

Thermal Weight Determination and Interstate Coupling in State-Averaged ADAPT-VQE

Harper R. Grimsley* and Francesco A. Evangelista*

Cite This: *J. Chem. Theory Comput.* 2025, 21, 12557–12572

Read Online

ACCESS |



Metrics & More

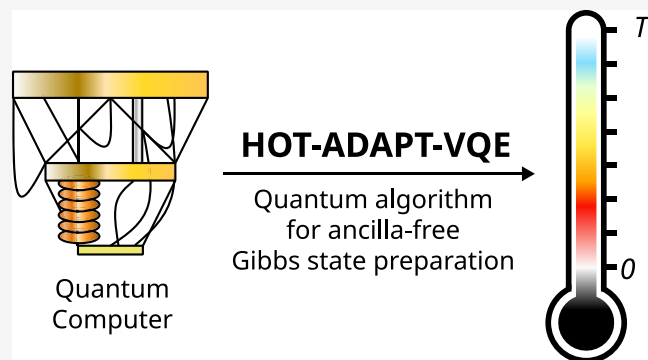


Article Recommendations



Supporting Information

ABSTRACT: Characterizing electronic thermal states at low temperatures is an important but challenging task in quantum chemistry and condensed matter physics, making it a prime candidate for a useful application in quantum computing. One of the most successful methods for state preparation on quantum computers is the Adaptive, Problem-Tailored (ADAPT) Variational Quantum Eigensolver (VQE), which has recently been generalized to treat excited states within a state-averaged framework as well as Gibbs states. In this work, we introduce Helmholtz-Optimized Thermal (HOT) ADAPT-VQE, an ancilla-free strategy for preparing Gibbs states that directly minimizes the Helmholtz free energy by targeting the dominant eigenstates of the thermal ensemble. We demonstrate the usefulness of HOT-ADAPT-VQE by predicting the free energy of two model systems with strongly correlated ground states: (1) the Fe^{2+} cation in a magnetic field and (2) a $[\text{Cu}_2\text{O}_7]^{10-}$ fragment of the Mott insulator La_2CuO_4 . Our results demonstrate that HOT-ADAPT-VQE significantly improves upon Gibbs-state estimates from multistate variants of ADAPT-VQE, often with substantially shallower quantum circuits, making it a promising candidate for thermal-state calculations.



1. INTRODUCTION

One of the most important applications of quantum chemistry is characterizing the ground- and excited-state electronic structures of molecules. While the description of individual states of chemical systems is important, chemical processes occur at finite temperatures, making it necessary in many situations to consider thermal ensembles. Understanding the effect of temperature on molecular systems is crucial in various applications, including the design of materials, the prediction of chemical reactions, and the interpretation of molecular spectra. Finite-temperature extensions of many classical methods have been proposed, including matrix product state methods,^{1–3} quantum Monte Carlo,^{4–6} dynamical mean-field theory,^{7,8} density functional theory,^{9,10} perturbation theory,^{11–13} coupled cluster theory,^{14–16} nonperturbative thermal cluster cumulant methods,^{17–20} and configuration interaction methods.^{21,22} However, the characterization of finite-temperature chemical systems with classical computers faces several challenges. Thermal states are largely difficult to predict for the same reasons as for individual excited states: They may display strong electron correlation, and the Hilbert or Fock space of all determinants is combinatorially large.²³ This difficult, relevant classical problem motivates a quantum solution.

Quantum chemistry has long been pursued as a potential application of quantum computers. Twenty years ago, the quantum phase estimation (QPE) algorithm²⁴ was applied in the context of electronic structure theory.²⁵ Since then,

substantial efforts have been directed toward the design of algorithms for the prediction of ground and excited electronic states of molecules.^{26–37} While less studied, various quantum methods exist for predicting thermal states. These usually follow one of two strategies: (1) variational minimization of a cost function^{38,39} or (2) sampling states obtained with quantum simulation of imaginary time evolution.^{31,40–42} Unfortunately, these methods are plagued by various issues such as deep circuits, ancilla qubit requirements, and difficulties in performing exact entropy measurements.

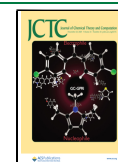
The first issue, circuit depth, is shared by all quantum algorithms during the ongoing “noisy, intermediate-scale quantum” (NISQ)⁴³ era. Deep quantum circuits correspond to lengthy time evolutions of the quantum registers. The longer the evolution, the more difficult it is to avoid decoherence of the qubits or buildup of gate errors. An early quantum algorithm proposed to address this problem was the variational quantum eigensolver (VQE).²⁶ In VQE, a correlated state is generated from a reference state $|\phi_0\rangle$ via a unitary operator $\hat{U}(\theta)$ that

Received: August 6, 2025

Revised: November 10, 2025

Accepted: November 14, 2025

Published: December 4, 2025



depends on a set of classical, real-valued parameters θ . The energy of the state $\hat{U}(\theta)|\phi_0\rangle$ is given as

$$E[\theta] = \langle \phi_0 | \hat{U}^\dagger(\theta) \hat{H} \hat{U}(\theta) | \phi_0 \rangle \quad (1)$$

where $\hat{H}$ is the Hamiltonian of the system. In VQEs, $\hat{U}(\theta)$ is designed such that it is a relatively short circuit, and $|\phi_0\rangle$ is easily prepared on a quantum register. These choices make measurements of $E[\theta]$ tractable on the NISQ hardware. The energy and wave function of the system are approximated by variationally minimizing $E[\theta]$ with respect to θ , using a classical optimization algorithm to update θ and the quantum computer to measure $E[\theta]$ and optionally its gradient with respect to the parameters θ .

The VQE algorithm is not uniquely defined due to the freedom in the choice of $\hat{U}(\theta)$. It was originally proposed to use the unitary (U) version of the coupled cluster singles and doubles (CCSD) ansatz⁴⁴

$$\hat{U}(\theta) = \exp \left(\sum_{ia} \theta_i^a (\hat{a}_i^a - \hat{a}_a^i) + \sum_{i < j, a < b} \theta_{ij}^{ab} (\hat{a}_{ij}^{ab} - \hat{a}_{ab}^{ij}) \right) \quad (2)$$

where the indices i, j (a, b) correspond to occupied (empty) spin-orbitals of the reference state. The operators $\hat{a}_i^a = \hat{a}_a^\dagger \hat{a}_i$ and $\hat{a}_{ij}^{ab} = \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i$ are single and double Fermionic excitations, respectively, expressed in terms of creation ($\hat{a}_p^\dagger$) and annihilation ($\hat{a}_q$) operators. Since it is impractical to implement the UCCSD unitary on quantum computers, a factorized (disentangled) form of UCCSD is employed instead. This disentangled form of UCC is an independent ansatz that can be systematically improved by adding higher-order excitations⁴⁵ and with an operator ordering dependence that leads to different variational solutions and corresponding energies.⁴⁶

In practice, many of the operators in the factorized UCCSD ansatz will frequently have little bearing on the energy accuracy attainable with the ansatz. Including superfluous operators increases the circuit depth, making standard factorized UCCSD inefficient.

