JCTC Journal of Chemical Theory and Computation

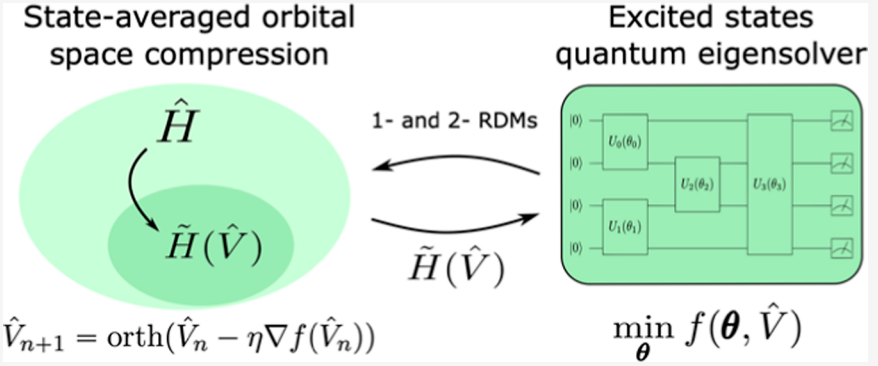
pubs.acs.org/JCTC Article

Qubit Count Reduction by Orthogonally Constrained Orbital Optimization for Variational Quantum Excited-State Solvers

Joel Bierman, Yingzhou Li,* and Jianfeng Lu*

Cite This: J. Chem. Theory Comput. 2024, 20, 3131−3143 Read Online

ACCESS | Metrics & More | Article Recommendations



ABSTRACT: We propose a state-averaged orbital optimization scheme for improving the accuracy of excited states of the electronic structure Hamiltonian for use on near-term quantum computers. Instead of parameterizing the orbital rotation operator in the conventional fashion as an exponential of an antihermitian matrix, we parameterize the orbital rotation as a general partial unitary matrix. Whereas conventional orbital optimization methods minimize the state-averaged energy using successive Newton steps of the second-order Taylor expansion of the energy, the method presented here optimizes the state-averaged energy using an orthogonally constrained gradient projection method that does not require any expansion approximations. Through extensive benchmarking of the method on various small molecular systems, we find that the method is capable of producing more accurate results than fixed basis FCI while simultaneously using fewer qubits. In particular, we show that for H_2, the method is capable of matching the accuracy of FCI in the cc-pVTZ basis (56 qubits) while only using 14 qubits.

1. INTRODUCTION

One of the early applications for quantum computers is expected to be the electronic structure problem;[1] however, the error stemming from the basis set truncation in the second quantization formulation will likely present a major obstacle for realizing accurate solutions for academically and industrially relevant chemical systems.[2−4] If no resource reduction techniques are employed, then one qubit is needed for each spin−orbital. As a result of the limited number of qubits on current hardware, experimental demonstrations have been limited to small molecules represented by small basis sets.[5−8] Several methods have been developed for more compact basis set representations in both classical and quantum settings. Explicitly correlated methods[9−12] apply a similarity transformation to the problem Hamiltonian that has an explicit dependence on the coordinates of the electrons. The intuition behind this is that such basis sets may be able to efficiently capture effects from electron−electron interactions which are often the cause of the inefficiency of fixed single-particle basis set representations.[13,14] Downfolded effective Hamiltonian techniques[15−19] take a full orbital space Hamiltonian, unitarily transform it according to an excitation operator that includes excitations outside of a given active space, and then project it onto the active space. Through this process, an effective Hamiltonian is produced that includes correlation effects from outside the active space but which is low dimensional and acts only on the active space. Orbital optimization methods[20−26] introduce the elements of a similarity transformation to be optimized in conjunction with the parameters of an eigensolver minimization problem. These two minimization problems are typically solved in an alternating fashion, until some stopping criteria are reached. Orbital optimization schemes such as quantum CASSCF[25] operate by representing the orbital rotation as an exponential of an $M \times M$ antihermitian matrix,

Received: November 27, 2023
Revised: March 19, 2024
Accepted: March 20, 2024
Published: April 10, 2024

approximating this exponential operator by a second-order Taylor expansion, performing successive Newton steps using this approximation until convergence of the energy is achieved, then choosing an active space of $N < M$ orbitals in which to solve for the eigenvalues and eigenstates. In this work, we extend the OptOrbVQE[20] method previously proposed by the authors, which finds the ground state in an optimized basis, to the problem of finding excited states of electronic structure Hamiltonians. The main differences between OptOrbVQE and other orbital optimization ground-state solvers can be summarized as follows. Instead of parameterizing the orbital rotation operator as an $M \times M$ exponential operator and carrying out successive Newton steps, we parameterize it directly as an $M \times N$ partial unitary matrix. This allows us to take advantage of modern optimization techniques that have been developed in recent years which have orthogonality constraints built in.[27−31] Such optimization methods render the conventional exponential parameterization as one choice of parameterization rather than a strict requirement. We note that aside from access to a wider range of optimization techniques, this approach comes with other potential advantages. The first is that the partial unitary transformation has NM constrained parameters, rather than the $(M-1)(M-2)/2$ independent parameters in the antihermitian matrix that appears in CASSCF. This results in a net reduction of the size of the parameter space involved for many problems, particularly when $M \gg N$. This may ease the convergence procedure. Additionally, the partial unitary nature of the parameterization implies that unlike CASSCF, we do not need to resort to the use of heuristics or chemical intuition when choosing the postoptimization active space. The dimensionality of the partial unitary matrix handles this automatically during the optimization procedure. We also note that there is a sense in which the dimensionality of this partial unitary operator is a degree of freedom that we can control. For example, we still retain the option to optimize a full $M \times M$ unitary matrix using orthogonally constrained optimizers and use heuristics to choose the active space of N orbitals. Alternatively, we could optimize over the set of $M \times m$ partial unitaries for $N < m < M$ and use heuristics to choose an active space of N orbitals from the space of m orbitals. This flexibility can be seen as an advantage in and of itself. Furthermore, the use of an orthogonally constrained projected gradient descent method[31] in conjunction with this orbital optimization procedure has been numerically demonstrated to be more adept at avoiding local minima and achieving better accuracy than CASSCF in a previous work by two of the authors of this work in a classical computing context.[24] This offers clear motivation for the continued study of this method in a broader range of contexts such as generalization to excited states and incorporation into quantum eigensolvers. For example, in a previous work by the authors, it was demonstrated that the convergence quality of state-averaged eigensolvers such as SSVQE in a fixed basis is highly sensitive to the ansatz expressiveness and choice of circuit initialization (compared to the ground-state VQE problem and the overlap-based qOMM excited-state solver).[32] In this work, we investigate the extent to which analogous observations hold true for the orbital-optimized case as well. Given that there is potentially a nontrivial interplay between the choice of basis (which is furthermore variable in orbital optimization) and the ability of any given ansatz to express the solution, this relation should be investigated.

2. EXCITED-STATE QUANTUM EIGENSOLVERS

Hybrid quantum-classical variational methods for finding eigenvalues of chemical Hamiltonians operate by classically minimizing an objective function constructed from quantities measured on a quantum computer. For example, to find the ground state of a Hamiltonian $\hat{H}$, we would first prepare a parameterized state $|\psi(\boldsymbol{\theta})\rangle$ on the quantum computer, measure the expectation value of $\hat{H}$, and carry out the minimization problem

$$\min_{\boldsymbol{\theta}} \langle \psi(\boldsymbol{\theta}) | \hat{H} | \psi(\boldsymbol{\theta}) \rangle \tag{1}$$

classically. This is the original formulation of the variational quantum eigensolver[33,34] (VQE). In order to extend this method to low-lying excited states, the mutual orthogonality of these states must be accounted for. Several methods have been proposed that accomplish this. SSVQE[35] and MCVQE[36] are state-averaged approaches which apply a parameterized circuit, $\hat{U}(\boldsymbol{\theta})$, to a set of mutually orthogonal initial states $\{|\psi_i\rangle\}$ and then minimize an objective function of the form

$$f(\boldsymbol{\theta}) = \sum_i w_i \langle \psi_i | \hat{U}^\dagger(\boldsymbol{\theta}) \hat{H} \hat{U}(\boldsymbol{\theta}) | \psi_i \rangle \tag{2}$$

where $\{w_i\}$ is a set of positive, real-valued weights. The main difference between MCVQE and SSVQE is that MCVQE chooses the weights $\{w_i\}$ to be equal, whereas SSVQE chooses them not to be equal. At first glance, this difference seems trivial; however, it should be noted that unequal weights correspond to a global minimum composed of the low-lying eigenvectors, whereas an equal weighting corresponds to a global minimum composed of states which span the low-lying eigenspace. MCVQE adds a classical postprocessing step which diagonalizes these states in this low-dimensional eigenspace to acquire the low-lying eigenvectors. It is unclear which of these approaches is advantageous or if their convergence is equivalent in practice. Other excited-state methods such as qOMM[32] and VQD[37] take overlap-based approaches to enforcing the mutual orthogonality of the solution by including penalty terms in the objective function which vanish when pairs of states are orthogonal. Thus, the orthogonality is enforced only at the global minimum rather than at every point in the cost function landscape.

