Symmetry Breaking Slows Convergence of the ADAPT Variational Quantum Eigensolver

Luke W. Bertels, Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Because quantum simulation of molecular systems is expected to provide the strongest advantage over classical computing methods for systems exhibiting strong electron correlation, it is critical that the performance of VQEs be assessed for strongly correlated systems. For classical simulation, strong correlation often results in symmetry breaking of the Hartree−Fock reference, leading to Löwdin's well-known "symmetry dilemma", whereby accuracy in the energy can be increased by breaking spin or spatial symmetries. Here, we explore the impact of symmetry breaking on the performance of ADAPT-VQE using two strongly correlated systems: (i) the "fermionized" anisotropic Heisenberg model, where the anisotropy parameter controls the correlation in the system, and (ii) symmetrically stretched linear H_4, where correlation increases with increasing H−H separation. In both of these cases, increasing the level of correlation of the system leads to spontaneous symmetry breaking (parity and $\hat{S}^2$, respectively) of the mean-field solutions. We analyze the role that symmetry breaking in the reference states and orbital mappings of the fermionic Hamiltonians have in the compactness and performance of ADAPT-VQE. We observe that improving the energy of the reference states by breaking symmetry has a deleterious effect on ADAPT-VQE by increasing the length of the ansatz necessary for energy convergence and exacerbating the problem of "gradient troughs".
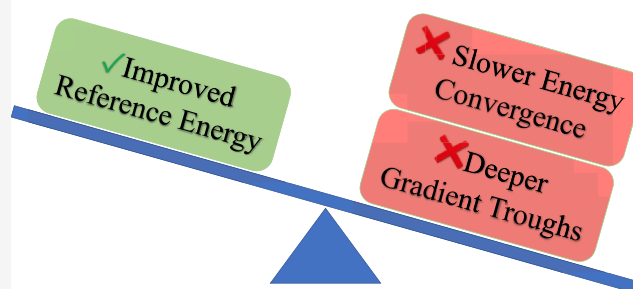
1. INTRODUCTION

The simulation of ground electronic states of molecular Hamiltonians is a fundamental goal of theoretical chemistry. Several classes of methods exist within the realm of classical computing for treating these systems, including density functional theory (DFT),[1,2] Hartree-Fock (HF) theory,[3,4] Møller−Plesset (MP) perturbation theory,[5] coupled cluster (CC) theory,[6,7] configuration interaction (CI), and many others. Although DFT and HF are by nature approximate, MP2, CC, and CI are systematically improvable. For systems with weakly correlated electrons, inclusion of only single and double excitations in MP or CC theory is sufficient for accurate results, while for systems with strongly correlated electrons, low-excitation rank methods fail to capture the physics required to provide accurate energies, although increasing the excitation rank leads to a rapid increase in the computational resources required, with exact treatment [full CI (FCI)] requiring combinatorial scaling with system size to cover the Hilbert space of the system.
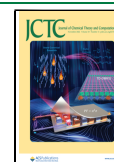
Simulation of chemical systems on quantum computers, however, offers an attractive alternative to classical simulations as the quantum mechanical structure is efficiently captured in the quantum nature of the device,[8] that is, the combinatorial growth of the Hilbert space with system size is absorbed by the quantum processor.[9] The promise of quantum computation for chemistry is currently limited by the small numbers of qubits (10^1-10^2) in existing quantum devices and the quality of these qubits (limited coherence times, limited connectivity between qubits, and gate errors).[10,11] In this noisy intermediate-scale quantum (NISQ) era, the success of quantum simulation depends on both the quality of qubits and the ability of quantum algorithms to cope with these limitations.[12]

The variational quantum eigensolver (VQE), originally proposed by Peruzzo et al.,[13] offers an approach for quantum simulation of chemical Hamiltonians. VQE is a hybrid quantum−classical algorithm in which the computational work is divided between a quantum processor and a classical co-processor.[14] In this scheme, one prepares parameterized trial states $|\psi(\vec{\theta})\rangle$ on the quantum processor and minimizes the expectation value of the Hamiltonian with respect to the ansatz parameters

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

$$= \min_{\vec{\theta}} \sum_i g_i \langle \psi(\vec{\theta}) | \hat{o}_i | \psi(\vec{\theta}) \rangle \tag{2}$$

where g_i represents the classically precomputed one- and two-electron integrals and $\hat{o}_i$ represents the corresponding one- and two-electron operators. As the different $\hat{o}_i$ terms do not generally commute, state preparation and measurement of the terms in eq 2 must be performed multiple times to statistically converge the expectation values. Several procedures for efficient measurement schemes[14−24] (well-summarized in ref 24) and state reconstruction via classical shadows[25,26] have been proposed to reduce the number of measurements required for convergence of the energy.

Here, the quantum computer is used to prepare trial states and measure the molecular Hamiltonian, while the classical computer is used to determine ansatz parameter updates. Trial states are prepared by applying a parameterized unitary operator, $\hat{U}(\vec{\theta})$, to a reference state $|\psi^{(0)}\rangle$

$$|\psi(\vec{\theta})\rangle = \hat{U}(\vec{\theta})|\psi^{(0)}\rangle \tag{3}$$

Several VQE ansätze have been explored for theoretical studies[13−15,27−34] and on quantum hardware,[15,27,31,35,36] many of which are modifications of the unitary coupled cluster (UCC) ansatz[37−40] from classical electronic structure theory. Although in principle a circuit implementation of an arbitrarily expressive ansatz can map the reference to any state in the Hilbert space (including the exact FCI state), practical limits of NISQ devices impose limits on the structure of $\hat{U}(\vec{\theta})$. "Hardware-efficient" ansätze,[15,34] while being more expressive and generally more resilient to short coherence times, are known to suffer from barren plateaus,[41−43] while "chemistry-inspired" ansätze[13,14,27,29−31,44] are less expressive and more limited by short coherence times but target a more chemically relevant region of Hilbert space. Therefore, although the reduced circuit depth of VQEs versus PEA (at the cost of many more measurements) makes VQEs attractive for NISQ devices, the limits imposed on $\hat{U}(\vec{\theta})$ mean that VQEs typically produce approximate solutions. The accuracy of VQEs is, therefore, ultimately limited by the variational flexibility of the predefined ansatz.

Unlike VQEs with statically defined ansätze, the adaptive problem-tailored VQE (ADAPT-VQE) method, developed by Grimsley et al.,[45] avoids a predefined unitary ansatz by constructing an arbitrarily accurate quasi-optimal ansatz on the fly. This is achieved by iteratively growing the ansatz by adding operators from a pool one-at-a-time as informed by the Hamiltonian. ADAPT-VQE has been shown to simultaneously provide smaller gate counts and errors than those of traditional VQE methods.[45,46] Further improvements in the ADAPT-VQE framework have come from the introduction of qubit-based operator pools (qubit-ADAPT-VQE)[47] and minimally complete pools to reduce measurement overhead.[48] The success of ADAPT-VQE has inspired the development of other adaptive VQEs as well, including iterative qubit excitation-based VQE (QEB-ADAPT-VQE),[49] mutual information-assisted adaptive VQE,[50] and the adaptive variational quantum imaginary time evolution (AVQITE) method.[51]

An additional motivation for the adaptive construction of ansätze is the ability to adapt to systems that are strongly correlated, where the performance of classical methods and even traditional VQEs is expected to suffer. In this work, we investigate the performance of ADAPT-VQE in two distinct systems that display variable amounts of strong correlation: (i) the fermionized anisotropic Heisenberg model, where the anisotropy parameter allows for control over the level of correlation in the system, and (ii) the symmetric dissociation of linear H_4. In both of these cases, increasing the level of correlation of the system leads to spontaneous symmetry breaking (parity and $\hat{S}^2$, respectively) of the mean-field solutions. We explore the roles played by these symmetries, both in the reference state and the operator pool, for ADAPT-VQE, highlighting their importance in generating compact ansätze and preventing premature convergence of the algorithm. Our results bolster the findings of Barron et al.[52] and Shkolnikov et al.[48] on the importance of building the symmetries of the Hamiltonian into the operator pools.

2. BACKGROUND

2.1. ADAPT-VQE Algorithm. Unlike traditional VQE, which begins with a predetermined form of the unitary $\hat{U}(\vec{\theta})$, ADAPT-VQE iteratively grows a problem-tailored unitary ansatz by adding operators one-at-a-time from a predetermined operator pool. Before the algorithm begins, the Hamiltonian coefficients are computed and mapped to a qubit representation, as in traditional VQE. The operators, $\hat{A}_k$, in the pools used in this work take the form of anti-Hermitian sums of generalized excitation and de-excitation operators, for example,

$$\hat{A}_q^{\,p} = \hat{a}_p^\dagger \hat{a}_q - \hat{a}_q^\dagger \hat{a}_p \tag{4}$$

$$\hat{A}_{rs}^{\,pq} = \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r - \hat{a}_r^\dagger \hat{a}_s^\dagger \hat{a}_q \hat{a}_p \tag{5}$$

where p, q, r, and s are arbitrary spin−orbital indices (see App. A). Exponentiation of these anti-Hermitian operators yields unitary operators. Although other operator pools have been explored,[47−49,53] we focus here on the fermionic operator pool. The ADAPT-VQE trial state is then initialized with a reference state that is easily prepared on the device, typically a product state corresponding to the HF determinant. To grow the ansatz, the current trial state, $|\psi^{(n)}\rangle$, is prepared on the device and the gradient of the energy with respect to the operator parameters θ_k for each operator $\hat{A}_k$ in the pool is measured. This is performed by measuring the expectation value of the commutator of the Hamiltonian and the operators for the current state

$$\frac{\partial E^{(n)}}{\partial \theta_k} = \langle \psi^{(n)} | [\hat{H}, \hat{A}_k] | \psi^{(n)} \rangle \tag{6}$$

This gradient measurement step of ADAPT-VQE is highly parallelizable over multiple uncoupled devices. The operator corresponding to the largest gradient magnitude is then used to form the new trial-state ansatz

$$\begin{aligned}
|\psi^{(n+1)}(\vec{\theta}^{(n+1)})\rangle &= e^{\theta_{n+1}\hat{A}_{n+1}}|\psi^{(n)}\rangle \\
&= e^{\theta_{n+1}\hat{A}_{n+1}} e^{\theta_n \hat{A}_n} \cdots e^{\theta_1 \hat{A}_1}|\psi^{(0)}\rangle
\end{aligned} \tag{7}$$

The new parameter θ_{n+1} is initialized to 0, while the initial values for the other parameters are taken to be the optimized values from the previous iteration. The new ansatz is then optimized over all $\vec{\theta}^{(n+1)}$ via a VQE subroutine to yield $|\psi^{(n+1)}\rangle$. From here, the algorithm repeats by returning to the operator gradient

measurement step. Convergence of the ADAPT-VQE algorithm may be determined in a number of ways, including the norm (L^2 or L^∞) of the operator gradient $\left| \frac{\partial E^{(n)}}{\partial \theta_k} \right| < \epsilon$, the variance of the ADAPT-VQE state $\langle \psi^{(n)} | \hat{H}^2 | \psi^{(n)} \rangle - (E^{(n)})^2 < \epsilon$, and the energy change between iterations $|E^{(n)} - E^{(n-1)}| < \epsilon$. The operator pools are not "drained" by the addition of an operator to the ansatz; a given operator may be added to the ansatz more than once, with different parameters for each occurrence. Because of this, ADAPT-VQE can be viewed as an algorithm that approximates the exact (FCI) ground state to arbitrary accuracy by appending multiple instances of the operators

$$|\psi^{FCI}\rangle = \prod_l^\infty \prod_k e^{\theta_k^{(l)} \hat{A}_k} |\psi^{(0)}\rangle, \tag{8}$$

where the parameters $\theta_k^{(l)}$ are allowed to vary independently for different l. This property of ADAPT-VQE, however, requires completeness of the operator pool (see ref 48 for a rigorous definition). Similar work on the exactness of general trotterized UCC variants has been explored by Evangelista et al.[54] and demonstrates the completeness of the UCC single- and double-excitation/de-excitation pool. ref 45 presents a more detailed explanation and demonstration of ADAPT-VQE and this connection to FCI.