The adaptive, problem-tailored (ADAPT-) VQE was introduced to address these issues for Fermionic ansätze.⁴⁷ The details of this algorithm will be addressed in Section 2, but the general idea is to start with a pool of simple anti-Hermitian generators $\{\hat{A}_\mu\}$, corresponding to primitive unitaries $\{\exp(\theta_\mu \hat{A}_\mu)\}$. The ansatz $\hat{U}(\theta)$ is iteratively constructed from these candidate unitaries based on greedy, first-order estimates of their ability to decrease the energy.

To obtain multiple states at once, the ansatz $\hat{U}(\theta)$ could be used to form a set of orthogonal many-body states by choosing a set of orthonormal references $\{|\phi_i\rangle\}$ and using the same unitary to form the set of correlated states $\{\hat{U}(\theta)|\phi_i\rangle\}$. Orthonormalization is ensured by the fact that for any pair of states, $\langle \phi_i | \hat{U}^\dagger(\theta) \hat{U}(\theta) | \phi_j \rangle = \langle \phi_i | \phi_j \rangle = \delta_{ij}$. Diagonalizing $\hat{H}$ in the basis of orthonormal states of this form can be used to obtain excited states in quantum methods. Moreover, state-averaged VQE methods exist, which optimize θ to minimize the average energy of a set of orthogonal states, such as the subspace-search (SS-)²⁸ and multistate, contracted (MC-)²⁹ VQE methods. VQEs based on a multideterminantal reference state have also been considered^{33,48–53} and, more recently, ADAPT-VQE has also

been generalized to use a state-averaged framework.^{54,55} In several state-averaged VQE methods, the user must define a set of reference states and specify the relative importance of the states by choosing a weight for the contribution of each state to the averaged energy. This formalism will be described more explicitly in Section 2 in the context of ADAPT-VQE. While the optimal choice of weights in state-averaged methods is a subject of ongoing research,^{56,57} the state-averaged ADAPT-VQE methods do not currently have a black box way of determining the weights of each state. The same open problem^{58–60} exists in traditional state-averaged complete-active-space self-consistent-field methods (SA-CASSCF),⁶¹ which simultaneously optimize the orbital basis and determinant coefficients of a multi-determinantal state.

ADAPT-VQE was first extended to thermal state preparation in the ADAPT-VQE-Gibbs method.^{39,62} This approach minimizes a proxy function for the Helmholtz free energy of a mixed state (ensemble) at a fixed temperature. Through purification, the mixed state is obtained as a partial trace of a pure state defined in a larger qubit space, necessitating the use of ancilla qubits. While the proxy function does not require entropy measurements, it still cannot be efficiently measured without approximation. Approximate estimates of the proxy objective function cause its minimum to no longer exactly correspond to that of the Helmholtz free energy.

In this work, we propose a different ADAPT-VQE strategy for thermal states that requires no ancilla, avoids problematic entropy measurements, and pursues the exact minimum of the Helmholtz free energy. As we will discuss in Section 2, rather than attempting to optimize a general mixed state in the 2^N Hilbert space represented by N qubits, we assume the special case that there are only a small number k of important references, which is realistic for low temperatures or systems with large excitation gaps. In making this assumption, we can design a unitary specifically to optimize the most important states in a physically informed way. We denote our method the Helmholtz-Objective Thermal ADAPT-VQE method (HOT-ADAPT-VQE). During preparation of this manuscript, Sambasivam et al.⁶³ introduced the Truncated Eigenvalue Parameterized Initial Density (TEPID-) ADAPT-VQE method, which, like our method, directly targets the Helmholtz free energy. While our proposed algorithm is similar to TEPID-ADAPT-VQE, it requires no ancilla qubits. As detailed in Section 2, HOT-ADAPT-VQE is also distinct in the use of a specific decoupling strategy. We explain in Section 2 how this decoupling allows us to reduce a naïvely k^2 -term density operator representation of the Gibbs state to only k terms without sacrificing accuracy. Finally, in this work, we extend the domain of application of thermal states beyond model Hamiltonians, focusing on ab initio Hamiltonians of two prototypical 3d transition metal systems: (1) the Fe^{2+} cation and (2) a cluster model of cuprates. We discuss the differences in the algorithms more extensively in Section 2.

2. THEORY

2.1. State-Averaged ADAPT-VQE. The most intuitive way to understand HOT-ADAPT-VQE is to begin from comparable multistate, temperature-agnostic alternatives. Two algorithms have been proposed to generalize the traditional ADAPT-VQE algorithm⁴⁷ to a multistate formalism that simultaneously optimizes a set of states. The first method is the state-averaged (SA-) ADAPT-VQE-SCF.⁵⁴ SA-ADAPT-VQE-SCF adaptively optimizes the molecular orbitals, as well as the ansatz itself.

Freezing the orbitals gives the simplified Multistate-Objective, Ritz-EigenSpectral (MORE-) ADAPT-VQE algorithm introduced by the authors.⁵⁵ For simplicity, we restrict our focus to MORE-ADAPT-VQE, ignoring the added benefits of orbital optimization.

The objective function to be minimized in MORE-ADAPT-VQE is the state-averaged energy (E_{MORE})

$$E_{\text{MORE}}[\theta] = \sum_i^k w_i \langle \phi_i | \hat{U}^\dagger(\theta) \hat{H} \hat{U}(\theta) | \phi_i \rangle \quad (3)$$

where $\{w_i\}_{i=1}^k$ is a set of k user-defined, positive weights which sum to 1, and $\{|\phi_i\rangle\}_{i=1}^k$ is a set of k user-specified, orthonormal references, typically chosen to be Slater determinants. We note in passing that one could consider the MORE-ADAPT-VQE objective function as the energy expectation value of an ensemble of states

$$E_{\text{MORE}}[\theta] = \text{Tr}(\hat{\rho}_{\text{MORE}}[\theta] \hat{H}) \quad (4)$$

where $\hat{\rho}_{\text{MORE}}[\theta]$ is a density matrix expressed in terms of state weights, reference states, and the unitary $\hat{U}(\theta)$ as

$$\hat{\rho}_{\text{MORE}}[\theta] = \sum_i^k w_i \hat{U}(\theta) |\phi_i\rangle \langle \phi_i| \hat{U}^\dagger(\theta) \quad (5)$$

To construct $\hat{U}(\theta)$, the user must specify a pool of anti-Hermitian generators $\{\hat{A}_\mu\}$, which is usually chosen to be universal for a pure state. In this work, we always use the universal pool of anti-Hermitian generalized single and double⁶⁴ excitations

$$\{\hat{A}_\mu\} = \{\hat{a}_p^q - \hat{a}_q^p\} \cup \{\hat{a}_{pq}^{rs} - \hat{a}_{rs}^{pq}\} \quad (6)$$

where the indices p, q, r , and s run over all spin-orbitals.