3. STATE-AVERAGED ORBITAL OPTIMIZATION

In OptOrbVQE, we take the electronic structure Hamiltonian in its Fermionic second-quantization representation

$$\hat{H} = \sum_{p,q=1}^{M} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{p,q,r,s=1}^{M} v_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \tag{3}$$

and rotate the set of M orbitals $\{\psi_1, \psi_2, ..., \psi_M\}$ according to the partial unitary transformation $\hat{V}$

$$\tilde{\psi}_i = \sum_j^M \hat{V}_{ji} \psi_j \tag{4}$$

resulting in a new set of $N < M$ orbitals $\{\tilde{\psi}_1, \tilde{\psi}_2, ..., \tilde{\psi}_N\}$. This is equivalent to transforming the Hamiltonian as

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

$$\tilde{H}(\hat{V}) = \sum_{p',q'=1}^{N}\sum_{p,q=1}^{M} h_{pq}\hat{V}_{pp'}\hat{V}_{qq'}\tilde{a}_{p'}^{\dagger}\tilde{a}_{q'} + \frac{1}{2}\sum_{p',q',r',s'=1}^{N}\sum_{p,q,r,s=1}^{M} v_{pqrs}\hat{V}_{pp'}\hat{V}_{qq'}\hat{V}_{ss'}\hat{V}_{rr'}\tilde{a}_{p'}^{\dagger}\tilde{a}_{q'}^{\dagger}\tilde{a}_{s'}\tilde{a}_{r'} \tag{5}$$

The orbital optimization then corresponds to minimizing the expectation value of this Hamiltonian with respect to a fixed quantum state $\hat{U}(\boldsymbol{\theta})|\psi_{\text{ref}}\rangle$ provided by a quantum eigensolver. The total minimization problem is then given by

$$\min_{\hat{V}\in\mathcal{U}(M,N)} \min_{\boldsymbol{\theta}} \langle\psi_{\text{ref}}|\hat{U}^{\dagger}(\boldsymbol{\theta})\tilde{H}(\hat{V})\hat{U}(\boldsymbol{\theta})|\psi_{\text{ref}}\rangle \tag{6}$$

where $\mathcal{U}(M, N)$ is the set of $M \times N$ real partial unitary matrices. The simplest way to generalize this problem is to consider eq 2 to be a function of both $\boldsymbol{\theta}$ and $\hat{V}$

$$f(\boldsymbol{\theta}, \hat{V}) = \sum_{i} w_i\langle\psi_{\text{ref},i}|\hat{U}^{\dagger}(\boldsymbol{\theta})\tilde{H}(\hat{V})\hat{U}(\boldsymbol{\theta})|\psi_{\text{ref},i}\rangle \tag{7}$$

and minimize the resulting state-averaged analogue problem of eq 6

$$\min_{\hat{V}\in\mathcal{U}(M,N)} \min_{\boldsymbol{\theta}} f(\boldsymbol{\theta}, \hat{V}) \tag{8}$$

Such state-averaged analogues of CASSCF-like orbital optimization schemes[21,26] have previously been explored in the literature. Thus, we expect a state-averaged analogue of OptOrbVQE to also perform well. It is worth noting that an overlap-based orbital optimization objective function has been proposed in the classical literature,[38] which allows for a separate optimal basis to be computed for each excited state. The authors claim that this allows for more accurate excitation energies to be computed. The method assumes the availability of the CI coefficients found by the eigensolver, which would require exponentially expensive full-state tomography to acquire in the quantum computing setting.

The total minimization problem in eq 8 is divided into two subproblems: minimization with respect to the ansatz parameters $\boldsymbol{\theta}$ and minimization with respect to $\hat{V}$. These two subproblems are solved in an alternating fashion, where one is fixed, while the other is varied. The optimal parameters for one subproblem are then used for the initialization of the next run of the other until some global stopping criteria are met. For example, for a given optimal $\hat{V}$, we can compute $\tilde{H}(\hat{V})$ and carry out a quantum excited-state solver to find an optimal $\boldsymbol{\theta}$ in the rotated basis. For a given $\boldsymbol{\theta}$ found by a quantum excited-state solver, we can compute the 1- and 2-RDMs with respect to each state in the set of computed excited states $\{\hat{U}(\boldsymbol{\theta})|\psi_{\text{ref},i}\rangle\}$, then vary eq 8 with respect to $\hat{V}$. The optimization with respect to $\boldsymbol{\theta}$ is handled via one of several known quantum excited-state solvers such as SSVQE,[35] MCVQE,[36] VQD,[37] or qOMM.[32] The optimization with respect to $\hat{V}$ (keeping $\boldsymbol{\theta}$ fixed) is carried out using an orthogonally constrained optimization procedure. In this work, we use an orthogonally constrained projected gradient method,[31] which has a parameter update step defined as

$$\hat{V}_{n+1} = \text{orth}(\hat{V}_n - \eta\nabla f(\hat{V}_n)) \tag{9}$$

where $\nabla f(\hat{V}_n))$ is the gradient of eq 7 with respect to $\hat{V}$ with fixed $\boldsymbol{\theta}$, η is a step size which is chosen adaptively in an alternating Barzilai−Borwein fashion, and the orth function is defined as[24,31]

$$\text{orth}(A) = AQ\Lambda^{-1/2}Q^{\dagger} \tag{10}$$

Here, Q is a matrix whose columns are the eigenvectors of $A^{\dagger}A$ and Λ is a diagonal matrix whose entries are the eigenvalues of $A^{\dagger}A$. As was done for OptOrbVQE, we explicitly state the superscript and subscript notation used for the total problem to avoid confusion:

- The subscript l will index the iteration number in the minimization problem where $\hat{V}$ is varied.
- The subscript m will index the iteration number in the minimization problem where $\boldsymbol{\theta}$ is varied.
- The subscript n will index a global "outer loop" iteration number that characterizes how many times both subproblems have been carried out.
- The superscript *opt* will denote the optimal parameter found in each subproblem for a given outer loop iteration number.

We now give an explicit step-by-step procedure for the total problem:

- Set $n = 0$. Choose an initial partial unitary $\hat{V}_{n=0,l=0}$, an initial set of ansatz parameters $\boldsymbol{\theta}_{n=0,m=0}$, and a stopping threshold ϵ.
- Calculate $\tilde{H}(\hat{V})$ on a classical computer and run a quantum eigensolver algorithm to obtain $\boldsymbol{\theta}_n^{\text{opt}}$.
- If $|f(\boldsymbol{\theta}_n^{\text{opt}}, \hat{V}_{n-1}^{opt}) - f(\boldsymbol{\theta}_{n-1}^{\text{opt}}, \hat{V}_{n-2}^{\text{opt}})| < \varepsilon$, halt the algorithm. Else, continue to the next step.
- Measure the 1- and 2-RDMs with respect to the set of states $\{\hat{U}(\boldsymbol{\theta}_n^{\text{opt}})|\psi_{\text{ref},i}\rangle\}$ on a quantum computer.
- Using the 1- and 2-RDMs from the previous step, minimize eq 8 with respect to $\hat{V}_n$ to obtain $\hat{V}_n^{\text{opt}}$.
- Set $n = n + 1$, $\hat{V}_{n+1,l=0} = \hat{V}_n^{\text{opt}}$, and $\boldsymbol{\theta}_{n+1,m=0} = \boldsymbol{\theta}_n^{\text{opt}}$. Optionally, a small random perturbation can be added to the latter two quantities. Repeat steps 2−6.

Step 5 requires the use of a classical optimizer, which constrains $\hat{V}$ to be a partial unitary. Several methods which do this exist,[27−30] but in our work, we use an orthogonal projection method.[31] In general, $\boldsymbol{\theta}$ and $\hat{V}_{n+1,l=0}$ could be any real vector and real partial unitary, respectively; however, it is intuitive to use information from the nth outer loop iteration to inform this choice. In our work, we choose $\boldsymbol{\theta}_{n+1,l=0} = \boldsymbol{\theta}_n^{\text{opt}}$ and $\hat{V}_{n+1,l=0} = \text{orth}(\hat{V}_n^{\text{opt}} + \text{Rand(M, N)})$, where Rand(M, N) is an $M \times N$ matrix whose elements are sampled from a normal distribution with average 0 and standard deviation 0.01. Additionally, although eq 8 is written as a state-averaged function of $\boldsymbol{\theta}$, step 2 does not necessarily need to be carried out using a state-averaged quantum eigensolver. The only requirement is that the solver returns solution states to be used for the calculation of 1- and 2-RDMs. Overlap-based methods such as qOMM[32] and VQD[37] could be used; however, for our work, we test MCVQE[36] and SSVQE.[35] We further note that because the method derives an optimized basis of $N < M$ orbitals from an initial large basis of M orbitals, it cannot match or exceed the accuracy obtained by solving the eigenvalue problem in the full M orbital active space. This is

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

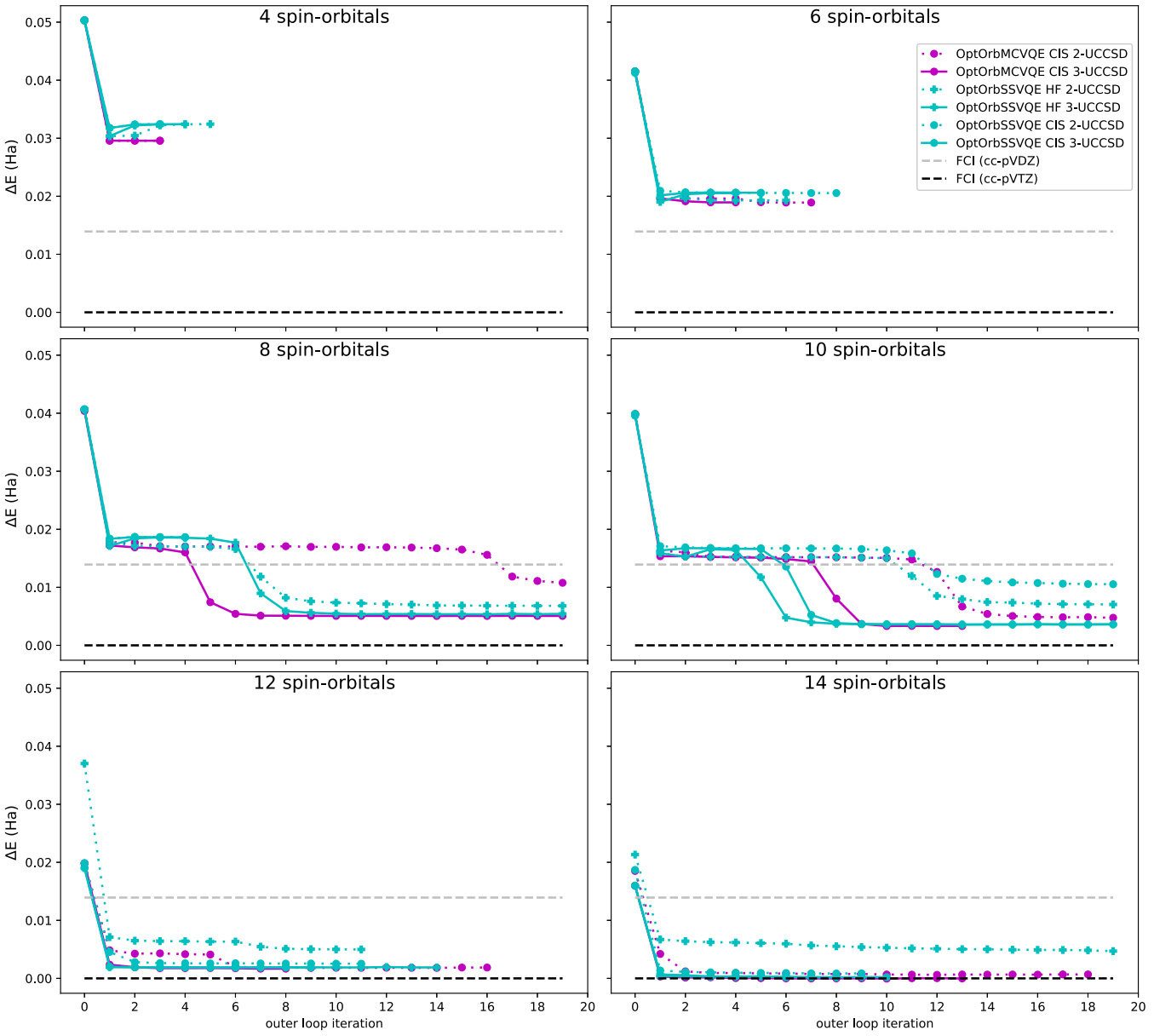


Figure 1. Convergence of orbital optimization methods for H_2 using various numbers of spin−orbitals (taken from the cc-pVQZ basis) as a function of the outer loop iteration. ΔE is the difference between the average energy and that of FCI in the cc-pVTZ basis (56 spin−orbitals).

