

A Quantum-Inspired Algorithm for the Factorized Form of Unitary Coupled Cluster Theory

Jia Chen,^{*,†,‡} Hai-Ping Cheng,^{*,†,‡} and James Freericks^{*,¶}

[†]*Department of Physics, University of Florida, Gainesville, FL 32611, USA*

[‡]*Quantum Theory Project, University of Florida, Gainesville, FL 32611, USA*

[¶]*Department of Physics, Georgetown University, 37th St. and O St., NW, Washington, DC 20057, USA*

E-mail: jiachen@ufl.edu; hping@ufl.edu; James.Freericks@georgetown.edu

Abstract

The factorized form of unitary coupled cluster theory (UCC) is a promising wave-function ansatz for the variational quantum eigensolver algorithm. Here, we present a quantum-inspired classical algorithm for UCC based on an exact operator identity for the individual UCC factors. We implement this algorithm for calculations of the H_{10} linear chain and the H_2O molecule with single and double ζ basis sets to provide insights about UCC as a wave-function ansatz. We find that for weakly-correlated molecules, the factorized form of the UCC provides similar accuracy to conventional coupled cluster theory (CC); For strongly-correlated molecules, where CC often breaks down, UCC significantly outperforms the configuration interaction (CI) ansatz. As a result, the factorized form of the UCC is an accurate, efficient and reliable electronic structure method in both the weakly and strongly correlated regions. This classical algorithm now allows robust benchmarking of anticipated results from quantum computers and application of coupled-cluster techniques to more strongly correlated molecules.

1 Introduction

Unitary coupled cluster theory (UCC) was proposed as a wave-function ansatz for quantum chemistry about four decades ago.¹⁻³ In spite of efforts to develop computational methods for this ansatz,⁴⁻⁷ UCC is not nearly as prevalent as the closely related coupled cluster theory (CC)^{8,9} in electronic structure calculations. One major reason behind this lack of usage is because the Baker-Campbell-Hausdorff expansion of the similarity-transformed Hamiltonian does not terminate in UCC (unlike what happens in CC), which often makes an exact calculation of this theory intractable on classical computers.

Quantum computing provides new ways to tackle the many-electron problem.¹⁰ For example, the quantum phase-estimation algorithm is capable of calculating the ground-state energy in polynomial time (provided a sufficiently good initial wavefunction is prepared, and time evolution is feasible for the Hamiltonian). Unfortunately, limitations on the maximal circuit depth that can be run on near-term quantum computers make the phase estimation algorithm unrealistic at this time. The variational quantum eigenvalue solver (VQE),¹¹ which is an approximate quantum-classical hybrid approach, has shown a lot of promise for the electronic structure problem in the near future. In VQE, a quantum circuit prepares a wavefunction that depends on a set of variational parameters, and multiple measurement circuits are employed to then evaluate the expectation value of the energy (multiple circuits are needed because the Hamiltonian is broken up into a sum of unitary pieces to carry out the computation on a quantum computer). Optimization of the energy with respect to the variational parameters is then carried out in concert with a classical computer. Since most operations on a quantum computer are unitary, UCC has been proposed as a low-circuit-depth state-preparation ansatz for VQE.¹²

This UCC-based state-preparation strategy has been successfully implemented on an ion-trap quantum computer^{13,14} for H₂O using a minimal basis and up to three UCC factors. This success further motivated research to improve both the UCC ansatz and the VQE algorithm; examples include, (i) the Bogoliubov transformed UCC¹⁵ and (ii) the k-UpCCGSD

approach¹⁶ (which utilizes generalized single and double excitations). An adaptive algorithm (ADAPT) has also been developed,¹⁷ which is able to adjust the ansatz by selecting the ordering of the most important operators to use.

Due to the current paucity of robust quantum hardware, most UCC calculations are performed on classical computers. One way to do this is to directly simulate the quantum circuits on classical computers. This approach is limited, because quantum circuits with more than 50 qubits cannot be efficiently simulated. The other method available requires a numerical computation of the matrix exponential (via a truncated power series, rescaling and squaring, or diagonalization). Since the dimension of the matrices is equal to the dimensionality of the Hilbert space, only small systems can be studied this way. As a result, the molecules that have been examined contain only a handful of atoms, and usually are represented by a minimal basis set. How the UCC will work for larger molecules and basis sets is critical to understanding how effective the VQE will be in advancing quantum chemistry on a quantum computer. Our work now allows us to do this using classical computation.

In this paper, we introduce a quantum-inspired algorithm for a factorized form of UCC that is based on an operator identity that recognizes a hidden $SU(2)$ symmetry.^{18,19} To illustrate how this approach work, we perform numerical calculations for the H_{10} linear chain and for the H_2O molecule with the minimal and the double- ζ basis set. Analysis of these results sheds light onto the finer points of how one can employ the UCC on quantum computers.

2 Theory and Method

In UCC, the trial wave-function is expressed in an exponential form, given by

$$|\Psi_{\text{UCC}}\rangle = \exp(\hat{\sigma})|\Psi_0\rangle, \tag{1}$$

where $|\Psi_0\rangle$ is a single reference state and the operator $\hat{\sigma}$ is an anti-Hermitian combination of particle-hole excitation and de-excitation:

$$\hat{\sigma} = \hat{T} - \hat{T}^\dagger; \quad (2)$$

$$\hat{T} = \sum_i^{\text{occ}} \sum_a^{\text{vir}} \theta_i^a \hat{a}_a^\dagger \hat{a}_i + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} \theta_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i + \cdots, \quad (3)$$

where the angles θ are the variational parameters. We use letters from the start of the alphabet $a, b, c, \dots$ to denote the virtual orbitals, with respect to the reference state, and letters from the middle of the alphabet $i, j, k, \dots$ to denote the occupied orbitals in the reference state. To simply notation, we express a general n -fold excitation operator as $\hat{a}_{ij\dots}^{ab\dots} = \hat{a}_a^\dagger \hat{a}_b^\dagger \dots \hat{a}_j \hat{a}_i$ (with the corresponding de-excitation operator being its Hermitian conjugate).

We work in a factorized form for the UCC, which is given by

$$|\Psi_{\text{UCC}}\rangle = \prod_{ij\dots}^{\text{occ}} \prod_{ab\dots}^{\text{vir}} \exp[\theta_{ij\dots}^{ab\dots} (\hat{a}_{ij\dots}^{ab\dots} - \hat{a}_{ab\dots}^{ij\dots})] |\Psi_0\rangle, \quad (4)$$

since this is the form of the UCC operator that can be easily generated on quantum computers. Note that this form is completely general, because we did not specify at all what the strategy is for determining the different factors. In particular, we can express the traditional UCC form (with all excitations appearing as a sum in the exponent) in this form, simply by using a Trotter breakup, which entails repeating many of the same factors in the expansion.