An ansatz $\hat{U}(\theta)$, parameterized by a vector θ of real numbers, is initialized to the trivial identity operator

$$\hat{U}^{(0)}(\theta) = \hat{\mathbf{I}} \quad (7)$$

To choose our first element from the pool, we consider each operator $\hat{A}_\mu$ as a candidate. For each candidate operator, we construct the candidate energy functional

$$E_{\text{candidate(SA)}}^\mu[\theta'] = \sum_i^k w_i \langle \phi_i | \hat{U}^\dagger(\theta) e^{-\theta' \hat{A}_\mu} \hat{H} e^{\theta' \hat{A}_\mu} \hat{U}(\theta) | \phi_i \rangle \quad (8)$$

where θ' is a new, undefined parameter. For each candidate energy, the gradient is computed with respect to the new parameter at $\theta' = 0$, giving a first-order estimate of the importance of the associated generator

$$\begin{aligned} g_\mu &= \left. \frac{\partial E_{\text{candidate(SA)}}^\mu[\theta']}{\partial \theta'} \right|_{\theta'=0} \\ &= \sum_i^k w_i \langle \phi_i | \hat{U}^\dagger(\theta) [\hat{H}, \hat{A}_\mu] \hat{U}(\theta) | \phi_i \rangle \end{aligned} \quad (9)$$

The operator $\hat{A}_\mu$ with the largest absolute gradient $|g_\mu|$ is then appended to the left side of $\hat{U}(\theta)$

$$\hat{U}^{(1)}(\theta) = e^{\theta_1 \hat{A}_\mu} \hat{\mathbf{I}} \quad (10)$$

where θ_1 has been appended to θ and $\hat{A}^{(1)} = \hat{A}_\mu$.

At this point, the state-averaged energy $E_{\text{MORE}}[\theta]$ (see eq 3) is optimized with a nonadaptive VQE, using $\hat{U}(\theta) = \hat{U}^{(1)}(\theta)$ as the ansatz.

After the VQE subroutine update of θ , another unitary operator is chosen based on its candidate gradient and appended to $\hat{U}(\theta)$. This process is repeated iteratively until some user-defined convergence criteria are achieved. One choice is to terminate the algorithm at some fixed number of operators in the ansatz. This is a natural choice in situations in which the main hardware limitation is an inability to implement the deeper circuits required for more operators. Another common choice for ADAPT-VQE methods is to terminate the algorithm when the norm (typically the L^2 or L^∞ norm) of the candidate gradients $\{g_\mu\}$ becomes sufficiently low. This choice is a more appropriate representation of situations where measurement error is the primary difficulty such that gradient resolution is the limiting factor of the hardware.

Once the convergence criteria are reached, a Ritz diagonalization of the Hamiltonian is performed in the basis of correlated reference states $\{\hat{U}(\theta) |\phi_i\rangle\}_{i=1}^k$, yielding k eigenstates $\{|\Psi_i\rangle\}$ and energies $\{E_i\}$. The corresponding eigenstates are given by

$$|\Psi_i\rangle = \sum_j^k \hat{U}(\theta) |\phi_j\rangle C_{ji} \quad (11)$$

where $\mathbf{C}$ is a $k \times k$ matrix of expansion coefficients. Due to the state-averaged formalism, $\hat{U}(\theta)$ is the same for all references. This allows the diagonalization to be written in terms of $\mathbf{C}$ and a diagonal matrix of the eigenenergies ($\mathbf{E}$)

$$\bar{\mathbf{H}} \mathbf{C} = \mathbf{C} \mathbf{E} \quad (12)$$

where the subspace Hamiltonian matrix $\bar{\mathbf{H}}$ has elements

$$\bar{\mathbf{H}}_{ij} = \langle \phi_i | \hat{U}^\dagger(\theta) \hat{H} \hat{U}(\theta) | \phi_j \rangle \quad (13)$$

The construction of $\bar{\mathbf{H}}$ does not require ancilla qubits since off-diagonal elements $\bar{\mathbf{H}}_{ij}$ can be obtained by performing measurements on the state $\frac{1}{\sqrt{2}} \hat{U}(\theta) (|\phi_i\rangle + |\phi_j\rangle)$ as discussed in refs 28 and 29. Note that when $k = 1$, there is only one reference, and MORE-ADAPT-VQE simplifies to a traditional ADAPT-VQE calculation.

Both SA-ADAPT-VQE-SCF and MORE-ADAPT-VQE rely on weights chosen by the user. The optimal choice of weights is a complicated and open problem,^{56,57} so a method that determines the weights for the user is an attractive feature. This partially motivates the HOT-ADAPT-VQE algorithm along with the desired application of predicting thermal states.

2.2. HOT-ADAPT-VQE. To generalize MORE-ADAPT-VQE to a system at a finite temperature, we begin by defining a variational energy functional whose minimum corresponds to a thermal state. For a system of N electrons at temperature T , the canonical ensemble Gibbs state $\hat{\rho}$ minimizes the Helmholtz free energy $\mathcal{F}$, given by

$$\mathcal{F} = U - k_B T S \quad (14)$$

where k_B is the Boltzmann constant. The internal energy U , not to be confused with the unitary operator $\hat{U}(\theta)$ which defines the ansatz, is the Hamiltonian expectation value, written in terms of Hamiltonian eigenstates $\{|\Psi_i\rangle\}$ and corresponding probabilities $\{p_i\}$

$$U = \text{Tr}(\hat{\rho}\hat{H}) = \sum_i^{\|\mathbb{H}\|} p_i \langle \Psi_i | \hat{H} | \Psi_i \rangle \quad (15)$$

where $\|\mathbb{H}\|$ indicates the dimension of the Hilbert space $\mathbb{H}$. The von Neumann entropy S is given as

$$S = -\text{Tr}(\hat{\rho} \log \hat{\rho}) = -\sum_i p_i \log p_i \quad (16)$$

The probability for the system to occupy the microstate $|\Psi_i\rangle$ follows from minimizing the free energy and is given by the Boltzmann distribution

$$p_i = \frac{1}{Z} e^{-\beta(E_i - E_0)} \quad (17)$$

Here, we have introduced the inverse temperature, $\beta = 1/k_B T$, and the canonical partition function

$$Z = \sum_i e^{-\beta(E_i - E_0)} \quad (18)$$

which ensures that the probabilities sum to one.