intuitive from the perspective that the ansatz and orbital rotation parameters constitute a joint variational space. Restricting the variational space will, in general, restrict the maximum attainable accuracy. In particular, it is known that orbital optimization methods, in general, are not efficient at capturing correlation effects arising from Coulomb repulsion between electrons. For this, one could consider combining the orbital optimization scheme with other methods such as explicitly correlated methods.[9−11]

4. NUMERICAL RESULTS

The code used for our numerical simulations is an extension of the functionality provided by the open source package Qiskit.[39] Qiskit provides an implementation of VQE,[33,34] which we have modified to produce implementations of SSVQE[35] and MCVQE.[36] The code for the state-averaged orbital optimization is a modification of the code used in our ground-state orbital optimization work,[20] with the main modification being the objective function to be minimized. The Qiskit package versions used are Qiskit-Aer version 0.12.0, Qiskit-Nature version 0.4.5, and Qiskit-Terra version 0.23.2. The 1- and 2-body integrals are obtained through the PySCF[40] electronic structure driver in Qiskit, which uses PySCF to perform a restricted Hartree−Fock problem to obtain the unoptimized molecular integrals. Configuration interaction circuits are obtained in two steps. First, the truncated Hamiltonians are constructed from the 1- and 2-body integrals using the Slater−Condon rules,[1] which are then exactly diagonalized using NumPy.[41] The reasons why we do not use PySCF's configuration interaction implementation are 2-fold: (1) PySCF does not have an CIS implementation and (2) we have found that PySCF's CISD implementation does not always produce orthogonal CI wave functions, with fidelity between two states being as large as on the order of 10^{-1}, even in the case where the corresponding eigenvalues are not degenerate. This is problematic for quantum algorithms such

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

as SSVQE and MCVQE which require that the initial states be mutually orthogonal.

This statevector can then be used to initialize a circuit using Qiskit's arbitrary statevector initialization implementation. We note that although this particular implementation requires the storage of an exponentially large statevector in classical memory, in principle configuration interaction-state preparation on a quantum computer could be done in a completely sparse manner with resources scaling polynomially with the number of qubits. For example, it has been shown that Givens rotations are universal for preparing chemically motivated states with the Jordan−Wigner mapping.[42] The authors also give a general procedure for preparing an arbitrary statevector. In Appendix B, we give an explicit example of how the particular case of arbitrary CIS statevectors can be prepared on a quantum computer, which may be of independent interest. Whether or not an efficient analogous procedure can be developed for CISD states is not discussed here; however, in our simulations, we include CISD initializations to investigate whether or not doing so would lead to further improvement. We also utilize an "excited Hartree−Fock" initialization that consists of the Hartree−Fock state and the lowest energy singly excited states from it. In this paper, we will refer to this initialization as just "Hartree−Fock" or "HF". In Qiskit, one can use any ansatz circuit as a base pattern to be repeated n times, increasing the circuit depth and number of parameters by a factor of n. In our simulations, we use Qiskit's implementation of the UCCSD ansatz[43] as a circuit block pattern to be repeated for various values of n. We denote this as n-UCCSD. The classical optimizer used for all test instances is L-BFGS-B.[44] The FCI reference values are calculated using CDFCI,[45] and all orbital-optimized tests are run using Qiskit's AerSimulator in noiseless statevector mode unless stated otherwise. In Section 4.5, we make slight adjustments to this methodology when calculating a potential energy surface of H_4.

4.1. H_2. We begin with our results for the simplest model tested, the first three energy levels of H_2 at the near-equilibrium bond distance of 0.735 Å, which are given in Figure 1. We use cc-pVQZ (120 spin−orbitals) as the starting basis and reduce the active space for even numbers of spin−orbitals from 4 to 14 using the proposed orbital optimization scheme. The difference between the average orbital optimized energy and that of FCI (over the ground and excited states) in the cc-pVTZ basis is plotted as a function of the outer loop iteration. Tests using both 2- and 3-UCCSD are included to investigate the effect of increasing the ansatz expressiveness in the algorithm. Both eigensolvers are initialized with configuration interaction singles (CIS) states. SSVQE is additionally tested using the Hartree−Fock initialization.

It is evident that orbital optimization has the potential to achieve more accurate average energies than FCI in the cc-pVDZ basis (20 spin−orbitals) and can even approach cc-pVTZ quality values, but this is highly dependent on the choice of eigensolver, ansatz, and number of optimized spin−orbitals. A minimum of 8 spin−orbitals are needed to achieve a higher accuracy than cc-pVDZ. At this point, OptOrbMCVQE can do this for both 2- and 3-UCCSD, although OptOrbSSVQE cannot. Using 10 spin−orbitals, both eigensolvers surpass cc-pVDZ for both 2- and 3-UCCSD, although MCVQE offers roughly a 5 mHartree improvement over SSVQE for 2-UCCSD. When 3-UCCSD is used, MCVQE offers a measurable but negligible improvement over SSVQE. At 14 spin−orbitals, cc-pVTZ quality results are achievable. Also notable is the effect that increasing the active space has on not only the quality of convergence but its rate of convergence. Note that for 8 and 10 spin−orbitals, the convergence appears to plateau, hovering just above cc-pVDZ accuracy for several iterations before rapidly surpassing it. This behavior is not present at 12 and 14 spin-orbitals, with the energy quickly converging to near or at cc-pVTZ accuracy for the majority of tests run. As a side note, we note that the 14 spin−orbital tests using 3-UCCSD with a CIS initialization were stopped manually at iterations 10 and 13 for SSVQE and MCVQE, respectively, as the runtime for these simulations proved to be the longest among these tests. However, we note that given that nearly all of the energy convergence occurred within the first 2 or 3 iterations, allowing the simulations to continue would likely not have resulted in further improvement.

4.2. H_4. We now present the results for the first three energy levels of H_4, a toy system composed of four hydrogen atoms arranged in a square with a nearest-neighbor distance of 1.23 Å. The starting basis set is cc-pVQZ (240 spin−orbitals) and an active space of 8 optimized spin−orbitals is used. Both 2- and 3-UCCSD are tested as ansatzes and both CIS and CISD are tested as initializations. SSVQE is additionally tested using the Hartree−Fock initialization. The results are listed in Figure 2.

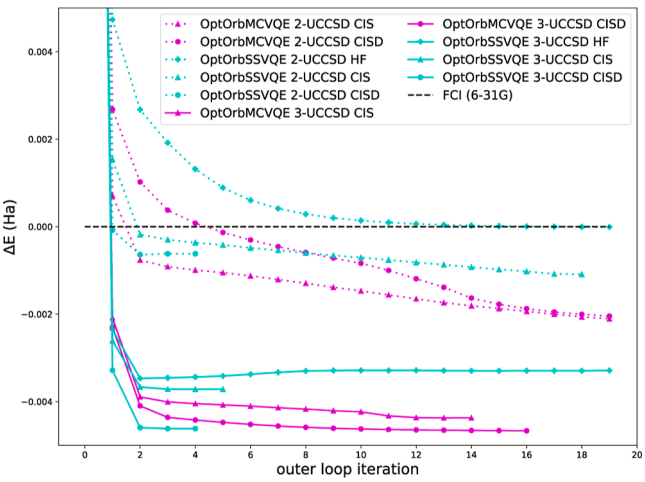


Figure 2. Convergence of orbital optimization methods for H_4 using 8 optimized spin−orbitals (taken from the cc-pVQZ basis) as a function of the outer loop iteration. ΔE is the difference between the average energy and that of FCI in the 6-31G basis (16 spin−orbitals).

We can see that for this system, orbital optimization can be used to achieve a more accurate average energy than the 6-31G basis (16 spin−orbitals), despite the fact that it is utilizing half the number of spin−orbitals. Convergence approaching the FCI cc-pVDZ (40 spin−orbitals) accuracy was not observed in our testing. Between the three different algorithmic choices considered (the eigensolver, the initialization, and the ansatz), increasing the ansatz expressiveness from 2-UCCSD to 3-UCCSD had the most significant effect on the converged accuracy. Changing the initialization from CIS to CISD offered a clear improvement when used with the 3-UCCSD ansatz; however, the same is not true for 2-UCCSD. With 2-UCCSD, OptOrbSSVQE using CISD converges quickly to a local minimum, whereas OptOrbSSVQE using CIS converges (albeit comparatively slowly) to a more accurate average energy. The final converged values for OptOrbMCVQE are

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

similar between CIS and CISD when using 2-UCCSD. Note also that for instances using the same ansatz and initialization, using MCVQE as the eigensolver typically offers an improvement over SSVQE. The one exception to this is using CISD with 3-UCCSD, where the difference between these two converged values is negligible.

4.3. LiH. We now present the results for the first two energy levels of LiH at a near-equilibrium interatomic distance of 1.595 Å. The starting basis set is cc-pVTZ (88 spin−orbitals), and an active space of 12 optimized spin−orbitals is used. Both 1- and 2-UCCSD are used to assess the effect of the ansatz expressiveness. CIS and CISD initializations are tested for MCVQE, whereas SSVQE additionally tests the Hartree−Fock initialization. The results are shown in Figure 3.