2.2. Heisenberg Model. Although spin Hamiltonians are most often associated with condensed matter physics, these Hamiltonians are also often used in the context of chemistry as model systems to develop a coarse-grained understanding of certain molecular interactions.[55−65] Such models are useful for describing the interactions between open-shell fragments, such as metal atoms in multi-metal organometallic complexes. These interactions are broadly classified as ferromagnetic coupling or antiferromagnetic coupling based on whether or not the ground state has a spin magnetic moment.

In this picture, the unpaired electrons on a given metal atom are aligned parallel to each other while the unpaired electrons on different metal atoms align either parallel (ferromagnetic coupling) or antiparallel (antiferomagnetic coupling) to each other. Although purely ab initio approaches to describe these systems must contend with strongly interacting electrons within nearly degenerate orbitals, the effective-Hamiltonian approaches reduce these to the interactions between the net spins on different fragments. The exchange interaction, a consequence of Fermi statistics, provides the energetic driving force behind this coupling. For fixed oxidation states, the Heisenberg−Dirac−van Vleck Hamiltonian (HDvV)[66−68] provides a simple model that depends on the net spin of the different metal centers

$$\hat{H}^{HDvV} = -2\sum_{ij} J_{ij} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j \tag{9}$$

$$= -2\sum_{ij} J_{ij} (\hat{S}_i^x \hat{S}_j^x + \hat{S}_i^y \hat{S}_j^y + \hat{S}_i^z \hat{S}_j^z). \tag{10}$$

$J_{ij} > 0$ couples sites i and j ferromagnetically while $J_{ij} < 0$ couples them antiferromagnetically. The problem then shifts (slightly) from describing the many interactions between the electrons to describing the interactions between the spins, namely, obtaining values for J_{ij}.

The Heisenberg spin Hamiltonian can also be considered as a model for strong fermionic correlation as well, when viewed at

the strong-correlation limit of the fermionic Hubbard model. The fermionic Hubbard Hamiltonian is given as follows

$$\hat{H}^{Hubbard} = t\sum_{ij} \sum_\sigma \hat{a}_{i,\sigma}^\dagger \hat{a}_{j,\sigma} + \frac{1}{2} U \sum_i \hat{n}_{i,\sigma} \hat{n}_{i,\bar{\sigma}}, \tag{11}$$

where t is the single-electron hopping and U is the two-electron repulsion. When $t/U \gg 1$, the system is dominated by hopping (kinetic energy-like), and at the limit $U \to 0$, it becomes a free-electron system, where the electrons delocalize over the entire lattice. In this regime, the delocalized state may be taken as the zeroth-order solution, with correlations between electrons being handled by perturbation theory. At the opposite limit, where $t/U \ll 1$, the system becomes localized. In this regime, degenerate perturbation theory may be used to treat delocalization as a perturbation to the $\binom{N}{k}$-fold degenerate localized ground states, where N is the number of sites and k is the number of electrons. This approach yields the Heisenberg Hamiltonian, and at the second order, $J_{ij}^{(2)} = -\frac{t^2}{U}$ (see derivation in ref 69). This connection between the Hubbard model at the limit of large electron−electron repulsion and the Heisenberg spin Hamiltonian suggests the latter to be a model for studying strong correlation. Although our use of the Heisenberg model is as a proxy for chemical systems, the use of quantum computers to simulate model Hamiltonians is an important field in its own right, with much recent work in this context.[44,52,70−77]

2.2.1. Anisotropic Heisenberg Hamiltonian. The HDvV Hamiltonian given in eqs 9 and 10 is referred to as being isotropic, meaning that the x, y, and z components of the total spin are treated equivalently. If interactions are present that break this equivalence (e.g., dipolar-like couplings), the resulting effective spin Hamiltonian becomes anisotropic. The anisotropic Heisenberg model, also known as the XXZ model, has a Hamiltonian given as follows

$$\hat{H}^{aniso} = -2J \sum_{\langle ij \rangle} (\hat{S}_i^x \hat{S}_j^x + \hat{S}_i^y \hat{S}_j^y) - 2K \sum_{\langle ij \rangle} \hat{S}_i^z \hat{S}_j^z, \tag{12}$$

where $\langle ij \rangle$ restricts the sum over nearest-neighbor sites.

2.2.2. Fermionization of 1D Spin Hamiltonians. Although application of degenerate perturbation theory to the Hubbard model at the $t/U \ll 1$ limit to yield the Heisenberg Hamiltonian provides a connection between fermions and spins, this connection is made more general using the Jordan−Wigner (JW) transformation.[78] The JW transformation has become ubiquitous in quantum simulation of chemistry Hamiltonians as a means to map fermionic Hamiltonians onto qubit Hamiltonians; however, the transformation was originally proposed as a means to map spins onto fermions. For a one-dimensional (1D) spin-1/2 lattice, the anisotropic Heisenberg Hamiltonian is "fermionized" by first writing eq 12 in terms of Pauli ladder operators and then substituting them with fermionic creation/annihilation operators:

$$\hat{H}^{aniso} = -\frac{J}{2} \sum_i (\hat{\sigma}_i^x \hat{\sigma}_{i+1}^x + \hat{\sigma}_i^y \hat{\sigma}_{i+1}^y)$$

$$- \frac{K}{2} \sum_i \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \tag{13}$$

Table 1. Different Cases of the Hamiltonian in eq 12 and Comparisons to Fermionic Models When Fermionized

case	Hamiltonian	features
K = 0	$\hat{H}^{XY} = -2J\sum_{\langle ij\rangle}(\hat{S}_i^x\hat{S}_j^x + \hat{S}_i^y\hat{S}_j^y)$	free fermion model, completely delocalized, no entanglement at the ground state
K = J	$\hat{H}^{HDvV} = -2J\sum_{\langle ij\rangle}(\hat{S}_i^x\hat{S}_j^x + \hat{S}_i^y\hat{S}_j^y + \hat{S}_i^z\hat{S}_j^z)$	isotropic, competition between localization and delocalization, entangled ground state
J = 0	$\hat{H}^{Ising} = -2K\sum_{\langle ij\rangle}\hat{S}_i^z\hat{S}_j^z$	degenerate Néel ground states, completely localized, no entanglement at the ground state

$$= -J\sum_i(\hat{\sigma}_i^+\hat{\sigma}_{i+1}^- + \hat{\sigma}_i^-\hat{\sigma}_{i+1}^+)$$
$$- \frac{K}{2}\sum_i\hat{\sigma}_i^z\hat{\sigma}_{i+1}^z \tag{14}$$

$$= -J\sum_i(\hat{a}_i^\dagger\hat{a}_{i+1} + \hat{a}_i\hat{a}_{i+1}^\dagger)$$
$$- K\sum_i\left(2\hat{a}_i^\dagger\hat{a}_i\hat{a}_{i+1}^\dagger\hat{a}_{i+1} - \hat{a}_i^\dagger\hat{a}_i\right.$$
$$\left. - \hat{a}_{i+1}^\dagger\hat{a}_{i+1} + \frac{1}{2}\right). \tag{15}$$

For $K = 0$, this fermionized Hamiltonian reduces to a one-electron Hamiltonian that is easily diagonalized to yield non-interacting fermions. This model is known as the XY model, and despite acting on the full Hilbert space in the spin representation, it has a trivial simple solution in the fermionic representation. Table 1 summarizes the different limits of the anisotropic Heisenberg Hamiltonian. The ratio of K/J is seen as a correlating parameter as below the isotropic point ($K/J = 1$), increasing K/J increases the correlation in the system by encouraging localization. Above the isotropic point, this correlation decreases with increasing K/J. Beginning at the isotropic point and for all larger K/J, the mean-field solution is seen to break spatial (parity) symmetry. In the first set of results, we apply ADAPT-VQE to the fermionized anisotropic Heisenberg model to investigate the role of parity symmetry in the performance of ADAPT-VQE for both local and non-local representations.

3. COMPUTATIONAL DETAILS

We employ an antiferromagnetically coupled ($J < 0$ and $K < 0$) anisotropic Heisenberg Hamiltonian on a 1D, eight-site lattice as a model system for our calculations. Calculations are performed within the $M_s = 0$ space, after which JW transformation yields a four-fermion system (half-filling). We survey correlation parameter values K/J ranging from 0.001 to 100. The OpenFermion[79] electronic structure theory package is used to build the nearest-neighbor (local) spin Hamiltonians and transform them into local fermionic Hamiltonians.[80] These local fermionic Hamiltonians are diagonalized to yield full configuration interaction (FCI) energies and wavefunctions. For each surveyed value of K/J, the ground-state wavefunction has even symmetry (*gerade*, *g*) with respect to the lattice, and the first-excited state has odd symmetry (*ungerade*, *u*). These two states become degenerate at the $K/J \rightarrow \infty$ limit. The local Hamiltonians are then read into the PySCF[81,82] electronic structure theory package to perform HF. For $K/J < 1$, a wavefunction stability analysis of the HF solutions confirms that the stable HF solution has an even symmetry with respect to the

center of the lattice. We follow this solution for $K/J \geq 1$ to yield symmetry-preserving HF solutions; however, performing a stability analysis on these solutions yields a more stable broken-symmetry solution. The "molecular" orbitals (MOs) from both the symmetry-preserving and symmetry-broken HF solutions are then used to transform the local Hamiltonians into non-local MO bases.

We perform second-order Møller−Plesset perturbation theory (MP2)[5] and traditional (non-unitary) CC theory with single and double excitations (CCSD)[83] on top of these HF solutions with PySCF. The non-local fermionic Hamiltonians are then translated into non-local qubit Hamiltonians within the ADAPT-VQE procedure. This process of transforming from local spin Hamiltonians to non-local spin Hamiltonians is illustrated in Figure 1.