MORE-ADAPT-VQE generalizes quite naturally to the prediction of thermal states due to its connection through the density matrix representation to a weighted statistical ensemble of states. In eq 5, we have already seen that one can use ADAPT-VQE to minimize the ensemble energy. To extend MORE-ADAPT-VQE to a thermal state, we approximate the canonical Gibbs state as

$$\hat{\rho} \approx \hat{\rho}_{\text{HOT}} = \sum_i^k p_i |\Psi_i\rangle \langle \Psi_i| \quad (19)$$

This form of $\hat{\rho}_{\text{HOT}}$ contains two assumptions. The first one is to restrict the sum over eigenstates to a set of k low-energy states $\{|\Psi_i\rangle\}_{i=1}^k$. For low temperatures, most of the eigenstates of the Hamiltonian with $E_i \gg E_0$ are not significantly populated and can be neglected. This assumption is also used in TEPID-ADAPT-VQE. The second assumption is that the k excited states $\{|\Psi_i\rangle\}_{i=1}^k$ can be represented, as in eq 11, as linear combinations of correlated references $\{\hat{U}(\theta)|\phi_i\rangle\}_{i=1}^k$ for an adaptively constructed ansatz $\hat{U}(\theta)$, where expansion coefficients $\mathbf{C}$ are obtained by solving eq 12. The HOT-ADAPT-VQE expression for $\hat{\rho}_{\text{HOT}}$ generalizes the expression of $\hat{\rho}_{\text{MORE}}$ (see eq 5) by replacing the reference states $|\phi_i\rangle$ with the set of multideterminantal states $\{\sum_{j=1}^k |\phi_j\rangle C_{ji}\}$. The additional flexibility of the latter ansatz is expected to lead to more accurate results for small values of k . As pointed out in ref 52, it is difficult to design circuits that implement arbitrary multideterminantal references of the form $\sum_{j=1}^k |\phi_j\rangle C_{ji}$. A significant advantage of the HOT-ADAPT-VQE method is that it does not require preparing complicated multireference states on a quantum register, as demonstrated later in this section. Note that HOT-ADAPT-VQE is guaranteed to converge to the exact finite temperature density matrix in the limit of $k \rightarrow \|\mathbb{H}\|$ since in this limit the basis of states $\{|\Psi_i\rangle = \sum_{j=1}^{\|\mathbb{H}\|} \hat{U}(\theta)|\phi_j\rangle C_{ji}\}$ can be exact eigenstates, irrespective of $\hat{U}(\theta)$ due to the availability of a sufficient number of parameters in the expansion coefficient matrix, albeit with exponential scaling. However, as discussed below, there are several scenarios where accurate thermal properties can be obtained with only a small number of low-

energy states (small k), such as low-temperature regimes, systems with large excitation gaps, or impurity problems, where thermal effects are localized to a small subsystem.

Introducing the definition of $\hat{\rho}_{\text{HOT}}$ into the Helmholtz free energy functional leads to the following expression for $\mathcal{F}$ in terms of the ensemble probabilities ($\{p_i\}$), expansion coefficients ($\{C_{ji}\}$), and unitary $[\hat{U}(\theta)]$

$$\mathcal{F}[\hat{U}(\theta)] = \sum_{ijl}^k p_i C_{ji} C_{li}^* \langle \phi_l | \hat{U}^\dagger(\theta) \hat{H} \hat{U}(\theta) | \phi_j \rangle + \frac{1}{\beta} \sum_i p_i \log p_i \quad (20)$$

While we have omitted their dependence on θ from the equations, the ensemble weights and expansion coefficients appearing in HOT-ADAPT-VQE are not truly independent variational parameters since, given $\hat{U}(\theta)$, they can be computed noniteratively from eqs 17 and 12, respectively. Throughout this work, we also assume that the orbitals are kept frozen. Therefore, all variational conditions involve the parameters in the HOT-ADAPT-VQE ansatz, the Boltzmann probabilities, and the expansion coefficients. The first term on the r.h.s. of eq 20 requires the measurement of k^2 elements of the subspace Hamiltonian $\bar{\mathbf{H}}$. As mentioned in the previous subsection, this can be accomplished without ancilla qubits using the method of ref 34. The second term (entropy) requires no additional quantum measurements to compute since the probabilities $\{p_i\}$ can be determined by classically diagonalizing $\bar{\mathbf{H}}$ and using the eigenvalues to predict the Boltzmann distribution. Having an efficient way to measure $\mathcal{F}[\hat{U}(\theta)]$, we now have what we need to build an ancilla-free finite-temperature VQE. For a general, fixed choice of the structure of $\hat{U}(\theta)$, we denote the process HOT-VQE and implement it in a one-step fashion, where $\{p_i\}$ and $\mathbf{C}$ are assigned their optimal values based on θ for any given free energy or gradient calculation according to eqs 12 and 17.

To create an adaptive variant of HOT-VQE, our last task is to identify a criterion for choosing which operator to add to the unitary. Following the approach used in TEPID-ADAPT-VQE, it is natural to use the single-parameter gradients of $\mathcal{F}$ in analogy with eq 9. We begin by turning eq 20 into a candidate energy functional for operator $\hat{A}_\mu$

$$\mathcal{F}_{\text{candidate}}^\mu[\theta'] = \sum_{ijl}^k p_i C_{ji} C_{li}^* \langle \phi_l | \hat{U}^\dagger(\theta) e^{-\theta' \hat{A}_\mu} \hat{H} e^{\theta' \hat{A}_\mu} \hat{U}(\theta) | \phi_j \rangle + \frac{1}{\beta} \sum_i p_i \log p_i \quad (21)$$

By always keeping $\mathcal{F}$ optimized (within the constraints of normalization and $\sum_i p_i = 1$) with respect to $\{p_i\}$ and $\mathbf{C}$, its gradient is greatly simplified. Because the optimal ensemble weights and CI coefficients correspond to a variational solution, $\mathcal{F}$ must have vanishing first derivatives with respect to both $\{p_i\}$ and $\mathbf{C}$. These conditions correspond to $\{p_i\}$ satisfying the Boltzmann distribution and $\mathbf{C}$ diagonalizing the subspace Hamiltonian. In particular, we note that the stationarity of $\mathcal{F}$ with respect to $\{p_i\}$ implies that there is no contribution from the entropy in gradients, so that while $\mathcal{F} \neq U$, we still have $\frac{\partial \mathcal{F}}{\partial \theta} = \frac{\partial U}{\partial \theta}$. This is actually a troubling feature of this free energy functional despite the simplified gradient. As we will discuss in Section 4, this can manifest in the form of premature convergence due to