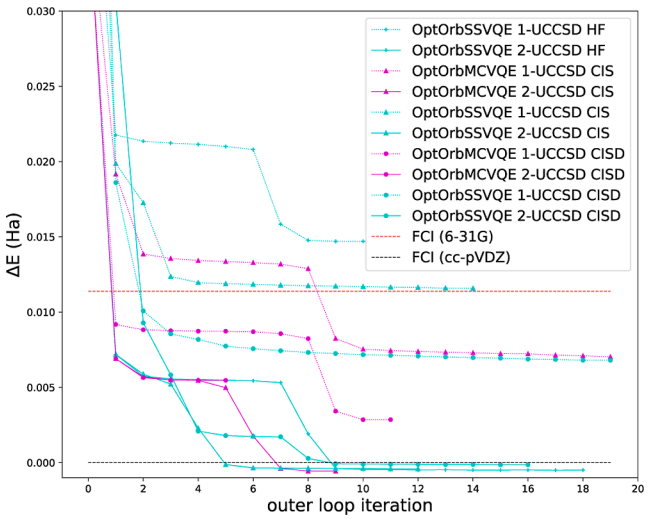


Figure 3. Convergence of orbital optimization methods for LiH using 12 optimized spin−orbitals as a function of the outer loop iteration. ΔE is the difference between the average energy and that of FCI in the cc-pVDZ basis (36 spin−orbitals).

The most notable feature of this plot is that orbital-optimized solvers can achieve more accurate results than FCI using much larger basis sets. For example, most tests run for this system outperform FCI 6-31G (22 spin−orbitals) using only 12 spin−orbitals. Depending on the choice of solver, ansatz, and initialization, some instances also outperform FCI cc-pVDZ (36 spin−orbitals). The second notable feature is that ansatz expressiveness (typically) has a greater influence on the final accuracy than does the choice of initialization. The 1- and 2-UCCSD tests almost form two cleanly separated accuracy tiers, except for OptOrbMCVQE using 1-UCCSD with a CISD initialization, which achieves a higher accuracy than MCVQE using CISD with 2-UCCSD. The choice of initialization has a greater impact when the less expressive 1-UCCSD ansatz is used and has little impact when the ansatz is sufficiently expressive to approximate the solution states well. The third notable feature is that OptOrbMCVQE typically outperforms OptOrbSSVQE when using the same ansatz and initialization, with the one exception being when CISD and 2-UCCSD are used. This effect is most noticeable when the less expressive 1-UCCSD analogue is used.

4.4. BeH_2. We now present the results for the first two energy levels of BeH_2 with a linear geometry at the near-equilibrium Be−H distance of 1.3264 Å. We find the full system with 14 spin−orbitals and 6 electrons to be intractable for our computational budget, so we freeze two electrons in the Hartree−Fock orbitals with the lowest energy and compare the active space energy against that of FCI using the same frozen core approximation. Because we do not wish for the quality of the frozen core approximation across different basis sets to influence the comparison against FCI values, here we will only compare the orbital optimized results starting with the cc-pVQZ basis with an active space of 12 spin−orbitals against FCI in the cc-pVQZ basis using an active space of 228 spin−orbitals. Because of the wide disparity in active space size, we do not expect the orbital optimized tests to approach chemical accuracy compared to FCI; however, we may still gain some insight as to what portion of the full basis set energy is attainable using a small active space and what kind of improvement orbital optimization offers over a naive approach which chooses a fixed active space based on Hartree−Fock orbital energies. These results are given in Figure 4.

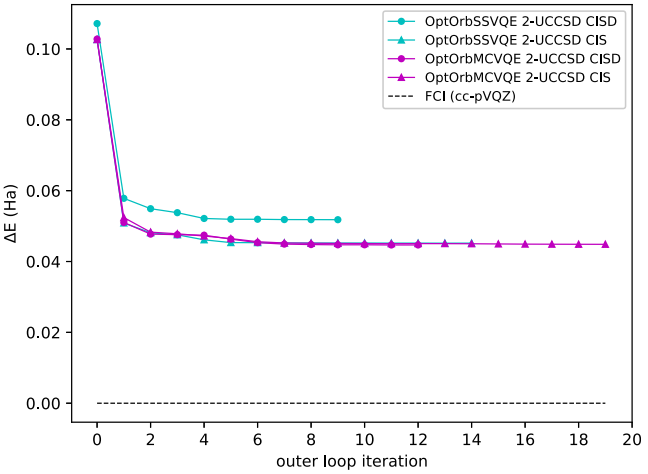
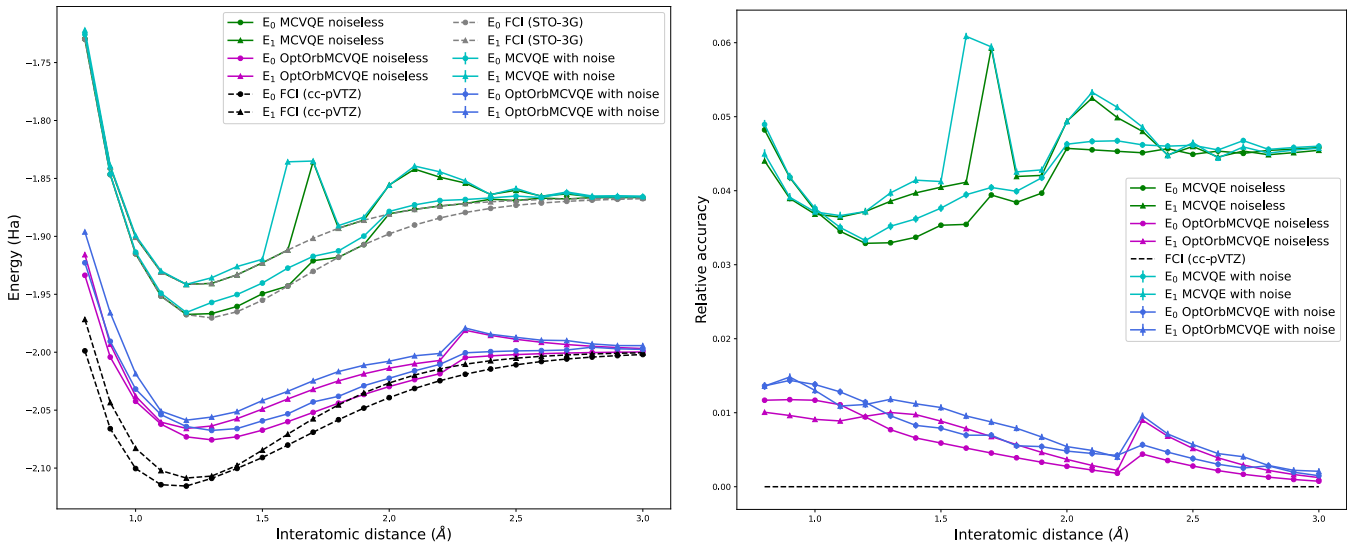


Figure 4. Convergence of orbital optimization methods for BeH_2 using 12 optimized spin−orbitals as a function of the outer loop iteration. ΔE is the difference between the average energy and that of FCI in the cc-pVTZ basis.

4.5. H_4 Noisy Binding Curve. In this section, we use OptOrbMCVQE to compute the potential energy surface resulting from uniformly stretching the nearest-neighbor interatomic distance of the H_4 toy model. We find that this potential energy surface is difficult for state-averaged solvers such as MCVQE to accurately describe in the STO-3G basis, even when starting from CISD states. This is especially true for stretched geometries, where MCVQE often converges to local minima. Thus, this may serve as an interesting test bench for orbital-optimized state-averaged eigensolvers such as OptOrbMCVQE. We study both MCVQE (STO-3G with 8 spin−orbitals) and OptOrbMCVQE (cc-pVTZ with 8 spin−orbitals) in the noiseless case as well as in the case where noise arises from statistical sampling. These results are compared against the exact FCI values in the STO-3G and cc-pVTZ basis sets for comparison. The methodology is largely the same as in previous sections with some notable exceptions. The first is that we now use the COBYLA[46] optimizer for the eigensolver subroutine as we find that it is more robust to noise than L-BFGS-B. Furthermore, we use PySCF's FCI implementation to generate exact comparison values. As opposed to previous sections, where we studied various initializations and ansatz expressiveness, here we restrict ourselves to CISD

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143



(a) Absolute energies (b) Relative energy accuracies

Figure 5. Potential energy surface of the first two energy levels of H_4 corresponding to uniformly stretching nearest-neighbor interatomic distances. (a) Absolute energies for OptOrbMCVQE (cc-pVTZ, 8 spin−orbitals) and MCVQE (STO-3G, 8 spin−orbitals) along with the FCI values in the STO-3G and cc-pVTZ basis sets for comparison. (b) Relative accuracies of E_0 and E_1 for the various methods compared to their exact FCI values in the full cc-pVTZ basis (112 spin−orbitals). Error bars are included but are small.

initializations with the 3-UCCSD ansatz. As we will present the individual energy levels instead of the state-averaged energy, this simplifies the presentation of the data. OptOrbMCVQE uses cc-pVTZ as the starting basis and uses an active space of 8 spin−orbitals. MCVQE uses the 8 spin−orbital STO-3G basis. For the tests which incorporate statistical sampling noise, we use 10^6 shots per observable evaluation using Qiskit's approximate treatment of sampling noise, which assumes a Gaussian distribution for observable measurements and returns the variance σ. We estimate the error associated with these measurements for a number of shots n as $\sqrt{\frac{\sigma}{n}}$. We find that the variance returned by Qiskit is typically on the order of 1×10^{-1}. Figure 5a shows the absolute energies for the binding curve for the first two energy levels of H_4 for OptOrbMCVQE and MCVQE. STO-3G (8 spin−orbitals) and cc-pVTZ (112 spin−orbitals) are included for comparison (Figure 5b). Note that OptOrbMCVQE uses cc-pVTZ as its starting basis but uses an active space of 8 spin−orbitals. Thus, these (along with the other tests in this work) represent an aggressive reduction in the active space size.