Next, we discuss the operator identity for an arbitrary UCC factor appearing in Eq. (4) and use it to develop a classical algorithm for the UCC.

2.1 Operator Identity For UCC Factors

We first examine UCC factors that correspond to single excitations:

$$\hat{U}_i^a = \exp[\theta_i^a (\hat{a}_i^a - \hat{a}_a^i)]. \quad (5)$$

This exponential operator can be reframed into the sum of a much simpler operator expression by using a hidden $SU(2)$ group structure associated with these UCC factors. First define the “hidden” spin operators as

$$\hat{S}_+ = i\hat{a}_a^\dagger \hat{a}_i; \quad \hat{S}_- = (\hat{S}_+)^\dagger = -i\hat{a}_i^\dagger \hat{a}_a; \quad \hat{S}_z = \frac{1}{2}[\hat{S}_+, \hat{S}_-] = \frac{1}{2}(\hat{a}_a^\dagger \hat{a}_a - \hat{a}_i^\dagger \hat{a}_i). \quad (6)$$

One can see that this is the conventional fermionic representation of spin, if we think of the virtual spin-orbital a as corresponding to spin-up and the real spin-orbital i as corresponding to spin down. The commutation relations of these operators can then be immediately determined to be

$$[\hat{S}_+, \hat{S}_-] = 2\hat{S}_z, \quad [\hat{S}_z, \hat{S}_+] = \hat{S}_+, \quad \text{and} \quad [\hat{S}_z, \hat{S}_-] = -\hat{S}_-, \quad (7)$$

which can be recognized as the conventional $SU(2)$ algebra. But these operators are not an independent $SU(2)$ algebra, instead they arise as a subgroup of the permutation symmetry of all the generators of the allowed UCC factors (in factorized form). Hence, they represent a direct sum of $S = 0$ and $S = 1/2$ representations when acting on any product state in the Hilbert space. We see this when we examine some additional operator identities given by

$$\begin{aligned} \hat{S}_+^2 &= \hat{S}_-^2 = 0; \quad (\hat{S}_+ + \hat{S}_-)^2 = \hat{S}_+ \hat{S}_- + \hat{S}_- \hat{S}_+; \\ (\hat{S}_+ + \hat{S}_-)^3 &= \hat{S}_+ \hat{S}_- \hat{S}_+ + \hat{S}_- \hat{S}_+ \hat{S}_- = 2\hat{S}_z \hat{S}_+ - 2\hat{S}_z \hat{S}_- = \hat{S}_+ + \hat{S}_-, \end{aligned} \quad (8)$$

which are not general operator identities of the $SU(2)$ algebra, but are specific to this direct-sum space. These identities immediately imply that odd powers of $\hat{S}_+ + \hat{S}_-$ are equal to $\hat{S}_+ + \hat{S}_-$ and even powers are equal to $\hat{S}_+ \hat{S}_- + \hat{S}_- \hat{S}_+$. These identities are similar to a spin-1 representation, where the cube of the Cartesian angular momentum operators are equal to the Cartesian angular momentum operators. Indeed, this same identity allows us to evaluate the exponentials exactly. Simply expand the exponential of the corresponding UCC factor in

a power series and use the fact that all nonzero even powers of the operator are the same and all odd powers of the operator are the same. Then the numerical factors can be immediately resummed to yield

$$\begin{aligned}\exp[-i\theta(\hat{S}_+ + \hat{S}_-)] &= 1 - i\theta(\hat{S}_+ + \hat{S}_-) + \frac{(i\theta)^2}{2}(\hat{S}_+ + \hat{S}_-)^2 - \frac{(i\theta)^3}{3!}(\hat{S}_+ + \hat{S}_-)^3 + \dots \\ &= 1 - i\sin\theta(\hat{S}_+ + \hat{S}_-) + (\cos\theta - 1)(\hat{S}_+\hat{S}_- + \hat{S}_-\hat{S}_+).\end{aligned}\quad (9)$$

For this single UCC excitation, inserting Eq. (6) into Eq. (9) yields:

$$\exp[\theta(\hat{a}_i^a - \hat{a}_a^i)] = 1 + \sin\theta(\hat{a}_i^a - \hat{a}_a^i) + (\cos\theta - 1)(\hat{n}_a + \hat{n}_i - 2\hat{n}_a\hat{n}_i), \quad (10)$$

where $\hat{n}_\alpha = \hat{a}_\alpha^\dagger \hat{a}_\alpha$ is the density operator for $\alpha = a$ or i . The result of applying Eq. (10) onto certain states can be found in Ref. 18. More importantly, this operator identity can be generalized to any UCC factor of arbitrary n -particle excitations via the recognition of the hidden $SU(2)$ algebra for the general case, which is given by

$$\begin{aligned}\hat{S}_+ &= i\hat{a}_{i_1 i_2 \dots i_n}^{a_1 a_2 \dots a_n}, \quad \hat{S}_- = -i\hat{a}_{a_1 a_2 \dots a_n}^{i_1 i_2 \dots i_n}, \quad \text{and} \\ \hat{S}_z &= \frac{1}{2}(\hat{n}_{a_1} \dots \hat{n}_{a_n}(1 - \hat{n}_{i_1}) \dots (1 - \hat{n}_{i_n}) - (1 - \hat{n}_{a_1}) \dots (1 - \hat{n}_{a_n})\hat{n}_{i_1} \dots \hat{n}_{i_n}).\end{aligned}\quad (11)$$

These spin operators satisfy the $SU(2)$ commutation relations in Eq. (7) and also satisfy the additional operator identities in Eq. (8). Hence, we can perform the exact same expansion of the power series for this UCC factor and find the same result as given in Eq. (9). Evaluating the spin operators in terms of the excitation and de-excitation operators then yields the final exact operator identity:

$$\begin{aligned}U_{i_1 \dots i_n}^{a_1 \dots a_n} &= \exp[\theta(\hat{a}_{i_1 \dots i_n}^{a_1 \dots a_n} - \hat{a}_{a_1 \dots a_n}^{i_1 \dots i_n})] = 1 + \sin\theta(\hat{a}_{i_1 \dots i_n}^{a_1 \dots a_n} - \hat{a}_{a_1 \dots a_n}^{i_1 \dots i_n}) \\ &+ (\cos\theta - 1)[\hat{n}_{a_1} \dots \hat{n}_{a_n}(1 - \hat{n}_{i_1}) \dots (1 - \hat{n}_{i_n}) + (1 - \hat{n}_{a_1}) \dots (1 - \hat{n}_{a_n})\hat{n}_{i_1} \dots \hat{n}_{i_n}].\end{aligned}\quad (12)$$

This identity acts in a direct sum space of $S = 0$ and $S = 1/2$: when $S = 0$, which happens when no excitation or de-excitation is possible, the operator acts as the identity, but when $S = 1/2$, which happens when an excitation or de-excitation is possible, the operator acts analogous to a spin-one-half spinor, which is rotated by the angle θ .

