energy error of $10^{-4}$.

D. Operator pool reduction

So far, a drawback of the qubit pool is its large size compared to the fermionic pool, which in turn leads to proportionally more measurements at each iteration. Even though the qubit pool is defined from the fermionic pool, which is a small subset of the full Pauli group, the pool size grows quickly with the number of orbitals. However, many of these operators are redundant in the sense that eliminating them has no effect on the convergence of the algorithm. For example, there are pairs of operators that are related by a global rotation, e.g., $X_0Y_1Y_2Y_3$ and $Y_0X_1X_2X_3$, so we only need to keep one of them in the pool; discarding the other does not have any significant effect on the performance of qubit-ADAPT. In the remainder of the paper, we will examine various approaches to reducing the pool size further by identifying additional redundancies among the operators.

The redundancy in the pool can be illustrated by removing randomly selected operators from the pool and monitoring the impact on convergence. First we randomly remove 3/4 of the operators in the pool. As shown in Fig. 3, despite this large reduction of the pool size, the performance of the algorithm is similar to that with the original pool. However, as is also evident in Fig. 3, if we further remove more operators, the pool is sometimes incomplete, and the energy may not converge to the ground state energy.

We can understand the tolerance of the algorithm to the drastic reduction of the pool by studying the Hilbert space spanned by the pool operators. Starting from the expression for $|\psi^{\text{ADAPT}}(\tilde{\theta})\rangle$ in Eq. (1) and using the Baker-Campbell-Hausdorff formula repeatedly, we obtain

$$|\psi^{\text{ADAPT}}(\tilde{\theta})\rangle = e^{\sum_i \phi_i A_i} |\psi^{HF}\rangle,$$

where the $A_i$ include all the pool operators and their commutators, and the $\phi_i$ are functions of $\tilde{\theta}$. Therefore, the Hilbert space spanned by the pool is determined by the set $\{A_i\}$. Note that if the original pool is comprised of odd Pauli strings, then so is $\{A_i\}$. If the operators in $\{A_i\}$ can transform the reference state to any real state in the $n$-qubit Hilbert space, then the qubit-ADAPT ansatz is guaranteed to be exact, and it is capable of converging to the ground state. (Here, the only symmetry we impose is time-reversal.) Note that if $\{A_i\}$ includes all the odd Pauli strings (of which there are $2^{n-1}(2^n - 1)$, which scales exponentially with the number of qubits), then we could create an arbitrary orthogonal transformation in Eq. (9). However, spanning the Hilbert space requires only a subset of the odd Pauli strings, because only $2^n - 1$ real parameters are required to create an arbitrary real state. In particular, we need $\{A_i\}$ to be such that for an arbitrary state $|\psi\rangle$, the states $A_i |\psi\rangle$ form a complete basis. In this case, we refer to $\{A_i\}$ as a complete basis of operators.

The problem then is to determine the minimal pools that produce a complete basis of operators. For a given pool, we define the overlap matrix as $M_{ij} = \langle \psi | A_i^\dagger A_j |\psi\rangle$ where $|\psi\rangle$ is an arbitrary real state. If the rank of $M$ satisfies $\text{rank}(M) \geq 2^n - 1$, this pool is called complete. To determine the smallest complete pools, we randomly generate many different pools of increasing size and compute $\text{rank}(M)$ in each case to test for completeness. We did this for up to 7 qubits. Our numerical investigations reveal that, surprisingly, the minimal pool size required for the overlap matrix to have the required rank of $2^n - 1$ is only $2n - 2$. This is evident in Fig. 4 which shows the fraction of pools that are complete for various pool sizes and numbers of qubits. In each case, complete pools are found for pool sizes that contain at least $2n - 2$ operators. Note that not only is this much smaller than the Hilbert space, it is also much smaller than the size of the fermionic pool, which scales like $n^4$.