We can see that while OptOrbMCVQE offers a significant increase in the absolute energies over MCVQE on the fixed STO-3G basis, its relative accuracy compared to the full cc-pVTZ value is still on the order of 10^{-2} for many geometries. This demonstrates that although orbital optimization offers a compact basis set representation, basis sets beyond the minimal basis will still clearly be required in order to approach chemical accuracy in the infinite basis set limit. Interestingly, we also note that the potential energy surface of the orbital-optimized tests is much more smooth in comparison to the fixed-basis MCVQE tests. MCVQE in the STO-3G basis clearly struggles to reliably converge to the exact values, particularly for interatomic distances of 1.6 Å and larger. The jump in energies around 2.0 Å for OptOrbMCVQE similarly indicate that the method is likely converging to local minima for stretched bond distances, although this effect is qualitatively less severe than in the fixed basis case. We also note that as one would expect, the effect of noise arising from statistical sampling is to degrade the energy accuracies by an amount on the milli-Hartree level. Encouragingly, the majority of the additional accuracy attained through orbital optimization persists despite the presence of this noise.

5. DISCUSSION AND CONCLUSIONS

In this paper, we have proposed an orbital optimization scheme which uses a state-averaged approach to compute excited states of electronic structure Hamiltonians. We have shown that this method can achieve more accurate results than FCI using much larger fixed basis sets. We also investigated the effects of the choice of quantum eigensolver, ansatz expressiveness, and state initialization. While exceptions to these trends can be found in our results, we can make the following general observations:

- Increasing the ansatz expressiveness offers the most significant effect among these factors.
- MCVQE often offers an improvement in accuracy over SSVQE for lower ansatz expressiveness. When higher expressiveness is used, the difference is often less significant.
- CIS initializations often offer an improvement over Hartree−Fock initializations; however, the advantage of using CISD over CIS is unclear. There are several instances of CISD initialized tests achieving a lower accuracy than their CIS (and even Hartree−Fock) counterparts.

The first of these is not surprising. The alternative expressiveness is what primarily determines the variational flexibility of the quantum eigensolver at each outer loop iteration. The second observation can be explained by noting that for a given initialization and ansatz, the solution space of SSVQE is more restricted than that of MCVQE. The solution space of SSVQE with unequal weights consists of the low-lying

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

Table 1. Final Accuracy of the Average Energy for H_4 for Given Choices of Eigensolver, Initialization, and UCCSD Ansatz Expressiveness

		UCCSD repetitions			
eigensolver	initialization	1-rep	2-rep	3-rep	4-rep
MCVQE	CIS	4.25×10^{-2}	3.85×10^{-2}	3.70×10^{-2}	3.70×10^{-2}
MCVQE	CISD	5.16×10^{-3}	2.16×10^{-3}	2.53×10^{-4}	3.73×10^{-8}
SSVQE	HF	5.78×10^{-2}	4.67×10^{-2}	3.70×10^{-2}	3.70×10^{-2}
SSVQE	CIS	4.26×10^{-2}	3.81×10^{-2}	3.70×10^{-2}	3.70×10^{-2}
SSVQE	CISD	5.41×10^{-3}	2.45×10^{-3}	1.56×10^{-4}	3.67×10^{-8}

eigenvectors themselves, whereas the solution space of MCVQE consists of the subspace spanned by the low-lying eigenvectors. MCVQE utilizes a postprocessing step involving a low-dimensional diagonalization problem in this subspace. This additional variational flexibility may ease the convergence process and allow it to partially compensate for an insufficiently expressive ansatz. The third point, while less easily explained than the other two, can be conjectured about. While the initialization does have an effect on ansatz expressiveness as it determines which excitation operators in the UCCSD ansatz act nontrivially, it is not as variationally flexible as the parameterized ansatz itself is. Furthermore, this state is computed in the initial basis set guess, which is usually low quality compared to that of the optimized basis set. Thus, there is no guarantee that CISD computed in this initial basis will continue to be advantageous over CIS for successive basis set rotations. This is less likely to be the case when compared to the Hartree−Fock initial state, which consists of only one Slater determinant and remains the same in the Jordan−Wigner qubit encoding for all basis sets.

One compelling and well-motivated extension of this work would be to take a state-specific orbital optimization approach rather than a state-averaged one. State-specific orbital optimization (as the name implies) optimizes a different basis set for each excited state individually rather than optimizing one basis set for an ensemble of excited states by minimizing its average energy. State-specific orbital optimization has been developed in the context of classical orbital optimization algorithms;[38] however, this particular method relies on a full CI expansion of the wave function at every outer loop iteration. In the quantum computing setting, such an explicit wave function expansion (as opposed to the expectation value of observables used here) would involve exponentially costly tomography and classical storage. These CI wave function expansions are used to compute the overlap of two different excited states in two different basis sets and uses them to enforce their orthogonality. In the quantum computing setting, one would have to develop a method that can compute these overlaps without access to CI expansions or that does not require overlaps at all. This is an interesting problem and will be a direction for future investigation.

APPENDIX A. EXCITED-STATE INITIALIZATIONS AND ANSATZ EXPRESSIVENESS

Here, we test the effects of various initialization choices and levels of ansatz expressiveness on the convergence of MCVQE and SSVQE on a fixed minimal basis. These tests serve to illustrate our motivation for our particular choices in the orbital optimized tests in Section 4 of the main text. By "initialization", we mean the choice of nonparameterized subcircuit prepended to the parameterized ansatz. The ansatz parameters themselves are initialized to zero, as this corresponds to the identity subcircuit. Thus, this allows us to explore various chemically motivated initializations. MCVQE is tested with configuration interaction single (CIS) and configuration interaction single and double (CISD) state initializations. SSVQE is tested with CIS and CISD as well as an "excited Hartree−Fock" initialization used in a previous study by the authors.[32] This initialization applies single-particle Fermionic excitations to the Hartree−Fock state and chooses the resulting Slater determinants with the lowest energy to initialize the circuit. Such states are orthogonal and can thus be used with both MCVQE and SSVQE. The alternative expressiveness is varied by varying the number of times the base UCCSD circuit pattern is repeated, where we denote the circuit consisting of n UCCSD repetitions as n-UCCSD. L-BFGS-B is the optimizer used for these tests.

Table 1 shows the final average energy accuracy for the first three states of H_4 at a nearest-neighbor distance of 1.23 Å for various choices of the eigensolver, state initialization, and UCCSD expressiveness. We can see that Hartree−Fock and CIS initializations fail to produce an accuracy greater than 10^{-2} Ha for any eigensolver or level of ansatz expressiveness. Furthermore, increasing the ansatz expressiveness offers no meaningful improvement for these initializations. On the other hand, the CISD initialization offers the ability to achieve greater than chemical accuracy. With 2-UCCSD, both eigensolvers fall just short of chemical accuracy, but increasing the ansatz to 3- and 4-UCCSD offers further improvements. Thus, we can see that there is motivation for developing circuits that correspond to CISD states.

We now compare the speed of convergence between MCVQE and SSVQE for the four test instances in Figure 6 that were able to surpass chemical accuracy. Figure 6 plots the state-averaged energy accuracy as a function of the number of

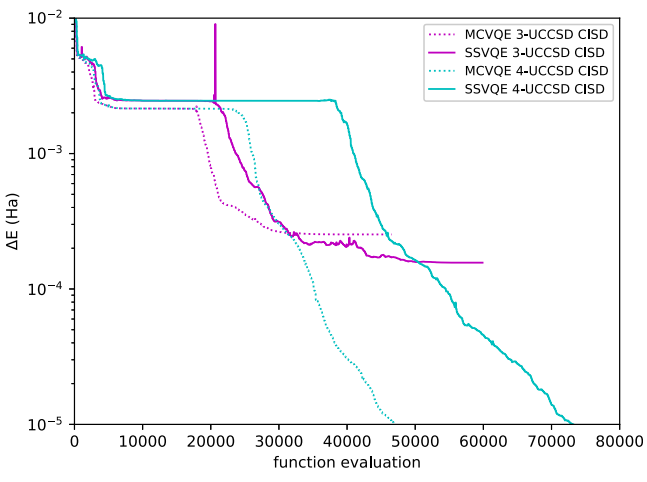


Figure 6. Convergence of the state-averaged energy accuracy (ΔE).

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

objective function evaluations. We can see that for all four instances, the state-averaged energy plateaus for many iterations before escaping and converging to (or closer to) the global minimum. This is consistent with previous studies which include SSVQE by the authors.[32] Notably, MCVQE is less prone to this issue.

APPENDIX B. CIS STATE PREPARATION

Here, we give an example of how configuration interaction single (CIS) states can be prepared as a quantum circuit on a quantum computer by using Givens rotations. It was proven that Givens rotations form a universal set of gates for chemically motivated statevectors.[42] The authors accomplish this constructively by giving a procedure for preparing an arbitrary state using Givens rotations controlled on the states of multiple qubits. They comment that for particular classes of states, the resources involved may be reduced by controlling the rotation only on certain qubit subsets. What remains to be done is to work out the details of how to apply this idea to specific classes of CI statevectors (CIS, CISD, CISDT, etc.) in a way that is as gate efficient as possible. Here, we give an example of how both dense and sparse CIS statevectors can be mapped to quantum circuits using Givens rotations.

We briefly note that the CIS state preparation circuit outlined in the MCVQE proposal paper[36] assumes a particular encoding where the reference state from which electrons are being excited is encoded as the "all-zero" state $|00\cdots0\rangle$ where the qubit registers encode the occupation number of orbitals unoccupied in the reference state but not those occupied in the reference state. Thus, the singly excited wave function components contain no information about the particular Hartree−Fock occupied orbital from which the electron was excited. Here, we seek a CIS state preparation circuit in the Jordan−Wigner encoding where the reference state is the Hartree−Fock state and the occupation number of orbitals occupied in this state are included for all wave function components. Thus, each singly excited wave function component does contain information about the occupied Hartree−Fock orbital from which the electron was excited.