2.2 The Quantum-Inspired Algorithm

Based on the exact operator identity derived above, we devise an algorithm inspired by the VQE¹¹ for UCC in a factorized form; this algorithm can be carried out on classical computers. Eq. (12) guarantees that applying the UCC factor $U_{i_1 \dots i_n}^{a_1 \dots a_n}$ to the configuration $|\Psi_{i_1' \dots i_{n'}}^{a_1' \dots a_{n'}}\rangle$, where $|\Psi_{i_1' \dots i_{n'}}^{a_1' \dots a_{n'}}\rangle$ is the result of applying $\hat{a}_{i_1' \dots i_{n'}}^{a_1' \dots a_{n'}}$ to reference state $|\Psi_0\rangle$, can only have three outcomes:

1. ($S = 1/2$ case, excitation) If sets $\{a_1 \dots a_n\}$ and $\{a_{1'} \dots a_{n'}\}$ have no common elements; and sets $\{i_1 \dots i_n\}$ and $\{i_{1'} \dots i_{n'}\}$ have no common elements, then applying the UCC factor generates the sum of two configurations—the original configuration $|\Psi_{i_1' \dots i_{n'}}^{a_1' \dots a_{n'}}\rangle$ with coefficient $\cos \theta$ and the excitation $|\Psi_{i_1' \dots i_{n'} i_1 \dots i_n}^{a_1' \dots a_{n'} a_1 \dots a_n}\rangle$ with coefficient $\pm \sin \theta$. Because excitations in the wavefunction can arise from multiple pathways, we must adopt a consistent ordering scheme for the fermionic operators that determine each determinant in the variational ansatz. We follow the convention used by Handy²⁰ to keep track of the signs; In some cases, there is an overall negative sign for a specific determinant that produces a $-\sin \theta$.
2. ($S = 1/2$ case, de-excitation) If set $\{a_1 \dots a_n\}$ is subset of $\{a_{1'} \dots a_{n'}\}$ with complement set $\{a_{1''} \dots a_{n''}\} = \{x \in \{a_{1'} \dots a_{n'}\} | x \notin \{a_1 \dots a_n\}\}$; and set $\{i_1 \dots i_n\}$ is a subset of $\{i_{1'} \dots i_{n'}\}$ with complement set $\{i_{1''} \dots i_{n''}\} = \{x \in \{i_{1'} \dots i_{n'}\} | x \notin \{i_1 \dots i_n\}\}$, then applying UCC factors generates two configurations—the original $|\Psi_{i_1' \dots i_{n'}}^{a_1' \dots a_{n'}}\rangle$ with coefficient $\cos \theta$ and the de-excitation $|\Psi_{i_{1''} \dots i_{n''}}^{a_{1''} \dots a_{n''}}\rangle$ with coefficient $-\sin \theta$. Similar to the excitation case, a sign (+1 or -1) needs to be determined for each the de-excitation

determinant.

3. ($S = 0$ case, nothing) If the above two conditions are not satisfied, then the result is one configuration, the original $|\Psi_{i_1, \dots, i_{n'}}^{a_1, \dots, a_{n'}}\rangle$ with coefficient 1.

Equipped with these exact rules, we can efficiently produce UCC wavefunctions (in the factorized form) on classical computers. Configurations generated by sequentially applying UCC factors fit into a tree-type data structure. The root node of the tree is the initial reference configuration. Each level except the last one corresponds to the application of one of the UCC factors in sequential order. Each parent node gives rise to one or two child nodes according to Eq. (12) and the three rules stated above. Starting from the root node with coefficient 1, we can calculate the coefficients for all the nodes on the tree according to the exact rules. The last level of the tree (leaf nodes) is the final UCC wave-function expressed as a linear combination of configurations. It is important to condense all final leaves (those indicated by the interrupted line pairs in the last row of the figure) to save space during the tree construction. One example of this tree structure built on three UCC factors can be found in Fig. 1.

This tree structure can not only be used to generate the UCC wave-function, but also the derivatives of the wavefunction with respect to the variational parameters, which is extremely useful in optimizing the variational ansatz. We can take the derivative of Eq. (12):

$$\begin{aligned} \frac{d\hat{U}_{i_1 \dots i_n}^{a_1 \dots a_n}(\theta)}{d\theta} &= \cos \theta (\hat{a}_{i_1 \dots i_n}^{a_1 \dots a_n} - \hat{a}_{a_1 \dots a_n}^{i_1 \dots i_n}) \\ &\quad - \sin \theta [\hat{n}_{a_1} \dots \hat{n}_{a_n} (1 - \hat{n}_{i_1}) \dots (1 - \hat{n}_{i_n}) + (1 - \hat{n}_{a_1}) \dots (1 - \hat{n}_{a_n}) \hat{n}_{i_1} \dots \hat{n}_{i_n}]. \end{aligned} \quad (13)$$

The result of applying this operator to a configuration is similar to what has been discussed above, but with the coefficients modified. The derivative of the wavefunction with respect to each variational parameter immediately follows as

$$\frac{d|\Psi_{\text{UCC}}\rangle}{d\theta_i} = \hat{U}_n \dots \hat{U}_{i+1} \frac{d\hat{U}_i}{d\theta_i} \dots \hat{U}_2 \hat{U}_1 |\Psi_0\rangle. \quad (14)$$