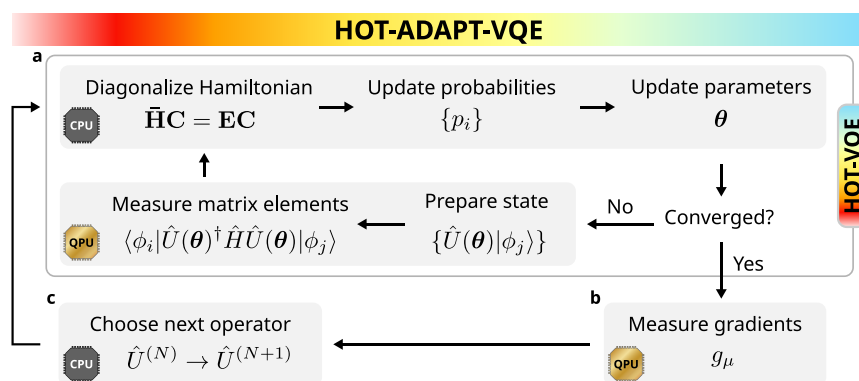


Figure 1. HOT-ADAPT-VQE procedure. (a) In the HOT-VQE subroutine, the CPU (central processing unit) uses classical optimization techniques to minimize the free energy $\mathcal{F}$ with respect to the coefficient matrix $\mathbf{C}$, ensemble weights $\{p_i\}$, and parameters θ , for a fixed unitary ansatz $\hat{U}^{(k)}(\theta)$ containing a product of k unitaries. The QPU (quantum processing unit) is responsible for preparing states and measuring the Hamiltonian matrix elements and gradients with respect to the parameters. (b) Once the HOT-VQE-VQE subroutine is converged, the QPU is used to measure the free energy gradient (g_μ) for all candidate operators. (c) The next operator is chosen to maximize the absolute gradient ($|g_\mu|$) and added to $\hat{U}^{(k)}(\theta)$, altering the ansatz to $\hat{U}^{(k+1)}(\theta)$. The procedure continues by entering the HOT-ADAPT-VQE subroutine to optimize the updated ansatz. The algorithm is terminated when all user-specified criteria are met.

the depopulation of excited states. Following these considerations, the gradient expression simplifies to

$$g_\mu = \left. \frac{\partial \mathcal{F}_{\text{candidate}}^\mu[\theta']}{\partial \theta'} \right|_{\theta'=0} = \sum_{ijl} p_i C_{ji} C_{li}^* \langle \phi_l | \hat{U}^\dagger(\theta) [\hat{H}, \hat{A}_\mu] \hat{U}(\theta) | \phi_j \rangle \quad (22)$$

Choosing the operator $\hat{A}_\mu$ with the largest $|g_\mu|$ gives us our prescription for ansatz construction. The overall algorithm is shown graphically in Figure 1 and in pseudocode in Alg. 1. Note that since in HOT-ADAPT the Boltzmann probabilities are generally unequal, the energy expression cannot be simplified by using the trace invariance property as done in MORE-ADAPT.

For an ensemble $\hat{\rho}_{\text{HOT}}$ of the form in eq 19, we have described how to obtain the individual energies $\{\langle \Psi_i | \hat{H} | \Psi_i \rangle\}$, from which the quantities U , S , and $\mathcal{F}$ can all be obtained through trivial classical calculations. The ensemble expectation value of any other observable $\hat{O}$ can be computed in the same fashion

$$\text{Tr}(\hat{\rho}_{\text{HOT}} \hat{O}) = \sum_i^k p_i \langle \Psi_i | \hat{O} | \Psi_i \rangle \quad (23)$$

where $\{\Psi_i\}$ is the set of eigenstates of $\hat{H}$.

We briefly remark on the energy additivity of HOT-ADAPT-VQE for noninteracting fragments represented by product states (size consistency). Consider the case where the Hamiltonian for a system A , $\hat{H}_A$, has a well-converged HOT-ADAPT-VQE solution

$$\hat{\rho}_{\text{HOT}}^A = \sum_i^k p_i^A |\Psi_i^A\rangle \langle \Psi_i^A| \quad (24)$$

where $|\Psi_i^A\rangle$ is the exact i th lowest-energy eigenstate of $\hat{H}_A$. As a simple example, let us consider the situation where all the energies are degenerate with the value E_A . It follows that all the ensemble weights $\{p_i^A\}$ will be equal to $\frac{1}{k}$ and the corresponding free energy will be equal to $\mathcal{F}_A$. Suppose a noninteracting copy of

A is introduced (which we refer to as B). In that case, the corresponding Hamiltonian will be $\hat{H} = \hat{H}_A + \hat{H}_B$ and the free energy should be equal to $2\mathcal{F}_A$ with the ensemble given by the product state $\hat{\rho}_{\text{HOT}}^A \otimes \hat{\rho}_{\text{HOT}}^B$, with $\hat{\rho}_{\text{HOT}}^B = \hat{\rho}_{\text{HOT}}^A$. Let us consider the free energy for the case where we include only some number of eigenstates k' for the tensor product Hilbert space $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. The value of k' corresponds to the total number of eigenstates being used to describe a mixed state in the enlarged Hilbert space, not the number of additional eigenstates relative to k . Because the energy of each state for each subsystem is E_A , and there is no coupling between the two subspaces, the energy of each eigenstate in the product space is $2E_A$. Additionally, since the eigenstates within the product space are all degenerate, the probabilities will be uniform in a Boltzmann distribution, and because we are only using k' eigenstates, the probability of each one in our ensemble will be $p_i = \frac{1}{k'}$. Plugging into the definition of the Helmholtz free energy, we have a supersystem energy $\mathcal{F}_S$, given as

$$\mathcal{F}_S = \sum_i^{k'} \frac{1}{k'} 2E + \frac{1}{\beta} \sum_i^{k'} \frac{1}{k'} \log \frac{1}{k'} \quad (25)$$

$$= 2E + \frac{1}{\beta} \log \frac{1}{k'} \quad (26)$$

$$= 2E + 2 \frac{1}{\beta} \sum_i^k \frac{1}{k} \log \frac{1}{k} - 2 \frac{1}{\beta} \sum_i^k \frac{1}{k} \log \frac{1}{k} + \frac{1}{\beta} \log \frac{1}{k'} \quad (27)$$

$$= 2\mathcal{F}_A - 2 \frac{1}{\beta} \log \frac{1}{k} + \frac{1}{\beta} \log \frac{1}{k'} \quad (28)$$

$$= 2\mathcal{F}_A + \frac{1}{\beta} (2 \log k - \log k') \quad (29)$$

$$= 2\mathcal{F}_A + \frac{1}{\beta} \log \frac{k^2}{k'} \quad (30)$$

Note that we reintroduced the original k in order to describe the supersystem free energy $\mathcal{F}_S$ in terms of the subsystem free energy and the number of eigenstates k used for the subsystem. It follows that HOT-ADAPT-VQE is not generally size-consistent unless the second term in the final free energy expression, $\log \frac{k^2}{k'}$ is null, a condition met when the number of supersystem references scales quadratically with system size ($k' = k^2$).