The matrix representation of a Givens rotation involving qubits n and m with angle θ is given by[42]

$$G_{nm}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

where the basis ordering is $|0\rangle_m|0\rangle_n$, $|0\rangle_m|1\rangle_n$, $|1\rangle_m|0\rangle_n$, $|1\rangle_m|1\rangle_n$. or notational convenience, we will often omit the subscript n and m labels on qubit registers. We can also make use of Givens rotations controlled by the state of a target qubit t, which we denote by $C_tG_{nm}(\theta)$. This gate can be represented as

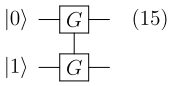
$$C_tG_{nm}(\theta) = |0\rangle\langle 0|_t \otimes \hat{I}_{nm} \tag{12}$$

$$+|1\rangle\langle 1|_t \otimes G_{nm}(\theta) \tag{13}$$

We also note that we adopt the convention of Qiskit where in the Jordan−Wigner encoding the qubits are ordered according to spin and Hartree−Fock energy. Orbitals with the same spin are ordered from right to left in the ascending Hartree−Fock energy. Thus, the relevant action of a Givens rotation is

$$G_{nm}(\theta)|01\rangle = \cos(\theta)|01\rangle + \sin(\theta)|10\rangle \tag{14}$$

We do not have to consider the action of Givens rotations on the state $|10\rangle$ as we are interested in exciting particles to orbitals of higher energies from the lower ones. The circuit notation for the single-excitation Givens rotation is given by[42]



(15)

We want to construct a circuit from Givens rotations which produces the state

$$|\mathrm{CIS}\rangle = C_{\mathrm{HF}}|\mathrm{HF}\rangle + \sum_{p\in O_{\mathrm{HF}}, q\in \mathcal{U}_{\mathrm{HF}}} C_p^q|\phi_q \leftarrow \phi_p\rangle \tag{16}$$

where $|\phi_q \leftarrow \phi_p\rangle$ means the computational basis state produced by exciting an electron from orbital ϕ_p to orbital ϕ_q from the Hartree−Fock ground state. O_{HF} and $\mathcal{U}_{\mathrm{HF}}$ denote the set of orbitals occupied and unoccupied in the Hartree−Fock state, respectively. We can solve for the coefficients C_p^q classically then set them equal to the parameterized coefficients of the wave function expansion produced by a circuit composed of Givens rotations. This produces a set of equations which can be solved to find the Givens angles which produce the circuit that prepares arbitrary CIS states.

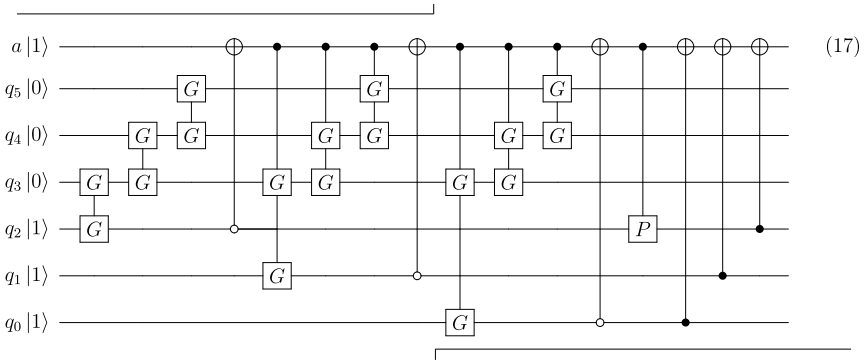
B.1. Example 1: 3 Particles, 6 spin−orbitals

We now give an example for the particular case where we want to generate the CIS wave function with 6 spin−orbitals and 3 particles, where all possible single-particle excitations are considered. The circuit for accomplishing this is given by eq 17where the register labeled a is an ancilla qubit and those labeled q_i are data qubits used to store the CIS state. CNOT gates with an open dot instead of the typical filled dot denote a CNOT gate where the NOT operation is controlled on the target qubit being in the state $|0\rangle$ instead of $|1\rangle$. Although it is not explicitly given in the circuit due to space constraints, each Givens rotation has its own parameter. The controlled phase gate P (implicitly $P(\lambda)$) is given in matrix form by

$$P(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\lambda} \end{pmatrix}$$

where the columns and rows are ordered as $|00\rangle$,$|01\rangle$,$|10\rangle$,$|11\rangle$. We will see later that we need λ to be 0 or π. $\lambda = 0$ corresponds to a 2-qubit identity gate, in which case, we could omit this gate entirely, whereas $\lambda = \pi$ corresponds to a controlled-Z gate. We denote this gate as P in order to keep full generality. The purpose of the final sequence of CNOT gates is to disentangle the ancilla qubit from the data qubits, putting the final state in the form $|\mathrm{CIS}\rangle\otimes|0\rangle$. The final state of the data qubits is given in eq 18

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143



(17)

$$
\begin{aligned}
|\mathrm{CIS}\rangle = {} & e^{i\lambda}\cos\theta_2^3\cos\theta_1^3\cos\theta_0^3|000111\rangle \\
& + \cos\theta_2^3\cos\theta_1^3\sin\theta_0^3\cos\theta_0^4|001110\rangle \\
& + \cos\theta_2^3\cos\theta_1^3\sin\theta_0^3\sin\theta_0^4\cos\theta_0^5|010110\rangle \\
& + \cos\theta_2^3\cos\theta_1^3\sin\theta_0^3\sin\theta_0^4\sin\theta_0^5|100110\rangle \\
& + \cos\theta_2^3\sin\theta_1^3\cos\theta_1^4|001101\rangle \\
& + \cos\theta_2^3\sin\theta_1^3\sin\theta_1^4\cos\theta_1^5|010101\rangle \\
& + \cos\theta_2^3\sin\theta_1^3\sin\theta_1^4\sin\theta_1^5|100101\rangle \\
& + \sin\theta_2^3\cos\theta_2^4|001011\rangle \\
& + \sin\theta_2^3\sin\theta_2^4\cos\theta_2^5|010011\rangle \\
& + \sin\theta_2^3\sin\theta_2^4\sin\theta_2^5|100011\rangle
\end{aligned} \tag{18}
$$

We denote the angle which first adds the component $|\phi_q \leftarrow \phi_p\rangle$ to the overall wave function as θ_p^q. By setting these coefficients equal to those of the CI wave function expansion given in eq 16, we arrive at the following set of equations in eq 19

$$
\begin{aligned}
e^{i\lambda}\cos\theta_2^3\cos\theta_1^3\cos\theta_0^3 &= C_{\mathrm{HF}} \\
\cos\theta_2^3\cos\theta_1^3\sin\theta_0^3\cos\theta_0^4 &= C_0^3 \\
\cos\theta_2^3\cos\theta_1^3\sin\theta_0^3\sin\theta_0^4\cos\theta_0^5 &= C_0^4 \\
\cos\theta_2^3\cos\theta_1^3\sin\theta_0^3\sin\theta_0^4\sin\theta_0^5 &= C_0^5 \\
\cos\theta_2^3\sin\theta_1^3\cos\theta_1^4 &= C_1^3 \\
\cos\theta_2^3\sin\theta_1^3\sin\theta_1^4\cos\theta_1^5 &= C_1^4 \\
\cos\theta_2^3\sin\theta_1^3\sin\theta_1^4\sin\theta_1^5 &= C_1^5 \\
\sin\theta_2^3\cos\theta_2^4 &= C_2^3 \\
\sin\theta_2^3\sin\theta_2^4\cos\theta_2^5 &= C_2^4 \\
\sin\theta_2^3\sin\theta_2^4\sin\theta_2^5 &= C_2^5
\end{aligned} \tag{19}
$$

The recursive structure of this circuit allows us to solve all of these parameters analytically in a recursive way. We can partition these 10 equations into 3 blocks of 3 equations and 1 block with 1 equation according to the occupied Hartree−Fock orbital from which the excitations are generated. We start with $p = 2$ and solve for θ_2^5, θ_2^4, and θ_2^3 in that order. This has the solution

$$
\begin{aligned}
\theta_2^5 &= \arctan\left(\frac{C_2^5}{C_2^4}\right) \\
\theta_2^4 &= \arctan\left(\frac{1}{\cos\theta_2^5}\frac{C_2^4}{C_2^3}\right) \\
\theta_2^3 &= \arcsin\left(\frac{C_2^3}{\cos\theta_2^4}\right)
\end{aligned} \tag{20}
$$

The equations corresponding to $p = 1$ are the same in structure to those of $p = 2$, except that the left-hand side is multiplied by a constant factor of $\cos\theta_2^3$, a quantity that we solved for in the $p = 2$ equations. We define $\alpha_2 = \cos\theta_2^3$ and divide both sides of these equations by α_2. We arrive at a second set of solutions

$$
\begin{aligned}
\theta_1^5 &= \arctan\left(\frac{C_1^5}{C_1^4}\right) \\
\theta_1^4 &= \arctan\left(\frac{1}{\cos(\theta_1^5)}\frac{C_1^4}{C_1^3}\right) \\
\theta_1^3 &= \arcsin\left(\frac{1}{\alpha_2}\frac{C_1^3}{\cos(\theta_1^4)}\right).
\end{aligned} \tag{21}
$$

The $p = 0$ block of equations also has the same form, but the left side is multiplied by a factor of $\alpha_1\alpha_2 = \cos\theta_1^3\cos\theta_2^3$. We divide by sides of each equation in this block by $\alpha_1\alpha_2$ and arrive at the solution for this third block

$$
\begin{aligned}
\theta_0^5 &= \arctan\left(\frac{C_0^5}{C_0^4}\right) \\
\theta_0^4 &= \arctan\left(\frac{1}{\cos\theta_0^5}\frac{C_0^4}{C_0^3}\right) \\
\theta_0^3 &= \arcsin\left(\frac{1}{\alpha_1\alpha_2}\frac{C_0^3}{\cos\theta_0^4}\right)
\end{aligned} \tag{22}
$$

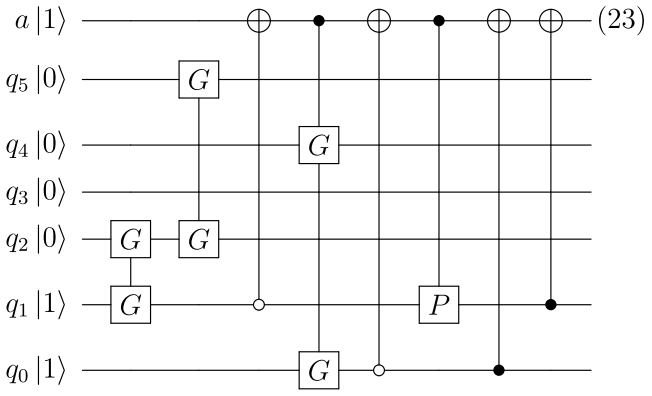
This leaves only the parameter λ to solve. The magnitude of C_{HF} will match that of $\alpha_2\alpha_1\alpha_0$ due to the normalization conditions, but the two may differ by a factor of either +1 or −1. The parameter λ will determine this phase. If the phase of the two quantities matches, then $\lambda = 0$ and the phase gate can be omitted entirely. If the two differ by a phase of −1, then $\lambda = \pi$.