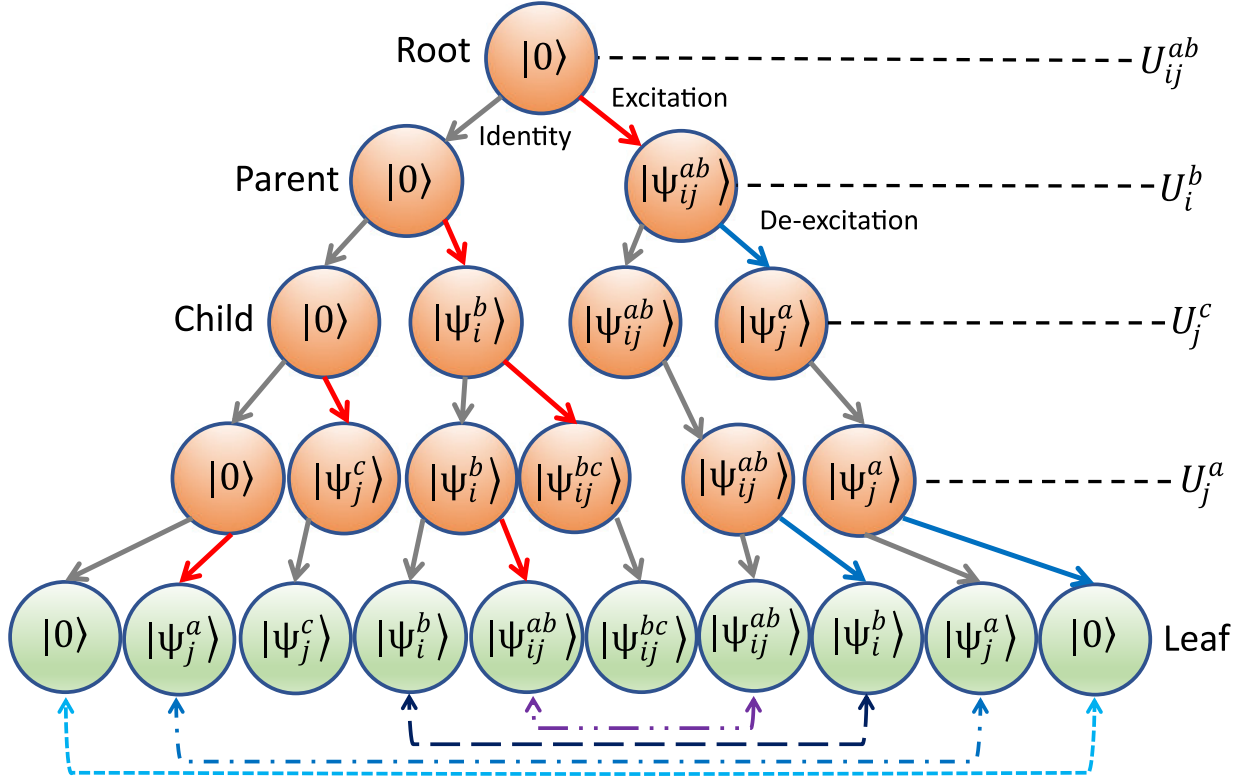


Figure 1: Tree structure of configurations generated by three UCC factors. Red arrows represent excitation, grey identity and blue de-excitation. Lines with two arrows at the bottom level represent elimination of the repeating configurations.

The tree structured wave-function and derivatives can also be used to carry out the ADAPT algorithm¹⁷ on classical computers, since repeating factors are allowed when constructing the tree. But, calculations using a generalized UCC ansatz¹⁶ cannot be performed with the same algorithm. The SU(2) operator identities still hold for each excitation/de-excitation pair in factorized form, but the algorithm to construct the tree for the wavefunction would need to be modified for general orbitals.

The Hamiltonian of the molecule is constructed from the reference state. We use canonicalized Hartree-Fock (HF) wave-functions as the reference state. the second quantized Hamiltonian becomes

$$\hat{\mathcal{H}} = h_0 + \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s, \quad (15)$$

where p, q, r, s label orthogonal spin-orbitals; h_0 is the nuclear repulsion energy, and h_{pq} and h_{pqrs} are one- and two-electron integrals generated by **PySCF** package:^{21,22}

$$h_{pq} = \int dr \phi_p(r) \left(-\frac{\nabla}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \phi_q(r); \quad (16)$$

$$h_{pqrs} = \int dr_1 dr_2 \frac{\phi_p(r_1) \phi_q(r_2) \phi_s(r_1) \phi_r(r_2)}{|r_1 - r_2|}. \quad (17)$$

To initialize these calculations, we need a set of UCC factors in a particular order. Different orderings of the same set of factors can possibly correspond to different wavefunction ansatzes, as pointed out previously²³ (and verified by simply looking carefully at how the tree is formed). In this work, we use second-order Møller–Plesset perturbation theory²⁴ (MP2) to choose the UCC factors and their ordering. MP2 provides amplitudes for double excitations, which are easy to obtain and serve as good estimations of their importance. UCC factors are chosen in descending order of the absolute value of the corresponding MP2 amplitudes.

This strategy to order UCC factors is well defined and provides a concrete selection and ordering scheme for the UCC factors, as the later factors are assumed to be less relevant. A drawback is that MP2 only has amplitudes for double excitations; to include single excitations in calculations, we put the single factors after the double factors, with random order. As we will see later, variational parameters associated with single UCC factors are generally small, thus their ordering ends up being of minor significance. A non-HF starting point, like natural orbitals could be used to improve screening of UCC factors in the future.

With Hartree-Fock spin-orbitals, a parameterized Hamiltonian, UCC factors chosen in MP2 order, and MP2 amplitudes as the initial guess for variational parameters, the initialization step is complete. In the next step, the tree structure of the configurations is generated. When the number of UCC factors is large, the number of configurations on the tree becomes prohibitively large. Eliminating repeating configurations on each level of the tree greatly reduces the memory requirement of the calculation. But this also means one child node can have more than one parent node. In the third step, coefficients for all the configurations on the tree are calculated; the coefficients for the wavefunction and its derivatives are obtained from the leaf nodes. Then, the energy and its derivatives are evaluated from the expectation values of Hamiltonian. Subsequently, the energy and derivatives are fed into an optimization algorithm. We used the Broyden–Fletcher–Goldfarb–Shannon (BFGS) minimization scheme as implemented in **SciPy**.²⁵ If convergence is not achieved, an updated set of values for the variational parameters are employed to recalculate the coefficients. Since the tree structure is fixed and saved in memory, it is not necessary to re-generate it during optimization; this procedure greatly saves time in completing the calculation. This then is the classical algorithm for UCC in factorized form, and its flowchart is illustrated in Fig. 2. To speed up the energy and derivative calculations, we prune the leaf nodes and keep only configurations with an absolute value of their amplitude larger than a specified threshold. In this work, we set the threshold to 10^{-6} , and differences in total energies due to pruning are always found to be smaller than 10^{-5} Ha in several test cases.