To illustrate this point, a simple toy example can be constructed by considering an H_2 molecule in the STO-3G basis with a bond distance of 1 Å and computing the lowest 4 FCI energies. Using these four states to form an optimized Gibbs state gives a free energy of $\mathcal{F}_{H_2} = -5.229 E_h$ at $T = 1 \times 10^6$ K. It follows that for a Hamiltonian which is a tensor product of two isolated H_2 molecules (i.e., the two are infinitely separated), the free energy should be $-10.457 E_h$ for size consistency to be satisfied. Using only the 8 lowest-energy products of H_2 eigenstates, however, gives a free energy of $-8.437 E_h$. To obtain the full $-10.457 E_h$, 16 (4^2) lowest-energy product states must be used. It is worth noting that replacing any of the 16 lowest-energy product states with a state that is higher in energy also fails to give the $-10.457 E_h$ energy. This exemplifies that to satisfy energy additivity, it is necessary to simultaneously satisfy the condition $k' = k^2$ and include the correct k^2 states.

It should be noted that the previous situation assumes the worst-case scenario of a temperature sufficiently high that all eigenstates are relevant in both subsystems. Another more realistic situation is that of two coupled subsystems with very different energy scales, with the best case scenario being that at temperature T , subsystem A has k significantly populated eigenstates $\{|\Psi_i^A\rangle\}_{i=1}^k$, while subsystem B has only one low-lying eigenstate $|\Psi_0^B\rangle$. This scenario captures impurity problems in physics and chemistry, such as a high-spin transition metal surrounded by solvent molecules, where the first subsystem has far more relevant states than the second. In this situation, there will be only k significantly populated tensor product states $\{|\Psi_i\rangle_A \otimes |\Psi_0\rangle_B\}$ in the ensemble, such that k does not need to be increased at all to capture the Gibbs state of the supersystem. One could repeat this process, adding an additional copy of the second subsystem, e.g., adding another solvent molecule, but by the same reasoning, this would not increase the number of relevant tensor product states. This implies that as long as the thermal effects are localized on the original subsystem, k does not have to increase with the system size for HOT-ADAPT-VQE to prepare the exact ensemble, making the method relevant for a general class of impurity problems.

A toy example illustrating this idea is to consider a stretched H_2 molecule with an interatomic distance of 5 Å (subsystem A) and a weakly correlated H_2 molecule with an interatomic distance of 1 Å (subsystem B) at 1000 K, using a minimal basis (STO-3G). Computing the free energy from the lowest four exact eigenstates ($k = 4$), the results for the two individual subsystems are -0.938 and $-1.101 E_h$, respectively. Note that for the 1 Å geometry, this is just the ground-state energy since the energy levels are too far apart for excited states to be significantly populated. For the noninteracting composite system, an expansion based on the lowest four product states (built from the exact eigenstates) yields a free energy of $-2.039 E_h$, accurately recovering the exact sum of the free energies of the isolated subsystems.

A major difference between HOT- and TEPID-ADAPT-VQE is that the latter does not address the coupling between the correlated reference states $\{\hat{U}(\theta)|\phi_i\rangle\}$. That is, in analogy to the MORE-ADAPT-VQE method, it expresses its density matrix ($\hat{\rho}_{\text{TEPID}}$) as a sum of projectors onto the individual VQE states $\{\hat{U}(\theta)|\phi_i\rangle\}$

$$\hat{\rho}_{\text{TEPID}} = \sum_i p_i \hat{U}(\theta)|\phi_i\rangle\langle\phi_i|\hat{U}^\dagger(\theta) \quad (31)$$

HOT-ADAPT-VQE, however, uses projectors onto diagonalized states $\{|\Psi_i\rangle\}$ to form its density matrix $\hat{\rho}_{\text{HOT}}$. This means that the HOT-ADAPT-VQE ensemble comprises states which are already decoupled through $\hat{H}$, giving a potentially lower free energy for the same ADAPT-VQE ansatz $\hat{U}(\theta)$. The significance of this difference between the algorithms becomes apparent in the limit of $T = 0$. TEPID-ADAPT-VQE becomes equivalent to ADAPT-VQE, barring situations in which $\hat{U}(\theta)$ changes the energy orderings of the states $\{\hat{U}(\theta)|\phi_i\rangle\}$ as it is constructed. HOT-ADAPT-VQE, on the other hand, can still make use of its extra references by replacing $|\phi_0\rangle$ with a continuously optimized linear combination of all of the references, giving a more flexible ansatz, even at $T = 0$. One can more generally think about the flexibility of the two ansätze by considering the requirements to obtain an arbitrary k -term density matrix

$$\hat{\rho} = \sum_i^k p_i |\Psi_i\rangle\langle\Psi_i| \quad (32)$$

For TEPID-ADAPT-VQE to exactly describe this density matrix, it would have to obtain individual states $\hat{U}(\theta)|\phi_i\rangle$ such that

$$\{\hat{U}(\theta)|\phi_i\rangle\} = \{|\Psi_i\rangle\} \quad (33)$$

for all $i < k$. On the other hand, HOT-ADAPT-VQE requires only that

$$\text{span}\{\hat{U}(\theta)|\phi_i\rangle\} = \text{span}\{|\Psi_i\rangle\} \quad (34)$$

Because the k weights $\{p_i\}$ are not generally equal, the Helmholtz free energy is not trace-invariant, so that one will obtain different energies and gradients if the states $\{\hat{U}(\theta)|\phi_i\rangle\}$ are rotated among themselves.