For linear H_4, we survey H−H separations from 0.50 to 3.00 Å. Classical electronic structure calculations are performed with
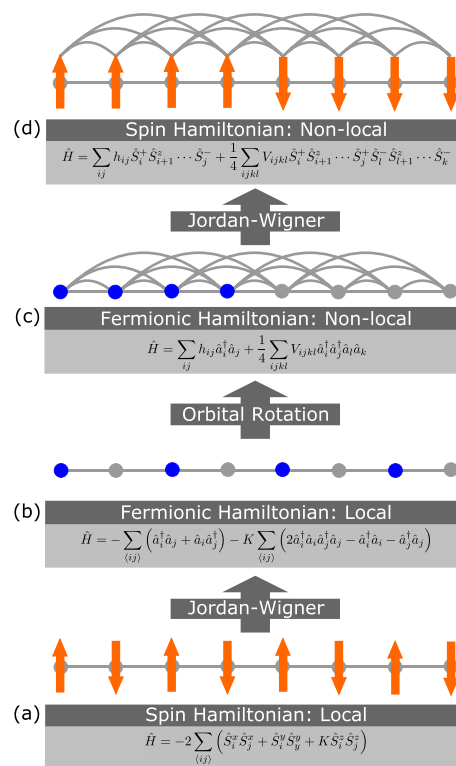

Figure 1. Depiction of the use of the JW transformation and Hartree−Fock to transform the anisotropic Heisenberg Hamiltonian from a local representation to a non-local representation. Gray lines between points depict interactions (not all are shown, but meant to represent the locality/non-locality of the system). Constant terms are neglected, and J is set to 1.
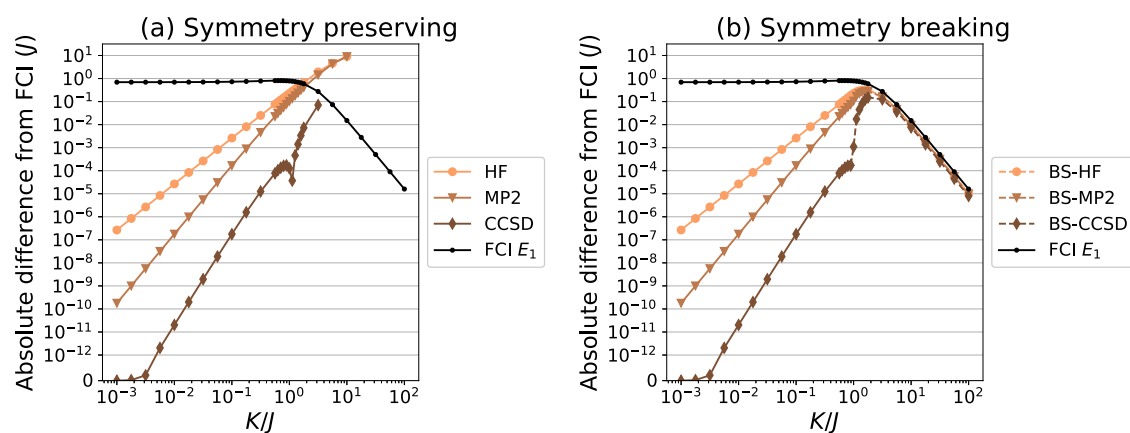
Figure 2. Absolute energy errors from FCI for HF, MP2, and CCSD versus the correlation parameter K/J presented alongside the energy gap between the ground and first-excited FCI states (FCI E_1). (a) Reference state is the symmetry-preserving orbitals. (b) Reference state is the broken-symmetry orbitals.

PySCF, and the STO-3G minimal basis[84] is used for both the classical methods and ADAPT-VQE simulations.

The ADAPT-VQE calculations are simulated without noise on a development branch of our in-house code,[85] which in turn uses OpenFermion[79] for the JW operator transformations and SciPy[86] for BFGS optimization of the parameters in the VQE subroutine.

For the fermionized, anisotropic Heisenberg Hamiltonians, we perform ADAPT-VQE simulations with five combinations of reference states and orbital bases (used to transform the Hamiltonians and define the fermionic operator pools) to investigate the roles that symmetry plays in ADAPT-VQE as correlation increases.

1. **Symmetry-preserving HF orbitals and the reference state**: The canonical HF orbitals are used to transform the Hamiltonians from the local-site orbital basis to the nonlocal, symmetry-preserving MO bases. When the parity of the system is enforced at the HF level, these canonical orbitals are of either g or u character, and as such, the determinants are eigenstates of the parity operator with parities determined from the occupied orbitals. Similarly, the parities of the pool operators are determined from the orbitals used to define the excitations/de-excitations. The reference state is the JW-transformed HF state,

$$|\psi_{HF}^{(0)}\rangle = |11110000\rangle, \qquad (16)$$

which has g symmetry and where orbitals are ordered in increasing energy from left to right. This reference state is the exact ground state at the free-fermion limit ($K \to 0$).

2. **Symmetry-breaking HF orbitals and the reference state:** The canonical HF orbitals are used to transform the Hamiltonian from the local site basis to the nonlocal, broken-symmetry HF orbitals. With the onset of symmetry breaking, these canonical orbitals are of neither g nor u character, and therefore, neither the determinants nor operators have parity symmetry. The reference state is the JW-transformed HF state (eq 16), which also breaks parity symmetry due to the symmetry breaking of the underlying orbitals.

3. **Local orbital basis and the Néel reference state**: The Hamiltonian is expressed on the local-site basis. As the site orbital basis has no parity symmetry, the determinant

basis and operators lack parity symmetry. The reference state is the Néel state,

$$|\psi_{N\acute{e}el}^{(0)}\rangle = |10101010\rangle \qquad (17)$$

which does not have parity symmetry. This reference state is energetically exact, although symmetry-broken, at the Ising limit ($J \to 0$), analogous to how unrestricted HF becomes exact for separated hydrogen atoms.

4. **Local orbital basis and the cat$_+$reference state**: The Hamiltonian is expressed on the local-site orbital basis. Although the basis and operators have no parity symmetry due to the asymmetry of the site orbital basis, the cat$_+$ state, given as the plus superposition of the two complementary Néel states, is used as the reference state and has g symmetry. This reference state lies in the twofold degenerate subspace of the exact ground state at the Ising limit ($J \to 0$).

$$|\psi_{cat_+}^{(0)}\rangle = \frac{1}{\sqrt{2}}(|10101010\rangle + |01010101\rangle) \qquad (18)$$

5. **SALC orbital basis and the cat$_+$reference state**: A symmetry-adapted basis formed by taking the plus and minus linear combinations of complementary site orbitals is used to transform the Hamiltonians from the site orbital basis to the SALC orbital basis. By construction, these SALC orbitals are of either g or u character, and as such, the determinant basis has parity symmetry. Similarly, the parities of the operators are determined from the orbitals used to define the excitations/de-excitations. The reference state is the cat$_+$ state in the SALC orbital basis,

$$\begin{aligned}
|\psi_{SALC,cat_+}^{(0)}\rangle = \frac{1}{2\sqrt{2}} & (|10101010\rangle + |01100110\rangle \\
& + |10011001\rangle + |01010101\rangle \\
& - |01101001\rangle - |10100101\rangle \\
& - |01011010\rangle - |10010110\rangle),
\end{aligned} \qquad (19)$$

which has g symmetry.

For all five combinations of orbital bases and reference states, we use a fermionic generalized singles and doubles (GSD)
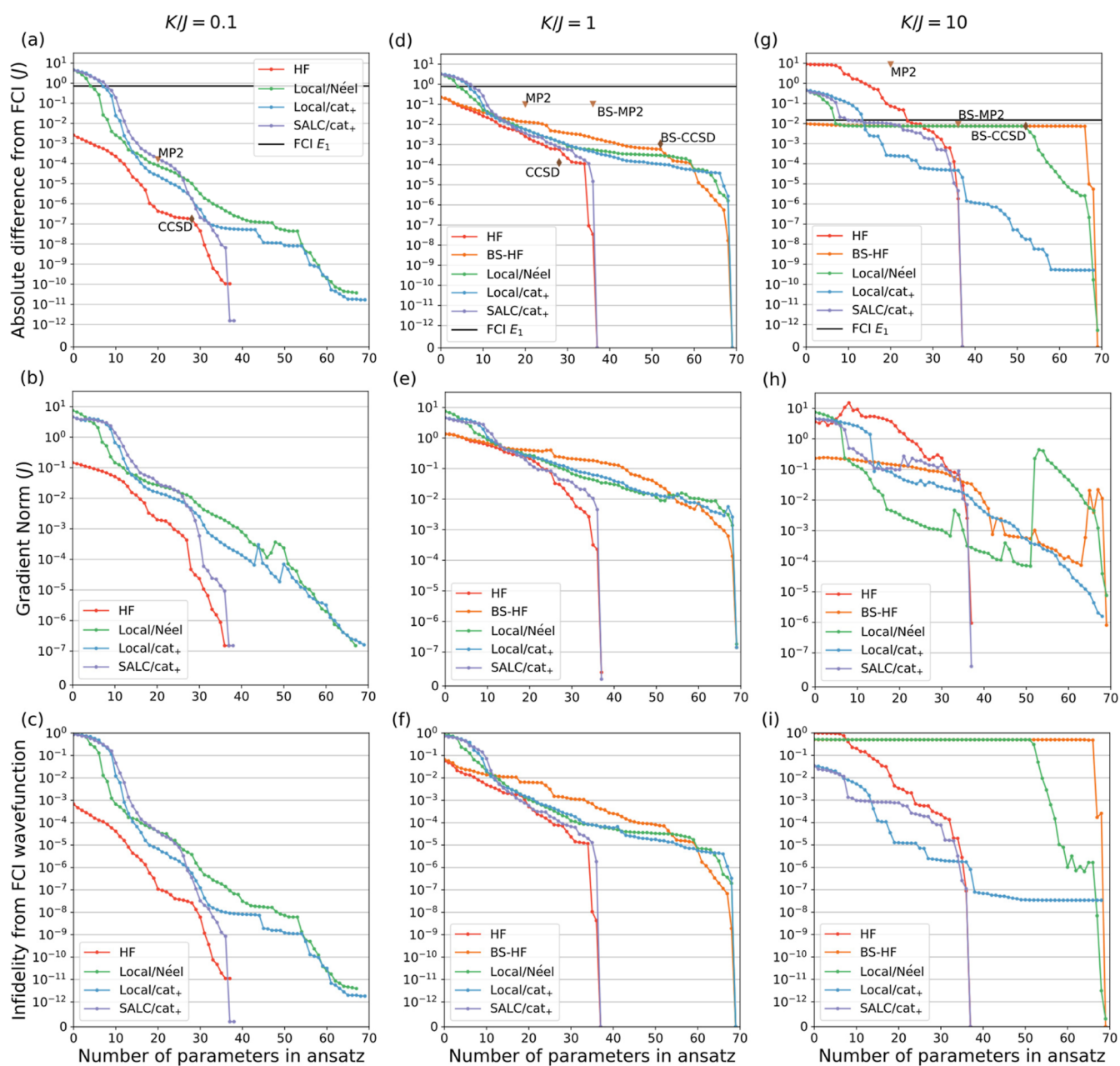
Figure 3. ADAPT-VQE results for the fermionized, anisotropic Heisenberg Hamiltonian. Absolute energy errors [(a), (d), and (g)], ADAPT-VQE gradient norms [(b), (e), and (h)], and infidelities from the exact wavefunction [(c), (f), and (i)] as the ADAPT-VQE ansätze grow are presented for $K/J = 0.1$, $K/J = 1$, and $K/J = 10$. The ADAPT-VQE methods surveyed use symmetry-preserving HF orbitals and the reference state (HF); broken-symmetry HF orbitals and reference state (BS−HF); local site orbital basis and the Néel reference state (local, Néel); local site orbital basis and the cat$_+$ reference state (local, cat$_+$); and symmetry-adapted linear combination orbital basis and the cat$_+$ reference state (SALC, cat$_+$). Absolute energy errors for the classical methods and the energy gap between the ground and first-excited FCI states are presented alongside the ADAPT-VQE errors.

operator pool without symmetry adaptation of the fermionic operators (App. A).