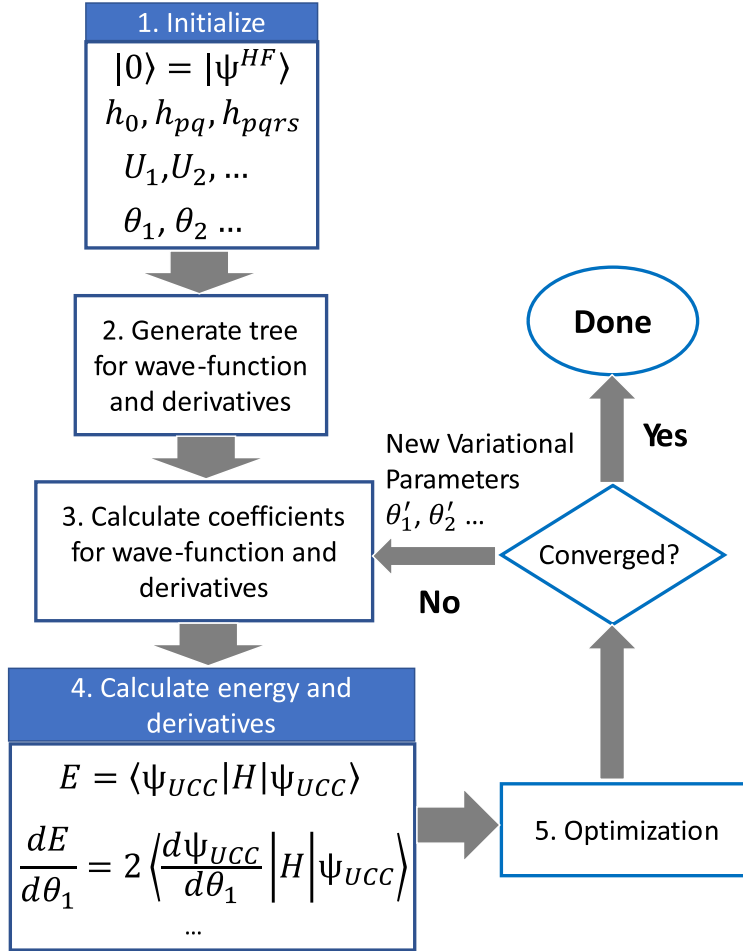


Figure 2: Flowchart of the quantum-inspired algorithm for factorized form of UCC.

3 Results and Discussion

3.1 H_{10} Linear Chain

We apply this quantum inspired UCC algorithm to calculate the ground state of the H_{10} linear chain with the minimal basis set STO-6G. We choose this model system because extensive benchmark calculations using many state-of-art computational methods have already been applied to it.²⁶ Comparison of our results with these standards provides insight into the accuracy of UCC as an electronic structure tool. With the minimal basis set, the reference state of the H_{10} linear chain has 10 molecular orbitals, and half of them are occupied. The Hilbert space for the $\langle \hat{s}_z \rangle = 0$ sector contains 63,504 determinants, with 825 of them doubles and 50 of them singles; we use a lower case $\hat{s}_z$ to refer to the physical z -component of spin of the different product states. Correlation energies (difference between the calculated energy and the Hartree-Fock energy) and the error (difference between the calculated energy and the full configuration interaction (FCI) energy) of several methods as a function of bond length between the H atoms are plotted in Fig. 3.

The correlation energy of the hydrogen chain increases with the bond length, which indicates that the system crosses over from being weakly correlated to being strongly correlated. In the weakly correlated regime, UCCSD results are very close those obtained by CCSD, which is known to be the method of choice for weakly correlated molecules. When the correlation strength goes beyond a certain point (where the bond length is approximately 2.5 Bohr in this case), CCSD energy is no longer bounded by FCI from below, and the error can become quite large. As constructed, the UCCSD calculation is always variational. Calculations show that the UCCSD curve follows the FCI results closely. Compared to the configuration interaction singles and doubles (CISD), which is also variational, and the MP2 method, which is not, the UCCSD provides significantly improved energies for all the bond lengths covered here.

Additional insights about the UCC can be obtained by examining the final variational

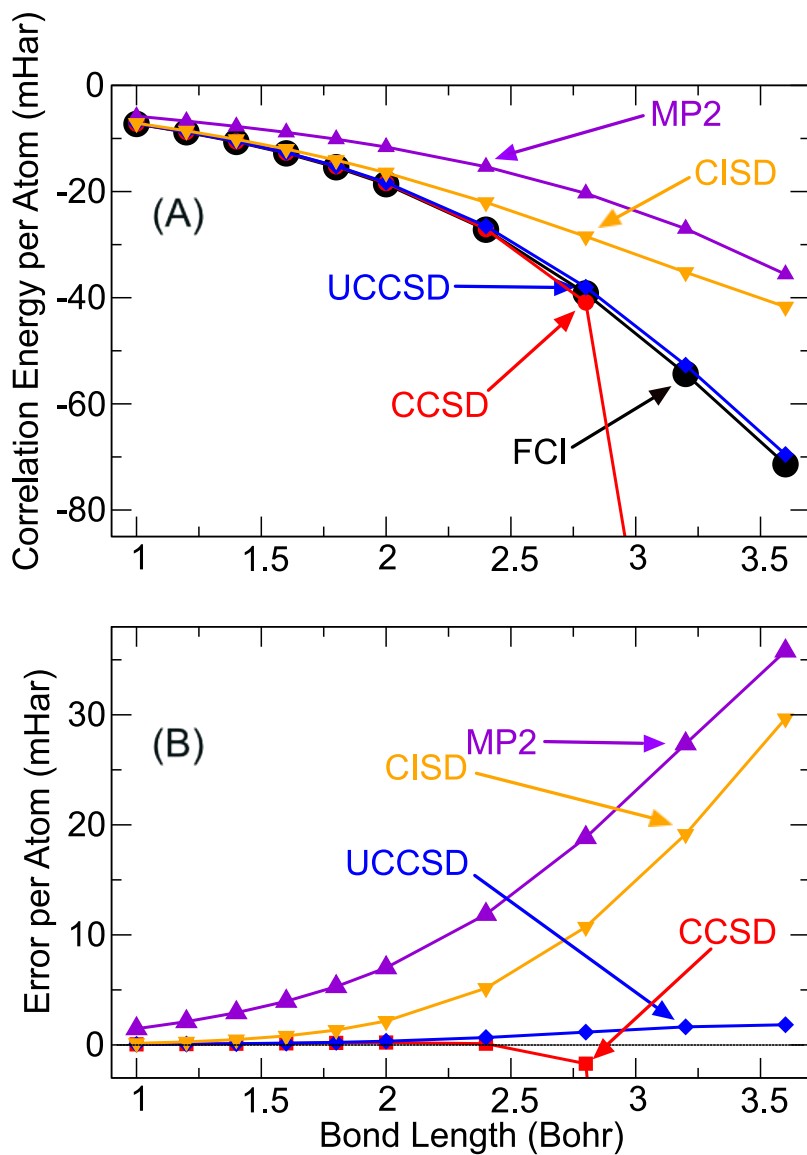


Figure 3: Correlation energy (panel A) and error (panel B) calculated with several quantum chemistry methods as a function of the bond length of the hydrogen linear chain. Data for CCSD are from Ref. 26. Data for other methods are from calculations performed in this work.

wavefunctions generated in the calculations. Values of the variational parameters θ after optimization, for the weakly and strongly correlated systems are plotted in Fig. 4. A couple of observations can be made about the values of θ . First, a large number of UCC factors of double excitations have θ very close to zero (here essentially all UCC doubles factors with index larger than 450; recall the singles come last). Those factors can be ignored in the UCC calculations if we have good estimations of each UCC factor's importance beforehand. Second, the absolute values of θ are tied to correlation energies. Strongly correlated systems (such as the bond length of 3.6 Bohr, plotted with red squares) have much larger θ values than the weakly correlated system (such as the bond length of 1.0 Bohr, plotted with black circles). It is interesting that all of these angles, even in strongly correlated systems, are clustered close to $\theta = 0$.