To show this numerically, we calculated the operator addition gradients from HOT-ADAPT-VQE, with and without coupling, for a linear H_4 molecule at 1000 K. A uniform spacing of 2 Å was used between each pair of adjacent hydrogen atoms. Canonical orbitals were used, and three references were chosen: the Aufbau determinant and the singly excited HOMO–LUMO determinants ($|\phi_0\rangle$, $|\phi_1^2\rangle$, and $|\phi_1^{\bar{2}}\rangle$). Using the same Fermionic pool used later in this work, HOT-ADAPT-VQE chose the operator $\hat{a}_{00}^{33} - \hat{a}_{33}^{00}$ with gradient $0.27 E_h$. Disabling the coupling, i.e., forcing $\mathbf{C}$ to be the identity matrix, led the algorithm to choose the operator $\hat{a}_{1\bar{1}}^{22} - \hat{a}_{2\bar{2}}^{11}$ with gradient $0.25 E_h$. In the coupled case, the formation of the spin-adapted triplet configuration state function (CSF) allowed HOT-ADAPT-VQE to identify a different operator than it would have otherwise chosen.

3. COMPUTATIONAL DETAILS

Single-particle states (orbitals) used to represent many-body states are obtained either from Hartree–Fock computations

performed with the Psi4 software package⁶⁵ or from complete-active-space self-consistent-field computations performed with the FORTE plugin.⁶⁶ In ref 8, we provide JSON (JavaScript Object Notation) files containing the value of the scalar energy (V), one-electron integrals h_{pq} , and the antisymmetrized two-electron integrals $\langle pq||rs \rangle = \langle pqlrs \rangle - \langle pqlsr \rangle$ for the two systems considered in this work. These quantities are given on a spin-orbital basis, ordered in such a way that the spin-up and spin-down functions for the p th spatial function correspond to even ($2p$) and odd ($2p + 1$) indices. All quantum algorithms were simulated in the absence of noise using a local modification of the QFORTE package.⁶⁷ Simulations also did not include error due to finite sampling. We removed these error sources in order to consider the performance of the algorithm itself. We defer a more involved exploration of these topics to future work. This code is made available through local forks of the main FORTE and QFORTE repositories (see the Code Availability statement). In modeling molecular systems at finite temperatures, we partition the orbital basis into three sets: core (doubly occupied), active (with variable occupation), and virtual (unoccupied). The exact eigenstates and corresponding eigenenergies were obtained via complete-active-space configuration interaction (CASSCI), whereby we impose occupation restrictions consistent with our orbital partitioning to generate the full Hilbert space of N -electron determinants. All CASSCI computations use the corresponding ground-state orbitals optimized at the restricted open-shell Hartree–Fock (ROHF) or state-specific CASSCF levels. All exact thermodynamic quantities were then computed by ensemble averaging, as shown in eq 17. For each system considered, we report the values of temperature (T), internal energy (U), entropy (S), free energy ($\mathcal{F}$), and the average value of the square of the total spin operator ($\hat{S}^2$) in CSV (comma-separated values) files in ref 8.

In this work, the performance of HOT-ADAPT-VQE is compared to that of a finite-temperature variant of MORE-ADAPT-VQE. In the latter, we first optimize $\hat{U}(\theta)$ using equal weights for the k chosen references and then predict ensemble averages using a Boltzmann distribution of the decoupled eigenstates $\{|\Psi_i\rangle\}$ according to eq 17. Reported ensemble properties are then computed according to eqs 14–16 and 23. Note that because the same individual states are being used to form the MORE-ADAPT-VQE ensembles at different temperatures, only one MORE-ADAPT-VQE calculation needs to be performed for a given system. In the formal limit of an exact ansatz, HOT-ADAPT-VQE and MORE-ADAPT-VQE would yield the same thermal state. The advantage of HOT-ADAPT-VQE is therefore not in the exact limit but in converging faster, that is, producing accurate finite-temperature results with significantly shallower circuits by directly targeting the Helmholtz free energy.

Reported critical temperatures are computed as

$$T_c = \frac{1}{k_B} (E_N - E_0) \quad (35)$$

where E_0 is the energy of the ground state and E_N is the energy of an excited state, with respective ensemble weights p_0 and p_N . Through algebraic manipulation of eq 35, we see that we can define the critical temperature to be such that p_0 is greater than p_N by a factor of e .

We discuss the choice of references for each molecular system in its respective subsection, but in both cases, we choose determinants rather than CSFs for the sake of simplicity,

particularly given the high multiplicities involved. As mentioned previously, the anti-Hermitian generalized single and double (GSD) operator pool was used for all MORE- and HOT-ADAPT-VQE calculations. The Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm was used for parameter optimization in all VQE subroutines.⁶⁸ As our primary goal is to model finite-temperature effects, the systems considered in this work were modeled without the inclusion of an environment or counterion effects since the quantitative accuracy of our models is less important than the accuracy of HOT- and MORE-ADAPT-VQE in obtaining solutions for those models.

4. NUMERICAL SIMULATIONS

4.1. Iron Cation. As our first illustrative example, we consider a model of the low-lying electronic states of the Fe^{2+} cation. This system was chosen due to its simplicity and the variety of spin states, making it an excellent benchmark for investigating spin-dependent thermal population effects. An active space consisting of six electrons in six orbitals ($6e, 6o$) was used, comprising the Fe 3d and 4s atomic orbitals from an ROHF computation in the cc-pVDZ basis.^{69,70} In our ROHF computation, we target the ground ^5D state, optimizing a determinant with quantum numbers $S = 2$ and $M_S = 2$ and orbital occupation $3d_{z^2}^2 3d_{x^2-y^2}^2 3d_{xy}^1 3d_{yz}^1 3d_{xz}^1$. Due to the double occupancy of the $3d_{z^2}$ orbital, the ROHF solution breaks spherical symmetry, yielding nondegenerate orbital levels. This splitting is analogous to the presence of a weak crystal field or weak orbital interactions in a metal complex. These orbitals, together with their associated Hartree–Fock energies (in E_h), are depicted in Figure 2. To weakly split the degenerate quintet ground state, we include a magnetic field aligned with the z -axis in the effective Hamiltonian