For the symmetric dissociation of H$_4$, we compare the performance of ADAPT-VQE when performed with spin-restricted HF (rHF) and spin-unrestricted HF (uHF) orbitals. This allows us to determine the impact of symmetry breaking in the representation. For rHF, we further explore the impact of spin-adapting the operator pool, by using both the singlet-GSD (sGSD) pool, where the pool operators are symmetry-adapted linear combinations of excitation/de-excitation operators, and the unrestricted-GSD (uGSD) pool, where the excitation/de-

excitation operators in the operator pool are not symmetry-adapted (App. A).

4. RESULTS

4.1. Anisotropic Heisenberg Model. Figure 2(a) presents the absolute errors for the symmetry-preserving HF, MP2, and CCSD on top of these references, and the energy gaps between the ground and first-excited FCI states. For $K/J > 10$, the HF solutions begin to spontaneously break the parity symmetry of the system even when using the previous solutions at lower values of K/J as initial guesses. We are also unable to converge the CCSD amplitude equations for $K/J > 3.16$. The kink in the

CCSD data observed at $K/J = 1.12$ corresponds to the onset of CCSD having a lower energy than that of the FCI ground state. This non-variational behavior persists for all larger values of K/J.

At the weakly correlated limit $K/J \ll 1$, MP2 reduces the energy error of HF by three orders of magnitude, demonstrating a success of simple perturbation theory. As more correlation is added to the system, the breakdown of MP2 becomes evident as its improvement over HF decreases significantly. For $K/J > 1$, the error of HF and MP2 continues to increase with increasing K/J. The HF and MP2 errors exceed the gap between the ground and first-excited FCI states for $K/J > 1.78$ and $K/J > 3.16$, respectively. The error reduction of CCSD is more resilient to increasing correlation, although scanning from $K/J = 0.001$ to $K/J = 1$, the improvement in errors over HF reduces from over seven orders of magnitude to three orders of magnitude.

For $K/J \geq 1$, the symmetry-preserving HF solutions are found to be unstable to symmetry breaking. Figure 2(b) presents the absolute errors for HF (BS−HF), MP2 (BS-MP2), and CCSD (BS-CCSD), where the HF solutions are allowed to break the parity symmetry of the lattice. The energy gap between the ground and first-excited FCI states is also plotted. For $K/J < 1$, the HF solutions do not break symmetry, and therefore, the HF, MP2, and CCSD curves in this region are identical to those in Figure 2(a). The cusp seen at $K/J = 1$ in the CCSD data arises from the change in character of the underlying HF reference due to symmetry breaking. Unlike in the symmetry-preserving case, the symmetry-broken CCSD remains above the FCI value for all surveyed values of K/J.

In the symmetry-breaking regime, the errors in the HF, MP2, and CCSD solutions are much closer to one another compared to those in the $K/J < 1$ regime. CCSD improves upon the HF error by two orders of magnitude ($1 \times 10^{-3} J$ vs $2 \times 10^{-1} J$) for $K/J = 1$, while at $K/J = 100$, CCSD only improves on the HF error by 20% ($8 \times 10^{-6} J$ vs $1 \times 10^{-5} J$). For all K/J, the errors in the broken-symmetry HF, MP2, and CCSD results fall below the energy gap between the ground and first-excited FCI states.

The instability of HF to symmetry breaking for $K/J > 1$ is an example of the "symmetry dilemma" discussed by Löwdin,[87] wherein the most energetically favorable single determinant (classical state) breaks an intrinsic symmetry of the system, while the most energetically favorable symmetry-preserving state is higher in energy.

Using the five calculation settings described above, we now investigate the roles that parity symmetry and locality play in the compactness of the ADAPT-VQE ansätze. In Figure 3, we report the absolute error from the exact ground-state energy, the norm of the ADAPT-VQE gradient, and the infidelity from the exact ground-state wavefunction all versus the number of parameters at the ADAPT-VQE trial state. The infidelity, presented as a measure of closeness of the ADAPT-VQE wavefunction $|\psi^{ADAPT}\rangle$ to the exact ground-state wavefunction $|\psi_0^{FCI}\rangle$, is given as follows

$$\text{Inf}(|\psi_0^{FCI}\rangle, |\psi^{ADAPT}\rangle) = 1 - F(|\psi_0^{FCI}\rangle, |\psi^{ADAPT}\rangle) \qquad (20)$$

$$= 1 - |\langle \psi_0^{FCI}|\psi^{ADAPT}\rangle|^2 \qquad (21)$$

where F is the fidelity of the two states. We note that for systems where there is a near degeneracy between eigenstates, a trial state may nearly converge to an excited eigenstate such that it simultaneously possesses a high infidelity (low overlap with the exact ground state) and small energy error.

4.1.1. Symmetry Breaking Slows Energy Convergence.
Looking across panels 3(a), 3(d), and 3(g), we see that in each case, both symmetry-preserving calculations (HF and SALC/cat$_+$) converged to the exact solution with only 37 parameters. This reduced number of parameters arises from the fact that when symmetry is preserved throughout the state preparation, we only need to parameterize states within the corresponding g-symmetry subspace, which for this eight-site lattice has a dimension of 38. For all the other cases, which involve some symmetry breaking (either in the orbitals, reference state, or both), the full 70-dimensional Hilbert space must be spanned to achieve exact convergence. The 52 CCSD amplitudes can be divided into 28 g excitations and 24 u excitations. In Figure 3, we only report CCSD as having 28 parameters since the remaining 24 u-symmetry excitations cannot contribute when applied to a symmetry-preserving reference state. As such, CCSD does not have enough parameters to span the 38-dimensional g subspace. This is also obvious from the fact that CCSD has neither connected triple nor quadruple excitations, interactions that ADAPT-VQE is able to include by sequential application of one- and two-particle rotations.

For the weaker-correlation case ($K/J = 0.1$), the symmetry-preserving calculations always outperform the symmetry-violating cases (the red curve is always below the rest). This is not too surprising, given that the HF reference state is the most stable product state available. With the onset of symmetry breaking in HF ($K/J = 1$), the broken-symmetry HF is slightly more favorable. However, ADAPT(HF) begins to outperform ADAPT(BS-HF) after a single iteration. On the other hand, when the correlation increases to $K/J = 10$, the symmetry-preserving HF reference is no longer the lowest-energy product state but instead is now the highest-energy reference state considered in our data. As a consequence, in this strong-correlation regime, the use of a broken-symmetry reference leads to lower energy at early stages of the algorithm. However, this energetic advantage of the broken-symmetry reference quickly becomes a disadvantage due to a very slow convergence at later stages [seen as the flat lining of the green and orange curves in Figure 3(g)].

4.1.2. Symmetry Breaking Worsens Gradient Troughs.
As reported recently,[88] strongly correlated systems are susceptible to exhibiting gradient troughs, whereby the gradients of the pool operators initially diminish before eventually increasing prior to convergence. This non-monotonic convergence is problematic because it appears to the user as false convergence. For the fermionized, anisotropic Heisenberg model studied here, we again observe the onset of gradient troughs when the correlation is increased, as is clearly evident in Figure 3(h). However, when we allow the symmetry to break, we find that problems with gradient troughs worsen.

For both symmetry-breaking references, ADAPT(BS-HF) (orange) and ADAPT(local/Néel) (green), the norm of the operator pool gradient is seen to decrease with the addition of operators to the ansatz and then suddenly increase by several orders of magnitude. Before escaping from the gradient trough, the energy errors for ADAPT(BS-HF) and ADAPT(local/Néel) lie close to those of broken-symmetry CCSD, which is approximately one-half of the gap between the exact ground and first-excited FCI states [panel 3(g)]. Additionally, the infidelities of the ADAPT(BS-HF) and ADAPT(local/Néel) states are approximately 0.5 before escaping the gradient trough [panel 3(i)]. In these cases, broken-symmetry CCSD, ADAPT-(BS-HF), and ADAPT(local/Néel) appear to be approximating
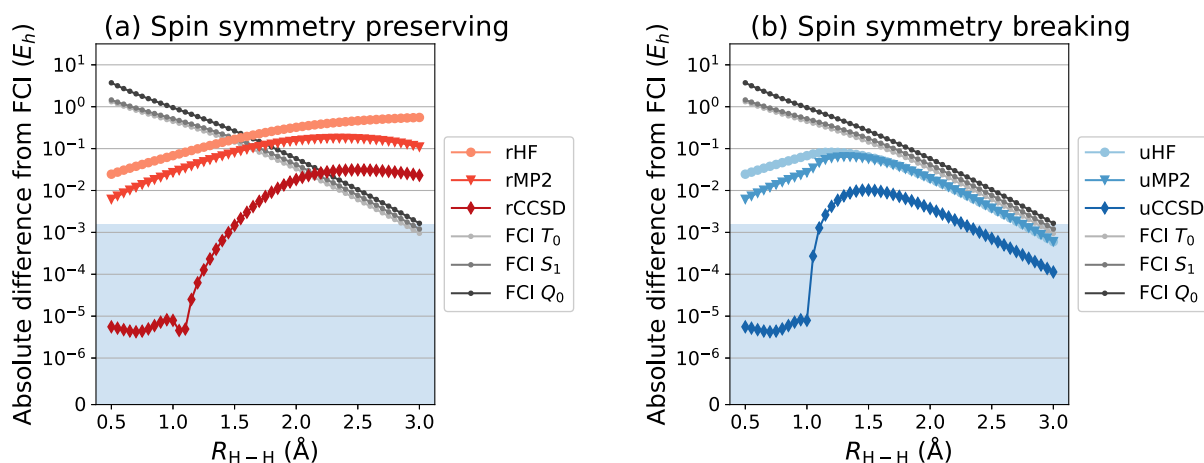
Figure 4. Absolute energy errors from FCI for HF, MP2, and CCSD versus the H−H separation presented alongside the energy gaps between the ground and next lowest singlet (S_1), triplet (T_0), and quintet (Q_0) excited states. HF spontaneously breaks $\hat{S}^2$ symmetry for $R_{H-H} > 1.00$ Å. The shaded area denotes "chemical accuracy" (errors less than 1 kcal mol^{-1}). (a) Reference state is restricted HF (rHF). (b) Reference state is unrestricted HF (uHF).

a broken-symmetry state, which is an equal superposition of the ground and first-excited states

$$|\psi^{BS}\rangle = \frac{1}{\sqrt{2}}(|\psi_0^{FCI}\rangle + |\psi_1^{FCI}\rangle). \qquad (22)$$

The energy error associated with this state is

$$
\begin{aligned}
\Delta E^{BS} &= \langle\psi^{BS}|\hat{H}|\psi^{BS}\rangle - E_0 \\
&= \frac{1}{2}(E_1 + E_0) - E_0 \\
&= \frac{1}{2}(E_1 - E_0)
\end{aligned}
\qquad (23)
$$

Also the infidelity of this state is

$$
\begin{aligned}
\text{Infidelity}(|\psi_0^{FCI}\rangle, |\psi^{BS}\rangle) &= 1 - |\langle\psi_0^{FCI}|\psi^{BS}\rangle|^2 \\
&= \frac{1}{2}.
\end{aligned}
\qquad (24)
$$

To explain this behavior, the Néel reference state $|10101010\rangle$ and its complement $|01010101\rangle$ are considered. The exact solution has equal weights for these states.[89] As K/J becomes large, the Hamiltonian more strongly penalizes states with occupation at consecutive sites. Because the operator pool includes only single and double excitations, a single operator cannot enact the quadruple excitation required to go between the reference state and its complement. The weight of the complementary state in the ADAPT-VQE wavefunction therefore is generated via products of multiple lower-rank excitation operators, putting it out of reach for a single pool operator. ADAPT-VQE(local/Néel) first touches the complementary Néel state after four operator additions. Despite having access to this determinant, ADAPT-VQE does not have the variational flexibility to significantly weigh this state as doing so would consequently weigh higher-energy intermediate determinants, raising the energy. As ADAPT-VQE continues to add operators, additional excitation pathways begin to form, although the VQE subroutine keeps the weight on the complementary state small. With the addition of the 52nd operator, ADAPT-VQE achieves the variational flexibility to substantially increase the weight of the complementary Néel state. This significant change in the

character of the ADAPT-VQE state is reflected in subplots 3(g), 3(h), and 3(i): the energy begins to significantly decrease again, the gradient norm jumps, and the infidelity drops.