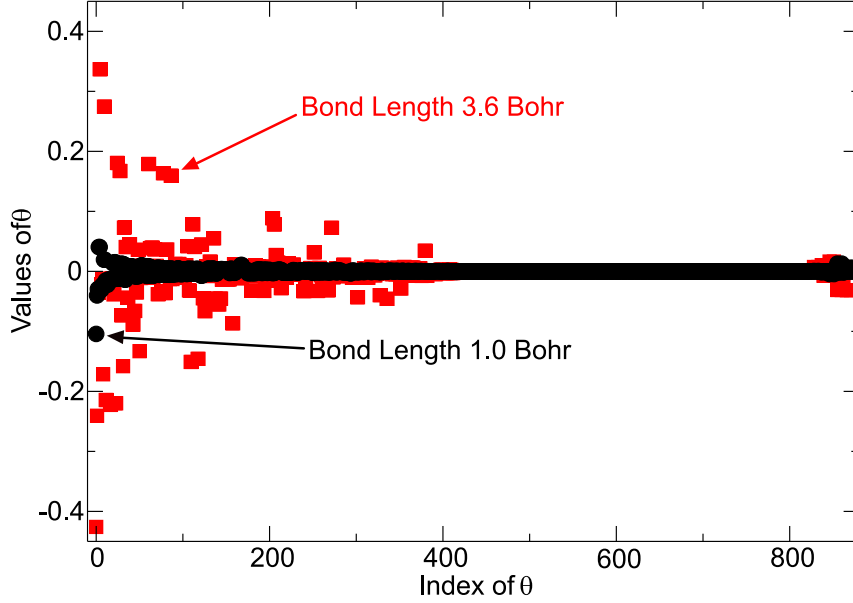


Figure 4: Variational parameters θ obtained from energy optimization at bond lengths of 1.0 (black circles) and 3.6 Bohr (red squares). The first 825 θ values are for UCC double factors and last 50 are for the UCC singles factors.

The quantum phase-estimation algorithm is the ideal approach to calculate the ground-

state energy (and to prepare the ground state) on a quantum computer; it should be possible to use this method once fault-tolerant quantum computation with high depth circuits is feasible. One potential application of the factorized form of UCC is to use it as the initial wave-function preparation method to start the phase-estimation algorithm. The success rates of the phase-estimation algorithm depend on the fidelity of the initially prepared wave-functions, which is determined by the squared overlap between the prepared wave-function and the exact ground state wave-function, (FCI wave-function): $F = |\langle \psi_{\text{approx.}} | \psi_{\text{FCI}} \rangle|^2$. Since, HF and CISD are variational, their states can in principle be prepared on quantum computers, with the Hartree-Fock being trivial in a second-quantized formalism. The fidelity of these approximate wavefunctions is plotted in Fig. 5. We see that factorized form of UCCSD provides a wavefunction with higher fidelity for all bond lengths studied here. The difference is significant in the strongly-correlated regime, which makes UCC a much more suitable state preparation method for the phase-estimation algorithm in those cases.

3.2 H₂O with single and double ζ basis sets

The next molecule we study is the H₂O molecule. With a minimum basis set STO-3g, a calculation using 3 UCC factors has been performed on a trapped ion quantum computer, and simulations (with more factors) on a classical computer. These latter calculations showed that only a small number of UCC factors (about 20) are necessary for UCC to achieve chemical accuracy, which is an energy window of 1.6 milliHartree (mHar) from the ground state.¹⁴ With the STO-3g basis set, the H₂O molecule has 7 molecular orbitals and 10 electrons, which leads to 120 double excitations and 20 single excitations. We put double and single UCC factors in descending order of their MP2 amplitudes, and applied our algorithm to this system. Similarly, we find that 20 UCC factors, (all doubles in our case) are needed to reach chemical accuracy (see Fig. 6). After incorporating all the double and single UCC factors, the UCCSD energies are slightly better than CCSD, and they are much better than CISD.

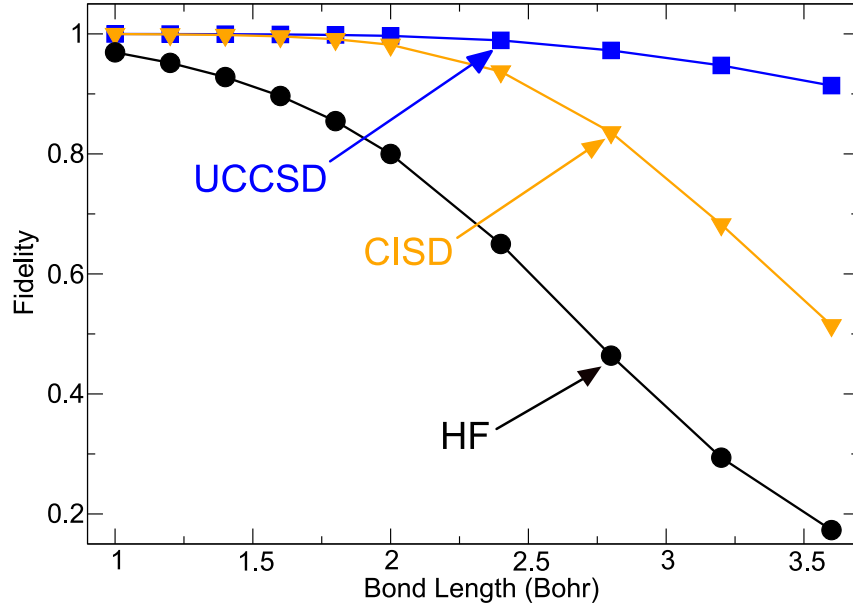


Figure 5: Fidelity of wavefunction prepared by HF, CISD and UCCSD, which are calculated as the squared overlap between the approximate wavefunction $|\psi_{\text{approx.}}\rangle$ and exact ground-state wavefunction (from FCI calculations): $F = |\langle\psi_{\text{approx.}}|\psi_{\text{FCI}}\rangle|^2$

It is known that the performance of the minimum basis sets is generally poor, and a double- ζ basis set can give a striking improvement.²⁷ Here, we perform a UCC calculation for the H_2O molecule with a 6-31g basis set, which has 13 molecular orbitals. To make the calculations more manageable, the orbital of lowest energy (oxygen 1s orbital) is frozen. In total, this system (with one frozen orbital) has 1360 double and 64 single excitations. Results of these calculations are plotted in panel (B) of Fig. 6. First, similar to the H_{10} linear chain, we find that after about 400 doubles, the remaining 1000 doubles UCC factors do not further improve the energy. The energy after including all the single and double UCC factors is again, slightly lower than CCSD, and much better than MP2 and CISD. Unlike the system with sto-3g basis set, the chemical accuracy can only be achieved after including all the important singles and doubles, which indicates chemical accuracy is harder to reach when we use larger basis sets.