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

B.2. Example 2: Sparse 2 Particles, 6 Spin−Orbitals

The previous example dealt with the particular case of 3 particles and 6 spin−orbitals where every single-particle excitation from any occupied Hartree−Fock orbital is possible. We now give an example for a different number of particles and spin−orbitals for the case where the CIS wave function is sparse and some of the coefficients are zero. This demonstrates that we can also generate approximate CIS wave functions at lower circuit depth in a straightforward, systematic way by omitting certain excitations if their CI coefficients are below a specified threshold.

Here, we suppose that we are preparing a CIS state of a system with 2 particles and 6 spin−orbitals, where ϕ_1 can only be excited to $\{\phi_2, \phi_5\}$ and ϕ_0 can only be excited to ϕ_4. The circuit for doing so is given in eq 23



(23)

This results in the data qubits being put in the state

$$\begin{aligned}|\mathrm{CIS}\rangle &= e^{i\lambda}\cos\theta_1^2\cos\theta_0^4|000011\rangle \\ &+ \cos\theta_1^2\sin\theta_0^4|010010\rangle \\ &+ \sin\theta_1^2\sin\theta_1^5|100001\rangle \\ &+ \sin\theta_1^2\cos\theta_1^5|000101\rangle \end{aligned} \tag{24}$$

This leads to the following set of equations

$$\begin{aligned} e^{i\lambda}\cos\theta_1^2\cos\theta_0^4 &= C_{\mathrm{HF}} \\ \cos\theta_1^2\sin\theta_0^4 &= C_0^4 \\ \sin\theta_1^2\cos\theta_1^5 &= C_1^2 \\ \sin\theta_1^2\sin\theta_1^5 &= C_1^5 \end{aligned} \tag{25}$$

We solve for θ_1^5, θ_1^2, θ_0^4, and λ recursively in that order. The solution is given by

$$\begin{aligned} \theta_1^5 &= \arctan\left(\frac{C_1^5}{C_1^2}\right) \\ \theta_1^2 &= \arcsin\left(\frac{C_1^2}{\cos\theta_1^5}\right) \\ \theta_0^4 &= \arcsin\left(\frac{C_0^4}{\alpha_1}\right) \\ e^{i\lambda} &= \frac{C_{\mathrm{HF}}}{\alpha_1\alpha_0} \end{aligned} \tag{26}$$

B.3. General Procedure

Based on the particular examples given, we can observe a general procedure for any number of particles and spin−orbitals. We first partition the spin−orbitals into two sets O_{HF} and $\mathcal{U}_{\mathrm{HF}}$, the set of spin−orbitals occupied and unoccupied in the Hartree−Fock reference state, respectively. For each $\phi_p \in O_{\mathrm{HF}}$, we generate an ordered set L_p of orbitals $\phi_q \in \mathcal{U}_{\mathrm{HF}}$ for which the CI amplitude C_p^q is not zero or is not below a desired truncation threshold. These orbitals are ordered in ascending Hartree−Fock energy. For every spin−orbital in each L_p, we map the orbital indices q to new indices n_p^q. This is simply so that we may write down a general analytical expression for the gate sequence which reflects the fact we may not want or need the full, dense CI wave function. n_p^q is the index of the list L_p which was mapped from the original index of the spin−orbital ϕ_q, e.g., the original set of unoccupied orbitals from which a particular occupied orbital may be given by $\{\phi_3, \phi_6, \phi_8\}$, but we map this ordered set to the list indices $\{0, 1, 2\}$.

The general sequence of gates is given in eq 27

$$\begin{aligned}|\mathrm{CIS}\rangle = [&\prod_{\phi_p\in O_{\mathrm{HF}}} \mathrm{CNOT}(a,p)]C_aP(\lambda)_0 \\ &\times \prod_{\phi_p\in O_{\mathrm{HF}}\setminus\{\phi_0\}}^{q_{\max}-1} [X_p\mathrm{CNOT}(a,p)X_p \prod_{\phi_q\in L_p} C_aG_{n_p^q,n_p^q+1}] \\ &\times X_0\mathrm{CNOT}(a,0)X_0 \prod_{\phi_q\in L_0}^{q_{\max}-1} G_{n_0^q,n_0^q+1}(\theta_0^q)|\mathrm{HF}\rangle|1\rangle_a \end{aligned} \tag{27}$$

The rightmost terms denote the fact that for the set of excitations from the first occupied spin−orbital, we do not have to apply the Givens rotations conditioned on the state of the ancilla qubit. Without loss of generality, we may take ϕ_0 to be the orbital which has the longest list L_p of possible excited orbitals. This will reduce the circuit depth, as compiling controlled Givens rotations into a sequence of 1- and 2-qubit basis gates will in general be more expensive than regular Givens rotations. After this, we apply a NOT gate to the ancilla qubit conditioned on the state of the qubit from which we just generated excitations being $|0\rangle$. This marks all data qubit wave function components that are not the Hartree−Fock component so that future Givens rotations will not apply excitations to these components. We then repeat this with Givens rotations controlled on the state of the ancilla for all of the other Hartree−Fock occupied orbitals. If there is an orbital for which there are no possible excitations, we simply skip it. We then apply a phase gate $P(\lambda)$ to any of the Hartree−Fock occupied orbitals conditioned on the state of the ancilla. This applies the relative phase $e^{i\lambda}$ to the Hartree−Fock component of the wave function. Finally, for each of the Hartree−Fock occupied orbitals, we apply a NOT gate controlled by the state of the ancilla. This disentangles the data qubits from the ancilla qubit, so that the final result is a product state of these registers.

Finally, we give a general procedure for mapping the CI coefficients C_p^q to Givens rotation angles θ_p^q. In order to do this, we temporarily reindex the CI coefficient indices in the same way that we did for the general circuit expression. For each L_p,

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143

we map the orbital indices $q \to n_p^q$. Here, $C_p^{n_p^q}$ is the reindexed CI coefficient mapped from C_p^q for list L_p. The sequence of steps for this procedure can be given by

(1) For each $\phi_p \in O_{\mathrm{HF}}$ (In the corresponding order applied in the circuit):

(a) If length $(L_p) = 1$

$$\theta_p^{n_p^q} = \arcsin\left(\frac{C_p^{n_p^q}}{\prod_{p'<p}\alpha_{p'}}\right)$$

(b) If length $(L_p) = 2$

$$\theta_p^{n_p^q=1} = \arctan\left(\frac{C_p^{n_p^q=1}}{C_p^{n_p^q=0}}\right)$$

$$\theta_p^{n_p^q=0} = \arcsin\left(\frac{1}{\prod_{p'<p}\alpha_{p'}}\frac{C^{n_p^q=0}}{\cos(\theta_p^{n_p^q=0})}\right)$$

(c) If length $(L_p) > 2$

$$\theta_p^{n_p^{q_{\max}}} = \arctan\left(\frac{C_p^{n_p^{q_{\max}}}}{C_p^{n_p^{q_{\max}}-1}}\right)$$

$$\theta_p^{n_p^{q_{\max}}-1} = \arctan\left(\frac{1}{\cos(\theta_p^{n_p^{q_{\max}}})}\frac{C_p^{n_p^{q_{\max}}-1}}{C_p^{n_p^{q_{\max}}-2}}\right)$$

$$\vdots$$

$$\theta_p^{n_p^q} = \arctan\left(\frac{1}{\cos(\theta_p^{n_p^q+1})}\frac{C_p^{n_p^q}}{C_p^{n_p^q-1}}\right)$$

$$\vdots$$

$$\theta_p^{n_p^q=0} = \arcsin\left(\frac{1}{\prod_{p'<p}\alpha_{p'}}\frac{C^{n_p^q=0}}{\cos(\theta_p^{n_p^q=0})}\right)$$

(2) Solve for λ

$$\lambda = \begin{cases} 0, & \text{if} \dfrac{C_{\mathrm{HF}}}{\prod_p \alpha_p} = 1 \\ \pi, & \text{if} \dfrac{C_{\mathrm{HF}}}{\prod_p \alpha_p} = -1 \end{cases}$$

Here, for the sake of notational convenience, we define $\prod_{p'<0}\alpha_{p'} = 1$.

AUTHOR INFORMATION

Corresponding Authors

Yingzhou Li − *School of Mathematical Sciences, Fudan University, Shanghai 200433, China; Shanghai Key Laboratory for Contemporary Applied Mathematics, Shanghai 200433, China; Key Laboratory of Computational Physical Sciences (MOE), Shanghai 200433, China;* orcid.org/0000-0003-1852-3750; Email: yingzhouli@fudan.edu.cn

Jianfeng Lu − *Department of Mathematics and Department of Chemistry, Duke University, Durham, North Carolina 27708, United States; Department of Physics, Duke University, Durham, North Carolina 27708, United States;* Email: jianfeng@math.duke.edu

Author

Joel Bierman − *Department of Physics, Duke University, Durham, North Carolina 27708, United States;* orcid.org/0000-0002-4311-4548

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.3c01297

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work of J.B. and J.L. are supported by the US National Science Foundation under awards CHE-2037263 and DMS-2309378. Y.L. is supported in part by the National Natural Science Foundation of China (12271109) and the Shanghai Pilot Program for Basic Research—Fudan University 21TQ1400100 (22TQ017).

REFERENCES

(1) Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley, 2014.

(2) Kühn, M.; Zanker, S.; Deglmann, P.; Marthaler, M.; Weiß, H. Accuracy and Resource Estimations for Quantum Chemistry on a Near-Term Quantum Computer. *J. Chem. Theory Comput.* **2019**, *15*, 4764−4780.

(3) Elfving, V. E.; Broer, B. W.; Webber, M.; Gavartin, J.; Halls, M. D.; Lorton, K. P.; Bochevarov, A. How Will Quantum Computers Provide an Industrially Relevant Computational Advantage in Quantum Chemistry?. 2020, arXiv:2009.12472. https://arxiv.org/abs/2009.12472.

(4) Gonthier, J. F.; Radin, M. D.; Buda, C.; Doskocil, E. J.; Abuan, C. M.; Romero, J. Measurements as a Roadblock to Near-Term Practical Quantum Advantage in Chemistry: Resource Analysis. *Phys. Rev. Res.* **2022**, *4*, 033154.

(5) Kandala, A.; Mezzacapo, A.; Temme, K.; Takita, M.; Brink, M.; Chow, J. M.; Gambetta, J. M. Hardware-Efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets. *Nature* **2017**, *549*, 242−246.