The suppression of these pathways leads to a deeper gradient trough with increasing K/J. To further explain this suppression of the operator gradient, we consider eq 6 with the ADAPT-VQE state being expressed in terms of the eigenstates $|\psi_i^{FCI}\rangle$ of $\hat{H}$

$$|\psi^{(n)}\rangle = \sum_i c_i^{(n)}|\psi_i^{FCI}\rangle \qquad (25)$$

$$\frac{\partial E}{\partial\theta_k} = \sum_i \sum_j c_i^{(n)*}c_j^{(n)}\langle\psi_i^{FCI}|[\hat{H}, \hat{A}_k]|\psi_j^{FCI}\rangle \qquad (26)$$

$$= \sum_{ij} c_i^{(n)*}c_j^{(n)}(E_i - E_j)\langle\psi_i^{FCI}|\hat{A}_k|\psi_j^{FCI}\rangle. \qquad (27)$$

The energy difference term here is seen to suppress the gradients when a contaminant state and the target state become close in energy. As K/J becomes large, the gap between the exact ground and first-excited states shrinks, suppressing the gradients in this regime when the first-excited state is a major contaminant in the ADAPT-VQE trial state, as in the cases with broken-symmetry reference states.

For even stronger correlation ($K/J = 100$), ADAPT(BS-HF) and ADAPT(local/Néel) become fatally trapped in a gradient trough (see Supporting Information). This can be seen as suppression of the operator gradient below the tolerance of the numerical noise of the VQE optimizer. As such, these methods with broken-symmetry reference states retain high infidelities and energy errors. We speculate that, with a numerically exact optimizer, ADAPT-VQE should be able to escape even these gradient troughs, although this would not be possible for a quantum computer with finite noise.

The emergence of deep gradient troughs is not seen for ADAPT(local/cat$_+$), even for large K/J. ADAPT(HF) exhibits a shallow gradient trough at the start of the ADAPT-VQE procedure. In this case, the g-symmetry reference state is a superposition of the ground state and excited states with g symmetry. The high infidelity (0.98) of the initial state indicates severe contamination. The symmetry of the operator pool made from symmetry-preserving HF orbitals, however, prevents
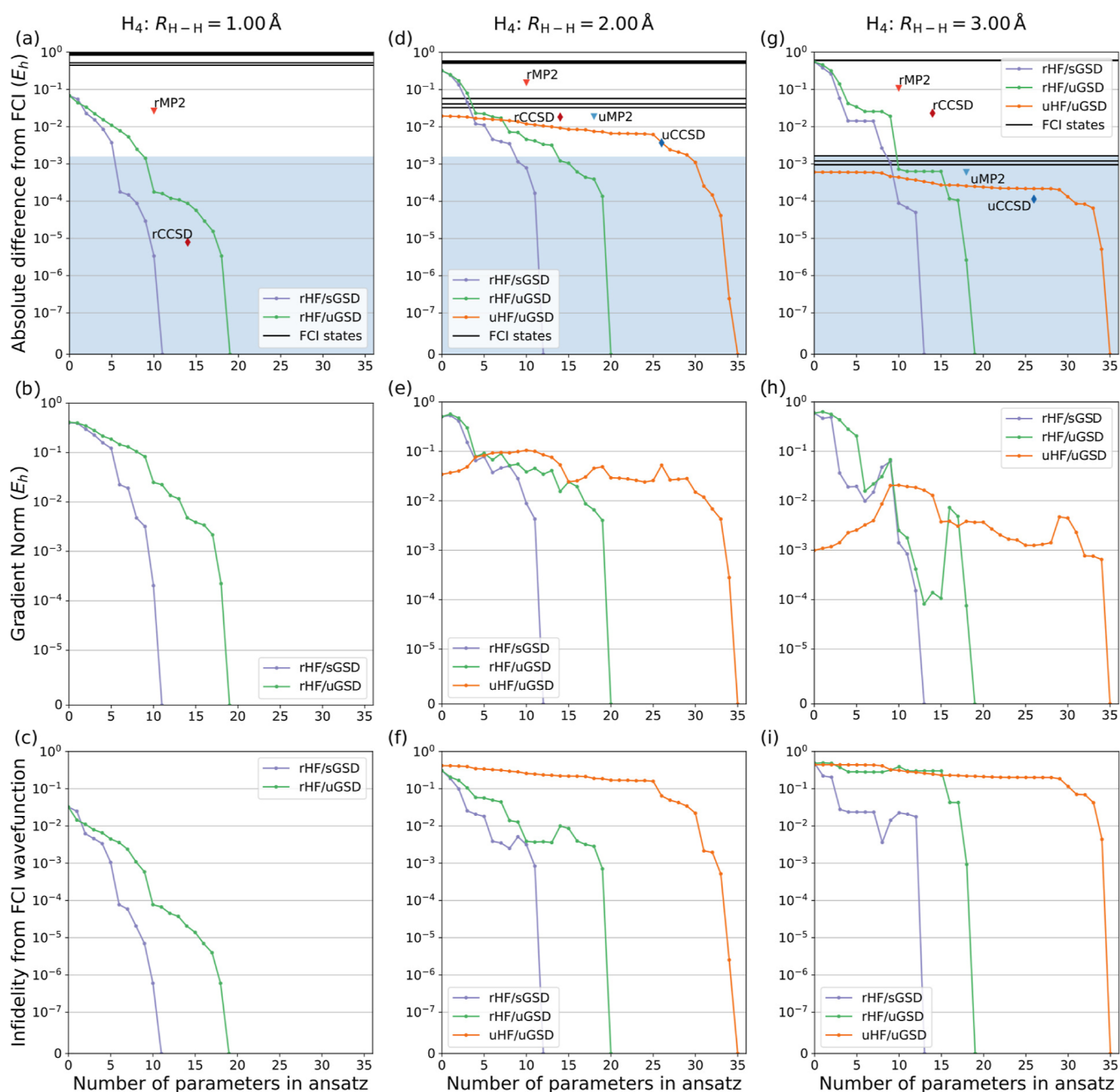
Figure 5. ADAPT-VQE results for the symmetric dissociation of linear H_4. Absolute energy errors [(a), (d), and (g)], ADAPT-VQE gradient norms [(b), (e), and (h)], and infidelities relative to the exact wavefunction [(c), (f), and (i)] as the ADAPT-VQE ansätze grow are presented for $R_{H-H} = 1.00$ Å, $R_{H-H} = 2.00$ Å, and $R_{H-H} = 3.00$ Å. The ADAPT-VQE methods with restricted HF orbitals, reference states, and singlet GSD operator pools (rHF/sGSD); restricted HF orbitals, reference states, and unrestricted GSD pools (rHF/uGSD); and unrestricted HF orbitals, reference states, and unrestricted GSD operator pools (uHF/uGSD) are presented. Absolute energy errors for the classical methods and the energy gaps between the ground state and the nine lowest excited states are presented alongside the ADAPT-VQE errors.

contamination from the low-lying, u-symmetry first-excited state. This restriction is seen to limit the depth of the trough. These results highlight the importance of symmetry in avoiding these deep gradient troughs.

4.2. H_4. In Figure 4 we present the absolute errors for a series of "classical" quantum chemistry methods. Figure 4(a) presents restricted HF (rHF), MP2, and CCSD on top of this reference (rMP2 and rCCSD, respectively) and the energy gaps between the FCI singlet ground state and the next lowest singlet (S_1), triplet (T_0), and quintet (Q_0) FCI excited states for the dissociation of H_4. These excited states become degenerate with

the ground state at the dissociation limit. CCSD yields a lower energy than the FCI ground state for values of $R_{H-H} > 1.05$ Å, yielding a kink in the absolute energy errors.

For $R_{H-H} = 0.90$ Å, near the equilibrium geometry, rMP2 improves upon the rHF reference error by more than a factor of 2 ($2 \times 10^{-2} E_h$ vs $6 \times 10^{-2} E_h$), while rCCSD improves upon the reference by nearly four orders of magnitude. A reduction in the error of rMP2 over rHF occurs in the intermediate regime, with the rMP2 energy peaking at $R_{H-H} = 2.45$ Å before ultimately diverging at the asymptotic limit. Beginning at the onset of the non-variational behavior, the improvement of rCCSD over rHF

absolute errors decreases with increasing H−H separation to just over one order of magnitude ($-0.03\ E_h$ vs $0.5\ E_h$) at R_{H-H} = 2.50 Å, the point at which the rCCSD energy begins to "turn up". The rCCSD energies are seen to be "chemically accurate" (errors less than 1 kcal mol^{-1}) for $R_{H-H} \leq 1.50$ Å. For large H−H separations, the rHF, rMP2, and rCCSD absolute errors exceed the energy gaps between the FCI ground state and the T_0, S_1, and Q_0 states.

Figure 4(b) presents the absolute errors for unrestricted HF (uHF), MP2, and CCSD on top of this reference (uMP2 and uCCSD, respectively), and the energy gaps between the FCI singlet ground state and the next lowest singlet (S_1), triplet (T_0), and quintet (Q_0) FCI excited states for the dissociation of H_4. The uHF references spontaneously break $\hat{S}^2$ symmetry for R_{H-H} > 1.00 Å. The kinks observed in the uMP2 and uCCSD errors at these points correspond to the change in the underlying HF reference. The uCCSD energy errors remain positive for all R_{H-H} being surveyed.