As apparent from the tree structure, UCC in a factorized form provides a low-rank representation for the many-body wave-function. This is similar to conventional coupled cluster theory, which also uses a low-rank representation, with the added benefit that the factorized form of UCC has higher accuracy for more strongly correlated systems. Computationally, two steps in the procedure are time consuming: the construction of the wave-functions and the calculation of the energy expectation values for the given wavefunction. For molecules studied here, the energy-expectation-value calculation, which involves multiplying a Hamiltonian matrix times a state vector, is the more time-consuming of the two. The iterative diagonalization method used in CI calculations is also limited by matrix-vector multiplications. So, UCC as implemented in this work, has similar computational cost as a selective configuration interaction method with the same number of determinants in the wave-function; the factorized form of the UCC approach is expected to have more matrix-vector multiplications than a typical selective CI will have, due to its use of an optimizer. Total number of determinants can grow exponentially as number of UCC factors increases, which eventually limits size of systems can be studied by this method.

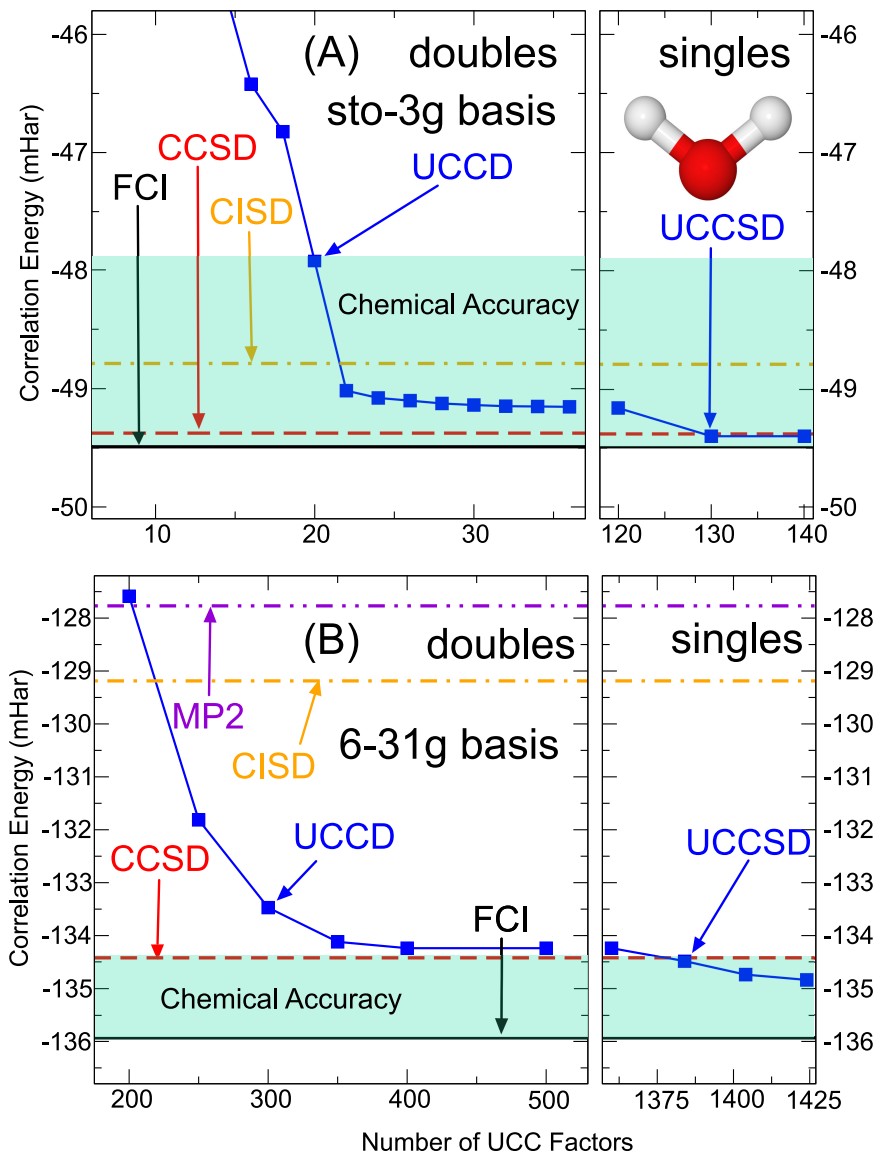


Figure 6: Panel(A): correlation energy of the H_2O molecule from a UCC calculation with the single- ζ basis set STO-3g. The first 120 UCC factors are doubles, and the last 20 are singles. Panel (B): correlation energy from a calculation with the double- ζ basis set 6-31g. The first 1360 UCC factors are doubles and last 64 are singles.

4 Conclusion

In this work, we presented a classical algorithm for the factorized form of UCC, which was inspired by the quantum VQE algorithm. The foundation of this algorithm is an operator identity for the general UCC factor, uncovered from a hidden $SU(2)$ symmetry. Implementation of this algorithm allowed us to apply UCC to larger systems than studied before on classical computers. Compared to more established quantum chemistry methods, the factorized form UCC is as accurate as CC for weakly correlated systems, and provides much better results than CI with the same excitations for strongly correlated systems. Since we can choose factors and their orders, the factorized form of UCC can be molded for different needs and computational budgets. This accuracy and flexibility make UCC a very valuable tool for chemistry, material science and condensed matter physics.

Acknowledgement

JC and HPC are supported by the Department of Energy, Basic Energy Sciences under contract DE-FG02-02ER45995. JKF is supported from the National Science Foundation under grant number CHE-1836497. JKF is also funded by the McDevitt bequest at Georgetown University. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. We also acknowledge useful discussions with Rodney Bartlett, Joseph Lee, John Staunton, Cyrus Umrigar, Luogen Xu and Dominika Zgid.

References

- (1) Kutzelnigg, W. In *Methods of Electronic Structure Theory*; Schaefer, H. F., Ed.; Springer US: New York, NY, USA, 1977; Chapter 5, pp 129–188.