To investigate the importance of the pool being complete, we run qubit-ADAPT for random Hamilto-
FIG. 5: Energy error curves for complete pools versus incomplete pools. Results for 3, 4 and 5-qubit random Hamiltonians are shown in panels (a), (b) and (c) respectively.

nians of 3, 4 and 5 qubits with random initial states and for pools consisting of $2n - 2$ operators randomly chosen from the set of odd Pauli strings. The operator coefficients in each Hamiltonian (which is taken to be real and symmetric) are obtained by sampling uniformly in the range $[-2, 2]$ with 20 samples for each of these 3 cases. The corresponding energy error curves are illustrated in Fig. 5. Each curve is the result for a different random Hamiltonian. All of the curves that fail to converge correspond to incomplete pools. For these cases, even though the gradient goes to zero the ground state is not reached because important operators are never generated. On the other hand, the runs with complete pools always converge, highlighting the importance of this criterion. For the cases considered, we find that 20-40% of pools containing $2n - 2$ operators are complete. These findings shed light on the question of what constitutes a good operator pool for ADAPT-VQE. These points were not appreciated in the original paper [30], and moreover, they suggest that the fermionic pool used in that work is overcomplete. Moreover, the fact that the minimal pool size appears to be linearly proportional to $n$ means that the number of additional measurements needed for each step of qubit-ADAPT is also linear in $n$. Thus, the extra measurement overhead of qubit-ADAPT remains modest as the problem size increases.

III. CONCLUSIONS

In conclusion, we introduced a more efficient and NISQ-compatible version of the ADAPT-VQE algorithm, called qubit-ADAPT. The basic idea of qubit-ADAPT is to use a pool consisting of Pauli strings so that the number of CNOT gates associated with each pool operator is reduced. We established a completeness condition that guarantees that a pool will generate an exact ADAPT ansatz, and we found that the smallest possible pool that obeys this criterion appears to scale linearly with the number of qubits. These results lead to a substantial reduction in the depths of state preparation circuits and in the number of measurements needed to run ADAPT-VQE on realistic hardware.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (Award No. 1839136) and the US Department of Energy (Award No. de-sc0019199).

[1] W. Kohn and L. J. Sham, “Self-consistent equations including exchange and correlation effects,” Phys. Rev., vol. 140, pp. A1133–A1138, Nov 1965.

[2] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas,” Phys. Rev., vol. 136, pp. B864–B871, Nov 1964.

[3] R. P. Feynman, “Simulating physics with computers,” International Journal of Theoretical Physics, vol. 21, pp. 467–488, Jun 1982.

[4] A. Y. Kitaev, “Quantum measurements and the Abelian Stabilizer Problem,” arXiv:quant-ph/9511026, Nov. 1995..arXiv: quant-ph/9511026.

[5] D. S. Abrams and S. Lloyd, “Simulation of many-body fermi systems on a universal quantum computer,” Phys. Rev. Lett., vol. 79, pp. 2586–2589, Sep 1997.

[6] D. S. Abrams and S. Lloyd, “Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors,” Phys. Rev. Lett., vol. 83, pp. 5162–5165.
Dec 1999.

[7] J. Preskill, “Quantum Computing in the NISQ era and beyond,” *Quantum*, vol. 2, p. 79, Aug. 2018.

[8] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” *Nature Communications*, vol. 5, Jul 2014.

[9] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, “The theory of variational hybrid quantum-classical algorithms,” *New Journal of Physics*, vol. 18, p. 023023, Feb 2016.

[10] S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin, and X. Yuan, “Quantum computational chemistry,” arXiv:1808.10402 [quant-ph], Aug. 2018. arXiv: 1808.10402.

[11] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, “Computation of molecular spectra on a quantum processor with an error-resilient algorithm,” *Phys. Rev. X*, vol. 8, p. 011021, Feb 2018.

[12] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets,” *Nature*, vol. 549, pp. 242 EP –, Sep 2017.

[13] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, “Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure,” *Phys. Rev. A*, vol. 95, p. 020501, Feb 2017.

[14] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, “Quantum chemistry calculations on a trapped-ion quantum simulator,” *Phys. Rev. X*, vol. 8, p. 031022, Jul 2018.

[15] R. J. Bartlett and M. Musial, “Coupled-cluster theory in quantum chemistry,” *Rev. Mod. Phys.*, vol. 79, pp. 291–352, Feb 2007.