Unlike in the restricted case, the improvement of uMP2 over uHF becomes negligible as the H−H separation increases. The improvement of uCCSD over uHF decreases from over two orders of magnitude at the onset of symmetry breaking (R_{H-H} = 1.05 Å; $3. \times 10^{-4} E_h$ vs $7 \times 10^{-2} E_h$) to a factor of 5 at R_{H-H} = 3.00 Å ($6 \times 10^{-4} E_h$ vs $1 \times 10^{-4} E_h$). The uCCSD energies are chemically accurate for $R_{H-H} \leq 1.10$ Å and $R_{H-H} \geq 2.30$ Å. Both uHF and uMP2 energies are chemically accurate for $R_{H-H} \geq$ 2.75 Å. The errors for uHF, uMP2, and uCCSD fall below the energy gaps between the FCI ground and excited states for all H−H separations investigated.

We now investigate the role that spin symmetry plays in the compactness of the ADAPT-VQE ansätze for the symmetric dissociation of linear H_4. Figure 5 presents the absolute energy errors [(a), (d), and (g)], ADAPT-VQE gradient norms [(b), (e), and (h)], and infidelities from the exact wavefunction [(c), (f), and (i)] as the ansätze grow for ADAPT-VQE applied to the symmetric dissociation of H_4 at R_{H-H} = 1.00 Å, R_{H-H} = 2.00 Å, and R_{H-H} = 3.00 Å. The ADAPT-VQE methods surveyed are ADAPT-VQE using rHF orbitals, rHF reference states, and singlet GSD operator pools [ADAPT(rHF/sGSD)]; ADAPT-VQE using rHF orbitals, rHF reference states, and unrestricted GSD operator pools [ADAPT(rHF/uGSD)]; and ADAPT-VQE using uHF orbitals, uHF reference states, and unrestricted GSD operator pools [ADAPT(uHF/uGSD)].

4.2.1. Spin Symmetry Breaking Slows Energy Convergence. Comparing panels 5(a), 5(d), and 5(g), we see that ADAPT-(rHF/sGSD) converges to the exact solution with 11, 12, and 13 parameters for R_{H-H} = 1.00 Å, R_{H-H} = 2.00 Å, and R_{H-H} = 3.00 Å, respectively. The use of the sGSD operator pool ensures that the ADAPT-VQE state remains an eigenstate of $\hat{S}^2$. Although there are 20 determinants in the basis of rHF orbitals that contribute to the exact ground state, by enforcing spin symmetry, ADAPT(rHF/sGSD) is able to converge to the exact solution with as few as 11 parameters. For R_{H-H} = 2.00 Å and R_{H-H} = 3.00 Å, ADAPT(rHF/sGSD) converges to local minima when 11 operators have been added, and the addition of one and two additional operators, respectively, increases the variational flexibility of the ADAPT-VQE state and allows it to recover the global minimum. For ADAPT(rHF/uGSD), the uGSD operator pool contains operators that break $\hat{S}^2$ symmetry. Therefore, despite beginning with the rHF reference state, as the ADAPT(rHF/uGSD) ansatz grows, the $\langle \hat{S}^2 \rangle$ expectation value is

seen to deviate from 0. Without the efficient parameterization offered by the sGSD operator pool, ADAPT(rHF/uGSD) requires at least 19 parameters to converge to the exact solution of the 20-dimensional subspace. For ADAPT(uHF/uGSD), the use of uHF orbitals breaks the symmetries between the α and β orbitals, and thus, there are 36 unrestricted determinants (corresponding to all $\hat{S}_z$-preserving determinants) that contribute to the exact ground state. In this case, the full 36-dimensional Hilbert subspace must be spanned to achieve convergence to the exact ground state, requiring at least 35 parameters.

On comparing ADAPT(rHF/sGSD) and ADAPT(rHF/uGSD), the use of the $\hat{S}^2$-preserving operator pool not only accelerates convergence to the exact ground state but is also seen to require fewer parameters to achieve chemical accuracy for all H−H separations surveyed. After the onset of symmetry breaking at the HF reference state (beginning near R_{H-H} = 1.05 Å), ADAPT(uHF/uGSD) initially outperforms ADAPT-(rHF/sGSD) and ADAPT(rHF/uGSD) by virtue of a more energetically favorable reference state. Despite this, the more efficient parameterization offered by preserving symmetries allows both ADAPT(rHF/sGSD) and ADAPT(rHF/uGSD) to outperform ADAPT(uHF/uGSD) after the addition of only a few operators. For R_{H-H} = 2.00 Å, these cross-overs occur before ADAPT(uHF/uGSD) has achieved chemical accuracy, whereas for R_{H-H} = 3.00 Å, ADAPT(uHF/uGSD) achieves chemical accuracy before ADAPT(rHF/sGSD) and ADAPT(rHF/uGSD) as the uHF reference state is already chemically accurate. Despite this, ADAPT(uHF/uGSD) shows very limited improvement in error as more operators are added, and it has not significantly improved upon the reference energy when the cross-overs with ADAPT(rHF/sGSD) and ADAPT(rHF/uGSD) are reached.

For R_{H-H} = 1.00 Å, the rHF reference state provides a reasonable zeroth-order description of the system, and as such, rCCSD is able to provide a competitive performance to ADAPT(rHF/uGSD) with the same number of parameters despite the lack of connected triple and quadruple excitations. As the H−H separation increases, the rHF reference state becomes a poorer description of the true ground state, as seen by the initial infidelities [panels 5(c), 5(f), and 5(i)]. Here, these excitations become more important in accurately describing the ansatz, and as such, the performance of rCCSD is seen to suffer relative to ADAPT(rHF/uGSD) for the same number of parameters.

4.2.2. Spin Symmetry Breaking Worsens Gradient Troughs. For large H−H separations, all three of the ADAPT-VQE methods surveyed exhibit gradient troughs, as is evident in Figure 5(h). These gradient troughs are accompanied by a flattening of the energy error curve [Figure 5(g)] and large infidelities [Figure 5(i)].

The rHF and uHF reference states in this geometry have high infidelities with respect to the exact ground state, (0.482 and 0.442, respectively), indicating that these trial ground states contain significant contributions from excited FCI states. Additionally, the lowest singlet (S_1), lowest triplet (T_0), and lowest quintet (Q_0) excited states lie close in energy to the ground state. Recalling eq 27, we see that the ADAPT-VQE gradients are suppressed when the overlap between the ADAPT-VQE state and the target state is small (small $c_0^{(n)}$) and when the contaminant states are close in energy to the target state [small $(E_j - E_0)$].

For ADAPT(rHF/sGSD), the rHF reference state is a singlet, and as such, all states contributing to it are singlet in nature. By enforcing the ADAPT-VQE trial state to be a singlet, the effect of the sGSD operator pool is to limit the possible contaminant states. This results in a gradient trough that is relatively shallow, and ADAPT(rHF/sGSD) acquires the variational flexibility to escape the trough by adding only a few operators. Although ADAPT(rHF/uGSD) utilizes the same singlet rHF reference state as ADAPT(rHF/sGSD) and as such begins with only singlet contaminant states, the use of the uGSD operator pool introduces contaminant states of higher spin multiplicities as the ADAPT-VQE procedure proceeds, beginning with the second operator addition. This is evidenced by the initial growth of the $\langle \hat{S}^2 \rangle$ expectation value and can be understood as a variational conversion of higher-energy, singlet contaminant states to lower-energy, higher-spin-multiplet states. The uHF reference state has an $\langle \hat{S}^2 \rangle$ expectation value of 1.996, indicating significant contamination from the T_0 and Q_0 excited states. ADAPT(rHF/uGSD) and ADAPT(uHF/uGSD) are both seen to exhibit two gradient troughs. Escaping each of these gradient troughs is accompanied by a drop in the $\langle \hat{S}^2 \rangle$ expectation value. In the case of ADAPT(uHF/uGSD), this can be understood as ADAPT-VQE acquiring the variational flexibility to project out contamination from Q_0, corresponding to a drop in $\langle \hat{S}^2 \rangle$ from ~2 to ~0.6, and subsequently acquiring the variational flexibility to project out contamination from T_0, corresponding to a drop in $\langle \hat{S}^2 \rangle$ from ~0.6 to 0 at convergence to the exact ground state.

5. CONCLUSIONS

In this work, we have investigated two strongly correlated systems that exhibit two different kinds of spontaneous symmetry breaking at the mean-field level as correlation increases. In each case, we explore the role that breaking/preserving these symmetries in the reference states, operator pools, and representations of the Hamiltonian has on the performance of ADAPT-VQE.

Although reducing symmetry through the use of UHF orbitals often improves the energy accuracy of classical electronic structure theory methods, the use of broken-symmetry HF solutions is a detriment to ADAPT-VQE. For fermionic operator pools without symmetry adaptation of the operators, the symmetry (or lack thereof) of the pools is determined by the symmetries of underlying orbitals. With the onset of symmetry breaking in the MO basis, the number of determinants with nonzero weights in the expansion of the exact ground state increases significantly. In order to create the exact ground state, each of the determinants contributing to the exact ground state requires the addition of an operator to the ansatz. Thus, the use of symmetry-broken HF as a reference for ADAPT-VQE, although improving the energy of the reference, leads to much larger exact ansätze compared to symmetry-preserving HF/rHF.

In the local representation of the Hamiltonian, the underlying site orbital basis is inherently symmetry-broken, and as such, the representation of the exact ground state in the determinant basis made of the site orbitals is dense. For these systems, the use of operator pools that are not symmetry adapted again requires a larger number of operators to converge ADAPT-VQE. This is the case whether the reference state is symmetrized (cat$_+$) or broken symmetry (Néel).

Symmetries can be introduced to the operator pool by changing the underlying orbital basis (SALC) or via symmetry

adaptation of the pool operators (using the sGSD pool for H$_4$). In both cases, the preservation of these symmetries leads to shorter ansätze with ADAPT-VQE. In the former, transformation of the site basis yields a sparser representation of the exact ground state in the new orbital basis, and thus, an operator pool without symmetry adaptation using this orbital basis leads to ADAPT-VQE convergence with a smaller number of operators compared to those of the original site orbital basis. In the latter, the singlet GSD pool more efficiently spans the subspace of determinants that overlap with the exact ground state by parameterizing a symmetry-adapted combination of fermionic operators with a single parameter.

With respect to the issue of gradient troughs in ADAPT-VQE, we make the following observations:

1. Gradient troughs appear when excited states become close in energy to the ground state, such as the large K/J limit of the fermionized, anisotropic Heisenberg model or in the limit of large H–H separation in linear H$_4$.
2. Reference states that have a high fidelity with the exact ground state do not exhibit deep gradient troughs (cat$_+$), while the reference states with low fidelities are seen to exhibit them when low-lying excited states are present.
3. For systems where the reference state is symmetry preserving, the use of symmetry-adapted operator pools leads to shallow troughs [ADAPT(HF) and ADAPT(rHF/sGSD)], while symmetry-agnostic operator pools can lead to deep gradient troughs when the overlap with the exact ground state is low [ADAPT(rHF/uGSD)]. Using symmetry-adapted operator pools limits the possible contaminant states in the ADAPT-VQE state to those that obey the symmetry in question, while symmetry-agnostic pools may introduce new contaminants into trial states that were not initially present.
4. For symmetry-broken reference states, the presence of deep gradient troughs is endemic [ADAPT(BS-HF), ADAPT(local/Néel), and ADAPT(uHF/uGSD)].