- (2) Koch, S.; Kutzelnigg, W. Comparison of CEPA and CP-MET methods. *Theor. Chim. Acta* **1981**, *59*, 387–411.
- (3) Kutzelnigg, W.; Koch, S. Quantum chemistry in Fock space. II. Effective Hamiltonians in Fock space. *J. Chem. Phys.* **1983**, *79*, 4315–4335.
- (4) Bartlett, R. J.; Kucharski, S. A.; Noga, J. Alternative coupled-cluster ansätze II. The unitary coupled-cluster method. *Chem. Phys. Lett.* **1989**, *155*, 133 – 140.
- (5) Watts, J. D.; Trucks, G. W.; Bartlett, R. J. The unitary coupled-cluster approach and molecular properties. Applications of the UCC(4) method. *Chem. Phys. Lett.* **1989**, *157*, 359 – 366.
- (6) Hoffmann, M. R.; Simons, J. Analytical energy gradients for a unitary coupled-cluster theory. *Chem. Phys. Lett.* **1987**, *142*, 451 – 454.
- (7) Taube, A. G.; Bartlett, R. J. New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* **2006**, *106*, 3393–3401.
- (8) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.
- (9) Shavitt, I.; Bartlett, J. R. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press, 2009.
- (10) Cao, Y.; Romero, J.; Olson, J. P.; Degroote, M.; Johnson, P. D.; Kieferová, M.; Kivlichan, I. D.; Menke, T.; Peropadre, B.; Sawaya, N. P. D.; Sim, S.; Veis, L.; Aspuru-Guzik, A. Quantum Chemistry in the Age of Quantum Computing. *Chem. Rev.* **2019**, *119*, 10856–10915, PMID: 31469277.
- (11) 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.

- (12) 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.
- (13) Hempel, C.; Maier, C.; Romero, J.; McClean, J.; Monz, T.; Shen, H.; Jurcevic, P.; Lanyon, B. P.; Love, P.; Babbush, R.; Aspuru-Guzik, A.; Blatt, R.; Roos, C. F. Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator. *Phys. Rev. X* **2018**, *8*, 031022.
- (14) Nam, Y.; Chen, J.-S.; Pienti, N. C.; Wright, K.; Delaney, C.; Maslov, D.; Brown, K. R.; Allen, S.; Amini, J. M.; Apisdorf, J.; Beck, K. M.; Blinov, A.; Chaplin, V.; Chmielewski, M.; Collins, C.; Debnath, S.; Hudek, K. M.; Ducore, A. M.; Keesan, M.; Kreikemeier, S. M.; Mizrahi, J.; Solomon, P.; Williams, M.; Wong-Campos, J. D.; Moehring, D.; Monroe, C.; Kim, J. Ground-state energy estimation of the water molecule on a trapped-ion quantum computer. *npj Quantum Inf.* **2020**, *6*, 33.
- (15) Dallaire-Demers, P.-L.; Romero, J.; Veis, L.; Sim, S.; Aspuru-Guzik, A. Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer. *Quantum Sci. Technol.* **2019**, *4*, 045005.
- (16) Lee, J.; Huggins, W. J.; Head-Gordon, M.; Whaley, K. B. Generalized Unitary Coupled Cluster Wave functions for Quantum Computation. *J. Chem. Theory Comput.* **2019**, *15*, 311–324.
- (17) Grimsley, H. R.; Economou, S. E.; Barnes, E.; Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* **2019**, *10*, 3007.
- (18) Evangelista, F. A.; Chan, G. K.-L.; Scuseria, G. E. Exact parameterization of fermionic wave functions via unitary coupled cluster theory. *J. Chem. Phys.* **2019**, *151*, 244112.

- (19) Xu, L.; Lee, J. T.; Freericks, J. K. Test of the unitary coupled-cluster variational quantum eigensolver for a simple strongly correlated condensed-matter system. *Mod. Phys. Lett. B* **2020**, *34*, 2040049.
- (20) Handy, N. Multi-root configuration interaction calculations. *Chem. Phys. Lett.* **1980**, *74*, 280 – 283.
- (21) 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.; Wouters, S.; Chan, G. K.-L. PySCF: the Python-based simulations of chemistry framework. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1340.
- (22) Sun, Q.; Zhang, X.; Banerjee, S.; Bao, P.; Barbry, M.; Blunt, N. S.; Bogdanov, N. A.; Booth, G. H.; Chen, J.; Cui, Z.-H.; Eriksen, J. J.; Gao, Y.; Guo, S.; Hermann, J.; Hermes, M. R.; Koh, K.; Koval, P.; Lehtola, S.; Li, Z.; Liu, J.; Mardirossian, N.; McClain, J. D.; Motta, M.; Mussard, B.; Pham, H. Q.; Pulkin, A.; Purwanto, W.; Robinson, P. J.; Ronca, E.; Sayfutyarova, E. R.; Scheurer, M.; Schurkus, H. F.; Smith, J. E. T.; Sun, C.; Sun, S.-N.; Upadhyay, S.; Wagner, L. K.; Wang, X.; White, A.; Whitfield, J. D.; Williamson, M. J.; Wouters, S.; Yang, J.; Yu, J. M.; Zhu, T.; Berkelbach, T. C.; Sharma, S.; Sokolov, A. Y.; Chan, G. K.-L. Recent developments in the PySCF program package. *J. Chem. Phys.* **2020**, *153*, 024109.
- (23) Grimsley, H. R.; Claudino, D.; Economou, S. E.; Barnes, E.; Mayhall, N. J. Is the Trotterized UCCSD Ansatz Chemically Well-Defined? *J. Chem. Theory Comput.* **2020**, *16*, 1–6, PMID: 31841333.
- (24) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (25) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.;

- Brett, M.; Wilson, J.; Jarrod Millman, K.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C.; Polat, İ.; Feng, Y.; Moore, E. W.; Vand erPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; Contributors, S. . . SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272.
- (26) Motta, M.; Ceperley, D. M.; Chan, G. K.-L.; Gomez, J. A.; Gull, E.; Guo, S.; Jiménez-Hoyos, C. A.; Lan, T. N.; Li, J.; Ma, F.; Millis, A. J.; Prokof'ev, N. V.; Ray, U.; Scuseria, G. E.; Sorella, S.; Stoudenmire, E. M.; Sun, Q.; Tupitsyn, I. S.; White, S. R.; Zgid, D.; Zhang, S. Towards the Solution of the Many-Electron Problem in Real Materials: Equation of State of the Hydrogen Chain with State-of-the-Art Many-Body Methods. *Phys. Rev. X* **2017**, *7*, 031059.
- (27) Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley; 1 edition, 2013.

Graphical TOC Entry