(6) Shen, Y.; Zhang, X.; Zhang, S.; Zhang, J.-N.; Yung, M.-H.; Kim, K. Quantum Implementation of the Unitary Coupled Cluster for Simulating Molecular Electronic Structure. *Phys. Rev. A* **2017**, *95*, 020501.

(7) O'Malley, P. J. J.; Babbush, R.; Kivlichan, I. D.; Romero, J.; McClean, J. R.; Barends, R.; Kelly, J.; Roushan, P.; Tranter, A.; Ding, N.; et al. Scalable Quantum Simulation of Molecular Energies. *Phys. Rev. X* **2016**, *6*, 031007.

(8) Ganzhorn, M.; Egger, D.; Barkoutsos, P.; Ollitrault, P.; Salis, G.; Moll, N.; Roth, M.; Fuhrer, A.; Mueller, P.; Woerner, S.; et al. Gate-Efficient Simulation of Molecular Eigenstates on a Quantum Computer. *Phys. Rev. Appl.* **2019**, *11*, 044092.

(9) McArdle, S.; Tew, D. P. Improving the Accuracy of Quantum Computational Chemistry Using the Transcorrelated Method. 2020, arXiv:2006.11181. https://arxiv.org/abs/2006.11181.

(10) Sokolov, I. O.; Dobrautz, W.; Luo, H.; Alavi, A.; Tavernelli, I. Orders of Magnitude Reduction in the Computational Overhead for Quantum Many-Body Problems on Quantum Computers via an Exact Transcorrelated Method. 2023, arXiv:2201.03049. https://arxiv.org/abs/2201.03049.

(11) Motta, M.; Gujarati, T. P.; Rice, J. E.; Kumar, A.; Masteran, C.; Latone, J. A.; Lee, E.; Valeev, E. F.; Takeshita, T. Y. Quantum Simulation of Electronic Structure with a Transcorrelated Hamiltonian: Improved Accuracy with a Smaller Footprint on the Quantum Computer. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24270−24281.

(12) Kumar, A.; Asthana, A.; Masteran, C.; Valeev, E. F.; Zhang, Y.; Cincio, L.; Tretiak, S.; Dub, P. A. Quantum Simulation of Molecular Electronic States with a Transcorrelated Hamiltonian: Higher Accuracy with Fewer Qubits. *J. Chem. Theory Comput.* **2022**, *18*, 5312−5324.
(13) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75−107.
(14) Kutzelnigg, W.; Morgan, J. D. Rates of convergence of the partial-wave expansions of atomic correlation energies. *J. Chem. Phys.* **1992**, *96*, 4484−4508.
(15) Metcalf, M.; Bauman, N. P.; Kowalski, K.; de Jong, W. A. Resource-Efficient Chemistry on Quantum Computers with the Variational Quantum Eigensolver and the Double Unitary Coupled-Cluster Approach. *J. Chem. Theory Comput.* **2020**, *16*, 6165−6175.
(16) Huang, R.; Li, C.; Evangelista, F. A. Leveraging Small Scale Quantum Computers with Unitarily Downfolded Hamiltonians. 2022, arXiv:2208.08591. https://arxiv.org/abs/2208.08591.
(17) Claudino, D.; Peng, B.; Bauman, N. P.; Kowalski, K.; Humble, T. S. Improving the Accuracy and Efficiency of Quantum Connected Moments Expansions. *Quantum Sci. Technol.* **2021**, *6*, 034012.
(18) Bauman, N. P.; Bylaska, E. J.; Krishnamoorthy, S.; Low, G. H.; Wiebe, N.; Granade, C. E.; Roetteler, M.; Troyer, M.; Kowalski, K. Downfolding of Many-Body Hamiltonians Using Active-Space Models: Extension of the Sub-System Embedding Sub-Algebras Approach to Unitary Coupled Cluster Formalisms. *J. Chem. Phys.* **2019**, *151*, 014107.
(19) Bauman, N. P.; Low, G. H.; Kowalski, K. Quantum Simulations of Excited States with Active-Space Downfolded Hamiltonians. *J. Chem. Phys.* **2019**, *151*, 234114.
(20) Bierman, J.; Li, Y.; Lu, J. Improving the Accuracy of Variational Quantum Eigensolvers with Fewer Qubits Using Orbital Optimization. *J. Chem. Theory Comput.* **2023**, *19*, 790−798.
(21) Omiya, K.; Nakagawa, Y. O.; Koh, S.; Mizukami, W.; Gao, Q.; Kobayashi, T. Analytical Energy Gradient for State-Averaged Orbital-Optimized Variational Quantum Eigensolvers and Its Application to a Photochemical Reaction. *J. Chem. Theory Comput.* **2022**, *18*, 741−748.
(22) de Gracia Triviño, J. A.; Delcey, M. G.; Wendin, G. Complete Active Space Methods for NISQ Devices: The Importance of Canonical Orbital Optimization for Accuracy and Noise Resilience. *J. Chem. Theory Comput.* **2023**, *19*, 2863.
(23) Mizukami, W.; Mitarai, K.; Nakagawa, Y. O.; Yamamoto, T.; Yan, T.; Ohnishi, Y.-y. Orbital Optimized Unitary Coupled Cluster Theory for Quantum Computer. *Phys. Rev. Res.* **2020**, *2*, 033421.
(24) Li, Y.; Lu, J. Optimal Orbital Selection for Full Configuration Interaction (OptOrbFCI): Pursuing the Basis Set Limit under a Budget. *J. Chem. Theory Comput.* **2020**, *16*, 6207−6221.
(25) Tilly, J.; Sriluckshmy, P. V.; Patel, A.; Fontana, E.; Rungger, I.; Grant, E.; Anderson, R.; Tennyson, J.; Booth, G. H. Reduced Density Matrix Sampling: Self-consistent Embedding and Multiscale Electronic Structure on Current Generation Quantum Computers. *Phys. Rev. Res.* **2021**, *3*, 033230.
(26) Yalouz, S.; Senjean, B.; Günther, J.; Buda, F.; O'Brien, T. E.; Visscher, L. A State-Averaged Orbital-Optimized Hybrid Quantum−Classical Algorithm for a Democratic Description of Ground and Excited States. *Quantum Sci. Technol.* **2021**, *6*, 024004.
(27) Wen, Z.; Yin, W. A Feasible Method for Optimization with Orthogonality Constraints. *Math. Program.* **2013**, *142*, 397−434.
(28) Gao, B.; Liu, X.; Yuan, Y.-x. Parallelizable Algorithms for Optimization Problems with Orthogonality Constraints. *SIAM J. Sci. Comput.* **2019**, *41*, A1949−A1983.
(29) Zhang, X.; Zhu, J.; Wen, Z.; Zhou, A. Gradient Type Optimization Methods For Electronic Structure Calculations. *SIAM J. Sci. Comput.* **2014**, *36*, C265−C289.
(30) Huang, W.; Gallivan, K. A.; Absil, P.-A. A Broyden Class of Quasi-Newton Methods for Riemannian Optimization. *SIAM J. Optim.* **2015**, *25*, 1660−1685.
(31) Gao, B.; Liu, X.; Chen, X.; Yuan, Y.-x. A New First-Order Algorithmic Framework for Optimization Problems with Orthogonality Constraints. *SIAM J. Optim.* **2018**, *28*, 302−332.
(32) Bierman, J.; Li, Y.; Lu, J. Quantum Orbital Minimization Method for Excited States Calculation on a Quantum Computer. *J. Chem. Theory Comput.* **2022**, *18*, 4674−4689.
(33) Peruzzo, A.; McClean, J.; Shadbolt, P.; Yung, M.-H.; Zhou, X.-Q.; Love, P. J.; Aspuru-Guzik, A.; O'Brien, J. L. A Variational Eigenvalue Solver on a Photonic Quantum Processor. *Nat. Commun.* **2014**, *5*, 4213.
(34) Tilly, J.; Chen, H.; Cao, S.; Picozzi, D.; Setia, K.; Li, Y.; Grant, E.; Wossnig, L.; Rungger, I.; Booth, G. H.; et al. The Variational Quantum Eigensolver: A Review of Methods and Best Practices. *Phys. Rep.* **2022**, *986*, 1−128.
(35) Nakanishi, K. M.; Mitarai, K.; Fujii, K. Subspace-Search Variational Quantum Eigensolver for Excited States. *Phys. Rev. Res.* **2019**, *1*, 033062.
(36) Parrish, R. M.; Hohenstein, E. G.; McMahon, P. L.; Martínez, T. J. Quantum Computation of Electronic Transitions Using a Variational Quantum Eigensolver. *Phys. Rev. Lett.* **2019**, *122*, 230401.
(37) Higgott, O.; Wang, D.; Brierley, S. Variational Quantum Computation of Excited States. *Quantum* **2019**, *3*, 156.
(38) Yalouz, S.; Robert, V. Orthogonally Constrained Orbital Optimization: Assessing Changes of Optimal Orbitals for Orthogonal Multireference States. *J. Chem. Theory Comput.* **2023**, *19*, 1388−1392.
(39) Qiskit Contributors. *Qiskit: An Open-Source Framework for Quantum Computing*. 2023.
(40) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; et al. PySCF: The Python-based Simulations of Chemistry Framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, No. e1340.
(41) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; et al. Array Programming with NumPy. *Nature* **2020**, *585*, 357−362.
(42) Arrazola, J. M.; Di Matteo, O.; Quesada, N.; Jahangiri, S.; Delgado, A.; Killoran, N. Universal Quantum Circuits for Quantum Chemistry. *Quantum* **2022**, *6*, 742.
(43) Romero, J.; Babbush, R.; McClean, J. R.; Hempel, C.; Love, P. J.; Aspuru-Guzik, A. Strategies for Quantum Computing Molecular Energies Using the Unitary Coupled Cluster Ansatz. *Quantum Sci. Technol.* **2018**, *4*, 014008.
(44) Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J. Sci. Comput.* **1995**, *16*, 1190−1208.
(45) Wang, Z.; Li, Y.; Lu, J. Coordinate Descent Full Configuration Interaction. *J. Chem. Theory Comput.* **2019**, *15*, 3558−3569.
(46) Powell, M. J. D. Direct Search Algorithms for Optimization Calculations. *Acta Numer.* **1998**, *7*, 287−336.

https://doi.org/10.1021/acs.jctc.3c01297
J. Chem. Theory Comput. 2024, 20, 3131−3143