While preparing this article for publication, a relevant preprint by Tsuchimochi et al.[90] appeared that also looks at spin-symmetry breaking in ADAPT-VQE. In their work, they highlight the unfavorable behavior of the "spin-dependent fermionic operator pool" (unrestricted operator pools in this work) and spin-complemented operator pools to break $\hat{S}^2$ symmetry. The authors similarly find that spin-symmetry breaking leads to an increase in the quantum computational resources (both parameter counts and CNOT gates required) compared to their spin-projected ADAPT-VQE, which applies a spin projection operator to restore the $\hat{S}^2$ symmetry. They further apply this spin-projected ADAPT-VQE to the computation of molecular properties and geometry optimization.

■ APPENDIX A

Operator pools

For the fermionized, anisotropic Heisenberg Hamiltonians, we utilize a fermionic generalized singles and doubles (GSD) pool without symmetry adaptation of the fermionic operators.

$$\{\hat{A}_k\} = \{\hat{A}_q^p\} \cup \{\hat{A}_{rs}^{pq}\} \tag{A1}$$

$$\hat{A}_q^p = \hat{a}_p^\dagger \hat{a}_q - \hat{a}_q^\dagger \hat{a}_p \tag{A2}$$

$$\hat{A}_{rs}^{pq} = \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r - \hat{a}_r^\dagger \hat{a}_s^\dagger \hat{a}_q \hat{a}_p \tag{A3}$$

For this generalized excitation pool the indices p, q, r, s index molecular orbitals of unspecified occupation in the reference state.

For the symmetric dissociation of linear H_4, we utilize both a singlet spin-symmetry-adapted generalized singles and doubles (sGSD) pool and a spin-unrestricted generalized singles and doubles (uGSD) pool. The operators in the sGSD pool map between states with $\langle \hat{S}^2 \rangle = 0$ by taking spin-adapted linear combinations of the excitation/de-excitation operators.

$$\{^1\hat{A}_k\} = \{^1\hat{A}_q^p\} \cup \{^1\hat{A}_{rs}^{pq}\}. \tag{A4}$$

The generalized spin-adapted singles operators are given as follows

$$^1\hat{A}_q^p = \frac{1}{\sqrt{2}}(\hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\alpha} + \hat{a}_{p_\beta}^\dagger \hat{a}_{q_\beta}) - \text{h. c} \tag{A5}$$

where p, q, ... index the spatial orbitals and p_α and p_β index the spin orbitals with spatial function p and α and β spin functions, respectively. The generalized spin-adapted doubles operators are given as follows

$$\{^1\hat{A}_{rs}^{pq}\} = \{^1\hat{A}_{rr}^{pp}\} \cup \{^1\hat{A}_{rr}^{pq}\} \cup \{^1\hat{A}_{rs}^{pp}\} \cup \{^{1A}\hat{A}_{rs}^{pq}\}$$
$$\cup \{^{1B}\hat{A}_{rs}^{pq}\}, \tag{A6}$$

$$^1\hat{A}_{rr}^{pp} = \hat{a}_{p_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{r_\alpha} - \hat{a}_{r_\alpha}^\dagger \hat{a}_{r_\beta}^\dagger \hat{a}_{p_\beta} \hat{a}_{p_\alpha}, \tag{A7}$$

$$^1\hat{A}_{rr}^{pp} = \frac{1}{\sqrt{2}}(\hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{r_\alpha} + \hat{a}_{q_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{r_\alpha}) - \text{h. c. }; \ p \neq q, \tag{A8}$$

$$^1\hat{A}_{rs}^{pp} = \frac{1}{\sqrt{2}}(\hat{a}_{p_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{s_\beta} \hat{a}_{r_\alpha} + \hat{a}_{p_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{s_\alpha}) - \text{h. c. }; \ r \neq s, \tag{A9}$$

$$^{1A}\hat{A}_{rs}^{pq} = \frac{1}{2\sqrt{3}}(2\hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\alpha}^\dagger \hat{a}_{s_\alpha} \hat{a}_{r_\alpha} + 2\hat{a}_{p_\beta}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{s_\beta} \hat{a}_{r_\beta}$$
$$+ \hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{s_\beta} \hat{a}_{r_\alpha}$$
$$- \hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{s_\alpha} - \hat{a}_{q_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{s_\beta} \hat{a}_{r_\alpha} + \hat{a}_{q_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{s_\alpha}) - \text{h}$$
$$. \text{ c. }; \ p \neq q, \ r \neq s, \tag{A10}$$

$$^{1B}\hat{A}_{rs}^{pq} = \frac{1}{2}(\hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{s_\beta} \hat{a}_{r_\alpha} + \hat{a}_{p_\alpha}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{s_\alpha}$$
$$+ \hat{a}_{q_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{s_\beta} \hat{a}_{r_\alpha} + \hat{a}_{q_\alpha}^\dagger \hat{a}_{p_\beta}^\dagger \hat{a}_{r_\beta} \hat{a}_{s_\alpha}) - \text{h. c. }; \ p \neq q, \ r$$
$$\neq s. \tag{A11}$$

The operators in the uGSD pool do not enforce $\hat{S}^2$ symmetry; however, they are constructed to preserve $\hat{S}_z$ symmetry

$$\{^u\hat{A}_k\} = \{\hat{A}_{q_\sigma}^{p_\sigma}\} \cup \{\hat{A}_{r_\sigma s_{\sigma'}}^{p_\sigma q_{\sigma'}}\}, \tag{A12}$$

where

$$\hat{A}_{q_\sigma}^{p_\sigma} = \hat{a}_{p_\sigma}^\dagger \hat{a}_{q_\sigma} - \hat{a}_{q_\sigma}^\dagger \hat{a}_{p_\sigma}, \tag{A13}$$

$$\hat{A}_{r_\sigma s_{\sigma'}}^{p_\sigma q_{\sigma'}} = \hat{a}_{p_\sigma}^\dagger \hat{a}_{q_{\sigma'}}^\dagger \hat{a}_{s_{\sigma'}} \hat{a}_{r_\sigma} - \hat{a}_{r_\sigma}^\dagger \hat{a}_{s_{\sigma'}}^\dagger \hat{a}_{q_{\sigma'}} \hat{a}_{p_\sigma}, \tag{A14}$$

and $\sigma, \sigma' \in \{\alpha, \beta\}$. Unlike the sGSD operator pool, the uGSD operator pool does not assume that the spin orbitals p_α and p_β have the same spatial functions, as is the case for molecular orbitals from a restricted HF reference but not necessarily the case for molecular orbitals from an unrestricted HF.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00709.

> Data tables for figures in the main text, plots of energy errors, parameter counts, infidelities for different K/J and R_{H-H}, and plot of spin-symmetry breaking for H_4 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

　　Nicholas J. Mayhall — *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States;* ⓞ orcid.org/0000-0002-1312-9781; Email: nmayhall@vt.edu

Authors

　　Luke W. Bertels — *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States;* ⓞ orcid.org/0000-0003-1776-2266

　　Harper R. Grimsley — *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States;* ⓞ orcid.org/0000-0001-5672-468X

　　Sophia E. Economou — *Department of Physics, Virginia Tech, Blacksburg, Virginia 24061, United States*

　　Edwin Barnes — *Department of Physics, Virginia Tech, Blacksburg, Virginia 24061, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.2c00709

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864.

(2) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133.

(3) Fock, V. Näherungsmethode zur Lósung des quantenmechanischen Mehrkörperproblems. *Z. Phys.* **1930**, *61*, 126−148.

(4) Dirac, P. A. M. Note on exchange phenomena in the Thomas atom. *Math. Proc. Camb. Philos. Soc.* **1930**, *26*, 376−385.

(5) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618−622.

(6) Coester, F.; Kümmel, H. Short-range correlations in nuclear wave functions. *Nucl. Phys.* **1960**, *17*, 477−485.

(7) Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **1966**, *45*, 4256−4266.

(8) Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **1982**, *21*, 467−488.

(9) Aspuru-Guzik, A.; Dutoi, A. D.; Love, P. J.; Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **2005**, *309*, 1704−1707.

(10) Wack, A.; Paik, H.; Javadi-Abhari, A.; Jurcevic, P.; Faro, I.; Gambetta, J. M.; Johnson, B. R.Quality, speed, and scale: Three key attributes to measure the performance of near-term quantum computers, **2021** arXiv 2110.14108v2 .

(11) Underwood, D.; Stehlik, J.; Phung, T.; Zajac, D.; Raftery, J. J.; Kumph, M.Gate error models for superconducting qubit architectures. *Bull. Am. Phys. Soc.*2021.

(12) Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2018**, *2*, 79.

(13) 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− 7.

(14) McClean, J. R.; Romero, J.; Babbush, R.; Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **2016**, *18*, 023023.

(15) 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.

(16) Rubin, N. C.; Babbush, R.; McClean, J. Application of fermionic marginal constraints to hybrid quantum algorithms. *New J. Phys.* **2018**, *20*, 053020.

(17) O'Gorman, B.; Huggins, W. J.; Rieffel, E. G.; Whaley, K. B.Generalized swap networks for near-term quantum computing, **2019**, 05118. arXiv preprint arXiv:1905.

(18) Gokhale, P.; Angiuli, O.; Ding, Y.; Gui, K.; Tomesh, T.; Suchara, M.; Martonosi, M.; Chong, F. T.Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families, **2019**, 13623. arXiv preprint arXiv:1907.

(19) Jena, A.; Genin, S.; Mosca, M.Pauli partitioning with respect to gate sets, **2019**, 07859. arXiv preprint arXiv:1907.

(20) Izmaylov, A. F.; Yen, T.-C.; Lang, R. A.; Verteletskyi, V. Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method. *J. Chem. Theory Comput.* **2019**, *16*, 190− 195.

(21) Izmaylov, A. F.; Yen, T.-C.; Ryabinkin, I. G. Revising the measurement process in the variational quantum eigensolver: is it possible to reduce the number of separately measured operators? *Chem. Sci.* **2019**, *10*, 3746−3755.

(22) Verteletskyi, V.; Yen, T.-C.; Izmaylov, A. F. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *J. Chem. Phys.* **2020**, *152*, 124114.

(23) Yen, T.-C.; Verteletskyi, V.; Izmaylov, A. F. Measuring all compatible operators in one series of single-qubit measurements using unitary transformations. *J. Chem. Theory Comput.* **2020**, *16*, 2400− 2409.

(24) Huggins, W. J.; McClean, J. R.; Rubin, N. C.; Jiang, Z.; Wiebe, N.; Whaley, K. B.; Babbush, R. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. *NPJ Quantum Inf* **2021**, *7*, 1−9.

(25) Huang, H.-Y.; Kueng, R.; Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **2020**, *16*, 1050−1057.

(26) Zhao, A.; Rubin, N. C.; Miyake, A. Fermionic partial tomography via classical shadows. *Phys. Rev. Lett.* **2021**, *127*, 110504.

(27) O'Malley, P. J. J.; et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **2016**, *6*, 031007.

(28) McClean, J. R.; Kimchi-Schwartz, M. E.; Carter, J.; De Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **2017**, *95*, 042308.