[16] M.-H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, “From transistor to trapped-ion computers for quantum chemistry,” *Scientific Reports*, vol. 4, pp. 3589 EP –, Jan 2014. Article.

[17] I. G. Ryabinkin, T.-C. Yen, S. N. Genin, and A. F. Izmaylov, “Qubit Coupled Cluster Method: A Systematic Approach to Quantum Chemistry on a Quantum Computer,” *J. Chem. Theory Comput.*, vol. 14, pp. 6317–6326, Dec. 2018.

[18] “[1906.11192] Iterative Qubit Cluster Coupled approach with efficient screening of generators.”

https://arxiv.org/abs/1906.11192.

[19] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, “Generalized Unitary Coupled Cluster Wave functions for Quantum Computation,” *J. Chem. Theory Comput.*, vol. 15, pp. 311–324, Jan. 2019.

[20] W. J. Huggins, J. Lee, U. Baek, B. O’Gorman, and K. B. Whaley, “A Non-Orthogonal Variational Quantum Eigensolver,” arXiv:1909.09114 [physics, physics:quant-ph], Sep. 2019.

[21] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, “Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer,” arXiv:1801.01053 [quant-ph], Jan. 2018.

[22] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz,” *Quantum Sci. Technol.*, vol. 4, p. 014008, Oct. 2018.

[23] “[1909.12410] A Jastrow-type decomposition in quantum chemistry for low-depth quantum circuits.”

https://arxiv.org/abs/1909.12410.

[24] F. A. Evangelista, G. Kin-Lic Chan, and G. E. Scuseria, “Exact Parameterization of Fermionic Wave Functions via Unitary Coupled Cluster Theory,” arXiv e-prints, p. arXiv:1910.10130 Oct 2019.

[25] H. R. Grimsley, D. Claudino, S. E. Economou, E. Barnes, and N. J. Mayhall, “Is the Trotterized UCCSD Ansatz usefully well-defined?,” arXiv e-prints, p. arXiv:1910.10329 Oct 2019.

[26] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, “Barren plateaus in quantum neural network training landscapes,” *Nature Communications*, vol. 9, no. 1, p. 4812, 2018.

[27] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Sais, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, S. Filipp, and I. Tavernelli, “Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions,” *Phys. Rev. A*, vol. 98, p. 022322, Aug 2018.

[28] M. Ganzhorn, D. J. Egger, P. K. Barkoutsos, P. Ollitrault, G. Sais, N. Moll, A. Fuhrer, P. Mueller, S. Wernner, I. Tavernelli, and S. Filipp, “Gate-efficient simulation of molecular eigenstates on a quantum computer,” *Physical Review Applied*, vol. 11, p. 044092, Apr. 2019. arXiv: 1809.05057.

[29] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, “Efficient Symmetry-Preserving State Preparation Circuits for the Variational Quantum Eigensolver Algorithm,” arXiv:1904.10910 [quant-ph], Apr. 2019. arXiv: 1904.10910.

[30] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, “An adaptive variational algorithm for exact molecular simulations on a quantum computer,” *Nature Communications*, vol. 10, no. 1, p. 3007, 2019.

[31] Because the gradient depends on the commutator with \( \hat{H} \) such that measurements associated with terms in \( \hat{H} \) which commute with a specific operator in the pool can be skipped, the actual number of measurements needed is difficult to predict, as it will depend on the operator pool chosen. Getting estimates for this number for various pools will be the focus of future work.

[32] “[1907.13117] Efficient and Noise Resilient Measurements for Quantum Chemistry on Near-Term Quantum Computers.”

https://arxiv.org/abs/1907.13117.

[33] W. J. Huggins, J. McClean, N. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush, “Efficient and Noise Resilient Measurements for Quantum Chemistry on Near-Term Quantum Computers,” arXiv:1907.13117 [physics, physics:quant-ph], Sept. 2019.

[34] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz,” *Quantum Science and Technology*, vol. 4, p. 014008, oct 2018.

[35] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. New York, NY, USA: Cambridge University Press, 10th ed., 2011.