(29) Barkoutsos, P. K.; Gonthier, J. F.; Sokolov, I.; Moll, N.; Salis, G.; Fuhrer, A.; Ganzhorn, M.; Egger, D. J.; Troyer, M.; Mezzacapo, A.; Filipp, S.; Tavernelli, I. Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions. *Phys. Rev. A* **2018**, *98*, 022322.

(30) 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.

(31) Colless, J. I.; Ramasesh, V. V.; Dahlen, D.; Blok, M. S.; Kimchi-Schwartz, M. E.; McClean, J. R.; Carter, J.; de Jong, W. A.; Siddiqi, I. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **2018**, *8*, 011021.

(32) Lee, J.; Huggins, W. J.; Head-Gordon, M.; Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **2018**, *15*, 311−324.

(33) 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.

(34) Ryabinkin, I. G.; Yen, T.-C.; Genin, S. N.; Izmaylov, A. F. Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer. *J. Chem. Theory Comput.* **2018**, *14*, 6317− 6326.

(35) 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.

(36) 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.

(37) 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.

(38) Kutzelnigg, W. Error analysis and improvements of coupled-cluster theory. *Theor. chim. acta* **1991**, *80*, 349−386.

(39) Taube, A. G.; Bartlett, R. J. New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* **2006**, *106*, 3393−3401.

(40) Harsha, G.; Shiozaki, T.; Scuseria, G. E. On the difference between variational and unitary coupled cluster theories. *J. Chem. Phys.* **2018**, *148*, 044107.

(41) McClean, J. R.; Boixo, S.; Smelyanskiy, V. N.; Babbush, R.; Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **2018**, *9*, 4812−6.

(42) Holmes, Z.; Sharma, K.; Cerezo, M.; Coles, P. J. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum* **2022**, *3*, 010313.

(43) Arrasmith, A.; Holmes, Z.; Cerezo, M.; Coles, P. J. Equivalence of quantum barren plateaus to cost concentration and narrow gorges. *Quantum Sci. Technol.* **2022**, DOI: 10.1088/2058-9565/ac7d06.

(44) Dallaire-Demers, P.-L.; Wilhelm, F. K. Method to efficiently simulate the thermodynamic properties of the Fermi-Hubbard model on a quantum computer. *Phys. Rev. A* **2016**, *93*, 032303.

(45) 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*, 3007−9.

(46) Claudino, D.; Wright, J.; McCaskey, A. J.; Humble, T. S. Benchmarking adaptive variational quantum eigensolvers. *Front. Chem.* **2020**, *8*, 1152.

(47) Tang, H. L.; Shkolnikov, V. O.; Barron, G. S.; Grimsley, H. R.; Mayhall, N. J.; Barnes, E.; Economou, S. E. qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. *PRX Quantum* **2021**, *2*, 020310.

(48) Shkolnikov, V. O.; Mayhall, N. J.; Economou, S. E.; Barnes, E.Avoiding symmetry roadblocks and minimizing the measurement overhead of adaptive variational quantum eigensolvers, **2021**, 05340. arXiv preprint arXiv:2109.

(49) Yordanov, Y. S.; Armaos, V.; Barnes, C. H.; Arvidsson-Shukur, D. R. Qubit-excitation-based adaptive variational quantum eigensolver. *Commun. Phys.* **2021**, *4*, 1−11.

(50) Zhang, Z.-J.; Kyaw, T. H.; Kottmann, J. S.; Degroote, M.; Aspuru-Guzik, A. Mutual information-assisted adaptive variational quantum eigensolver. *Quantum Sci. Technol.* **2021**, *6*, 035001.

(51) Gomes, N.; Mukherjee, A.; Zhang, F.; Iadecola, T.; Wang, C.-Z.; Ho, K.-M.; Orth, P. P.; Yao, Y.-X. Adaptive variational quantum imaginary time evolution approach for ground state preparation. *Adv. Quantum Technol.* **2021**, *4*, 2100114.

(52) Barron, G. S.; Gard, B. T.; Altman, O. J.; Mayhall, N. J.; Barnes, E.; Economou, S. E. Preserving symmetries for variational quantum eigensolvers in the presence of noise. *Phys. Rev. Appl.* **2021**, *16*, 034003.

(53) Yordanov, Y. S.; Arvidsson-Shukur, D. R. M.; Barnes, C. H. W. Efficient quantum circuits for quantum computational chemistry. *Phys. Rev. A* **2020**, *102*, 062612.

(54) 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.

(55) Calzado, C. J.; Cabrero, J.; Malrieu, J. P.; Caballol, R. Analysis of the magnetic coupling in binuclear complexes. II. Derivation of valence effective Hamiltonians from ab initio CI and DFT calculations. *J. Chem. Phys.* **2002**, *116*, 3985−4000.

(56) Maurice, R.; Bastardis, R.; Graaf, C. d.; Suaud, N.; Mallah, T.; Guihéry, N. Universal Theoretical Approach to Extract Anisotropic Spin Hamiltonians. *J. Chem. Theory Comput.* **2009**, *5*, 2977−2984.

(57) Monari, A.; Maynau, D.; Malrieu, J.-P. Determination of spin Hamiltonians from projected single reference configuration interaction calculations. I. Spin 1/2 systems. *J. Chem. Phys.* **2010**, *133*, 044106.

(58) Malrieu, J. P.; Caballol, R.; Calzado, C. J.; de Graaf, C.; Guihéry, N. Magnetic interactions in molecules and highly correlated materials: physical content, analytical derivation, and rigorous extraction of magnetic Hamiltonians. *Chem. Rev.* **2014**, *114*, 429−492.

(59) Mayhall, N. J.; Head-Gordon, M. Computational quantum chemistry for single Heisenberg spin couplings made simple: Just one spin flip required. *J. Chem. Phys.* **2014**, *141*, 134111.

(60) Mayhall, N. J.; Head-Gordon, M. Computational quantum chemistry for multiple-site Heisenberg spin couplings made simple: Still only one spin-flip required. *J. Phys. Chem. Lett.* **2015**, *6*, 1982−1988.

(61) Mayhall, N. J. From model Hamiltonians to ab initio Hamiltonians and back again: Using single excitation quantum chemistry methods to find multiexciton states in singlet fission materials. *J. Chem. Theory Comput.* **2016**, *12*, 4263−4273.

(62) Coen de Graaf, C. d. G.; Ria Broer, R. B. *Magnetic Interactions in Molecules and Solids*; Springer, 2016.

(63) Abraham, V.; Mayhall, N. J. Simple rule to predict boundedness of multiexciton states in covalently linked singlet-fission dimers. *J. Phys. Chem. Lett.* **2017**, *8*, 5472−5478.

(64) Pokhilko, P.; Krylov, A. I. Effective Hamiltonians derived from equation-of-motion coupled-cluster wave functions: Theory and application to the Hubbard and Heisenberg Hamiltonians. *J. Chem. Phys.* **2020**, *152*, 094108.

(65) Kotaru, S.; Kähler, S.; Alessio, M.; Krylov, A. I. Magnetic exchange interactions in binuclear and tetranuclear iron (III) complexes described by spin-flip DFT and Heisenberg effective Hamiltonians. *ChemRxiv* **2022**, DOI: 10.1002/jcc.26941.

(66) Heisenberg, W. Zur Theorie des Ferromagnetismus. *Z. Phys.* **1928**, *49*, 619−636.

(67) Dirac, P. A. M. On the theory of quantum mechanics. *Proc. R. Soc. Lond. A* **1926**, *112*, 661−677.

(68) van Vleck, J. H. *The Theory of Electric and Magnetic Susceptibilities*; Clarendon Press: Oxford, 1932.

(69) Cleveland, C. L.; Medina, A. R. Obtaining a Heisenberg Hamiltonian from the Hubbard model. *Am. J. Phys.* **1976**, *44*, 44−46.

(70) Cai, Z. Resource estimation for quantum variational simulations of the Hubbard model. *Phys. Rev. Appl.* **2020**, *14*, 014059.

(71) Cade, C.; Mineh, L.; Montanaro, A.; Stanisic, S. Strategies for solving the Fermi-Hubbard model on near-term quantum computers. *Phys. Rev. B* **2020**, *102*, 235122.

(72) Bilkis, M.; Cerezo, M.; Verdon, G.; Coles, P. J.; Cincio, L. A semi-agnostic ansatz with variable structure for quantum machine learning, **2021**, 06712. arXiv preprint arXiv:2103.

(73) Van Dyke, J. S.; Barron, G. S.; Mayhall, N. J.; Barnes, E.; Economou, S. E. Preparing Bethe ansatz eigenstates on a quantum computer. *PRX Quantum* **2021**, *2*, 040329.

(74) Van Dyke, J. S.; Barnes, E.; Economou, S. E.; Nepomechie, R. I. Preparing exact eigenstates of the open XXZ chain on a quantum computer. *J. Phys. A Math. Theor.* **2022**, *55*, 055301.

(75) Selvarajan, R.; Sajjan, M.; Kais, S. Variational quantum circuits to prepare low energy symmetry states, **2021** 12857. arXiv preprint arXiv:2112.

(76) Gyawali, G.; Lawler, M. J. Adaptive variational preparation of the Fermi-Hubbard eigenstates. *Phys. Rev. A* **2022**, *105*, 012413.

(77) Jattana, M. S.; Jin, F.; De Raedt, H.; Michielsen, K. Assessment of the variational quantum eigensolver: Application to the Heisenberg Model, **2022**, 05065. arXiv preprint arXiv:2201.

(78) Wigner, E.; Jordan, P. Über das paulische äquivalenzverbot. *Z. Phys* **1928**, *47*, 631.

(79) McClean, J. R.; et al. OpenFermion: the electronic structure package for quantum computers. *Quantum Sci. Technol.* **2020**, *5*, 034014.

(80) The Locality of the Spin Hamiltonian in the Fermionic Representation Arises Only from the Fact that the Spin Hamiltonian Is Local and 1D, Which Cancels All the Z Strings that Enforce Antisymmetry.

(81) 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*, No. e1340.

(82) Sun, Q.; et al. Recent developments in the PySCF program package. *J. Chem. Phys.* **2020**, *153*, 024109.

(83) Purvis, G. D., III; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910−1918.

(84) Hehre, W. J.; Stewart, R. F.; Pople, J. A. Self-consistent molecular-orbital methods. i. Use of gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.* **1969**, *51*, 2657−2664.

(85) https://github.com/mayhallgroup/adapt-vqe, 2022.(Accessed on May 17, 2022).

(86) Virtanen, P.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261−272.

(87) Löwdin, P.-O. Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects. *Phys. Rev.* **1955**, *97*, 1509.

(88) Grimsley, H. R.; Barron, G. S.; Barnes, E.; Economou, S. E.; Mayhall, N. J. ADAPT-VQE is insensitive to rough parameter landscapes and barren plateaus, **2022**, 07179. arXiv preprint arXiv:2204.

(89) Although Asymptotically Large Values of K/J Admit Arbitrary Mixtures of These Two States, Any Finite Value Will Contributions from Other Configurations Which Will Fix the Relative Weight between These Configurations to Be Equal.

(90) Tsuchimochi, T.; Taii, M.; Nishimaki, T.; Ten-no, S. L. Adaptive construction of shallower quantum circuits with quantum spin projection for Fermionic systems, **2022**, arXiv preprint arXiv:2205.07097.