$$\hat{H}_z = \hat{H} - 1 \times 10^{-4} \hat{S}_z \quad (36)$$

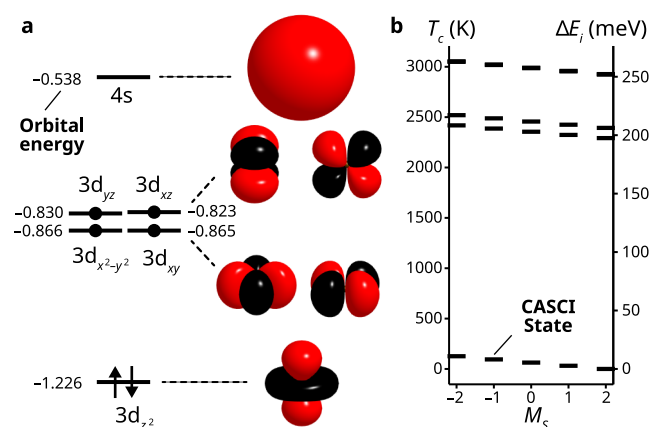


Figure 2. CASCI results for the Fe^{2+} cation using an active space containing six electrons in six orbitals. (a) Restricted open-shell Hartree–Fock (ROHF) orbitals used for the active space computations and a representative occupation pattern for the reference state. Orbital energies are shown in the absence of a magnetic field. The 3d orbitals obtained from ROHF are nondegenerate due to the lack of spherical symmetry (the $3d_{z^2}$ orbital being doubly occupied). (b) CASCI energy of the quintet states that arise from the atomic ^5D term. The highest two quintets in the plot are nearly degenerate, causing them to overlap at this energy scale. Orbital visualization was performed using the VMD software package.⁷¹

Table 1. References Used for the Fe²⁺ Cation MORE-/HOT-ADAPT-VQE Calculations

reference	determinant	reference	determinant
0	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\alpha 3d_{yz}^\alpha 3d_{xz}^\alpha\rangle$	8	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\alpha 3d_{yz}^\alpha 3d_{xz}^\beta\rangle$
1	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\alpha 3d_{yz}^\beta 3d_{xz}^\alpha\rangle$	9	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\alpha 3d_{yz}^\beta 3d_{xz}^\beta\rangle$
2	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\beta 3d_{yz}^\alpha 3d_{xz}^\alpha\rangle$	10	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\beta 3d_{yz}^\alpha 3d_{xz}^\beta\rangle$
3	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\beta 3d_{yz}^\beta 3d_{xz}^\alpha\rangle$	11	$ 3d_z^2 3d_{x^2-y^2}^\alpha 3d_{xy}^\beta 3d_{yz}^\beta 3d_{xz}^\beta\rangle$
4	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\alpha 3d_{yz}^\alpha 3d_{xz}^\alpha\rangle$	12	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\alpha 3d_{yz}^\alpha 3d_{xz}^\beta\rangle$
5	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\alpha 3d_{yz}^\beta 3d_{xz}^\alpha\rangle$	13	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\alpha 3d_{yz}^\beta 3d_{xz}^\beta\rangle$
6	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\beta 3d_{yz}^\alpha 3d_{xz}^\alpha\rangle$	14	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\beta 3d_{yz}^\alpha 3d_{xz}^\beta\rangle$
7	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\beta 3d_{yz}^\beta 3d_{xz}^\alpha\rangle$	15	$ 3d_z^2 3d_{x^2-y^2}^\beta 3d_{xy}^\beta 3d_{yz}^\beta 3d_{xz}^\beta\rangle$

The CASCI (exact diagonalization) energies of the lowest 25 states of the effective Hamiltonian are shown in Figure 2. Due to symmetry breaking, the 25-fold degenerate ⁵D ground state of Fe²⁺ artificially splits into five quintet states that span an energy range of ca. 260 meV (with the fourth and fifth groups being nearly degenerate). The field induces a 2.72 meV (100 μE_h) splitting within each group of the five components of the quintet state. Focusing on the thermodynamics of the lowest five states, this splitting corresponds to a critical temperature $T_c = 31.6$ K. To describe the Gibbs state within the range $0.5 \leq T < 5 T_c$, we consider as references for MORE- and HOT-ADAPT-VQE a set of 16 determinants (enumerated in Table 1) in which the $3d_z^2$ orbital is doubly occupied and the other 3d orbitals are singly occupied. These references span values of M_S ranging from -2 to 2 . The energy errors for MORE- and HOT-ADAPT-VQE are compared in Figure 3 for temperatures between 0.5 and $5 T_c$.

We begin our discussion of Figure 3 by analyzing the performance of HOT-ADAPT-VQE. For this problem, at iteration 0, HOT- and MORE-ADAPT-VQE yield the same energies and the errors are below chemical accuracy (1 mE_h). Therefore, we focus on the relative improvements brought by the unitary. As expected, the free energy converges monotonically with each iteration for any given temperature. Conversely, the internal energy and entropy contributions to the free energy do not improve monotonically, as HOT-ADAPT-VQE only tries to minimize their sum. At all temperatures, HOT-ADAPT-VQE is eventually unable to add additional operators after the ansatz includes 25 generators. This is the case except for $T = 2 T_c$, where an additional operator leads to a negligible decrease in the free energy. While premature convergence to a local minimum is possible, this problem is more pronounced at low temperatures, where the algorithm prematurely ignores important states with low weights in the early stages of the simulation. In this situation, it is more likely that the small error comes from the increasing population of the first excited quintet states in the CASCI solution at higher temperatures. For example, because the references used include only one $M_S = 2$ determinant, the lowest-lying state of the first excited quintet cannot be prepared with the M_S -conserving GSD operators. This is consistent with the fact that the internal energy and entropy values are both slightly below the CASCI values, suggesting that there are not enough states in the manifold to obtain the exact entropy, attributing a higher weight to the lower-energy states.

Compared with finite-temperature computations based on MORE-ADAPT-VQE states, HOT-ADAPT-VQE yields a

significant improvement in the number of operators required to obtain accurate free energies. This is consistent with the fact that the MORE-ADAPT-VQE energy functional does not prioritize the low-lying eigenstates of $\hat{H}$, leading to the selection of a sequence of unitaries that is not optimal to describe a finite-temperature state. This is illustrated quite dramatically in the fact that the internal energy, entropy, and free energy all become less accurate at 25 iterations of MORE-ADAPT-VQE than at 0 iterations. This is possible because MORE-ADAPT-VQE seeks to improve the unweighted average of the energies, which can cause the low-lying energies to increase as long as the high-lying energies decrease. This effect is shown in Figure 4, where a comparison of MORE-ADAPT-VQE at 0 and 25 iterations shows that the high-lying states (6–16) with low thermal populations improve at the expense of the important states (1–5). When the thermal ensemble is prepared, the high-lying energies will still be too high in energy to be significantly populated, but some of the low-lying energies will be much worse than they were at iteration 0. It is only at infinite temperature that equally weighted MORE-ADAPT-VQE is guaranteed to monotonically improve the internal energy U with each iteration. Given that the errors are greater than those of HOT-ADAPT-VQE, negative errors in U cannot be dismissed for MORE-ADAPT-VQE as coming from an insufficient number of references. It is possible, however, to obtain negative errors in U for MORE-ADAPT-VQE, despite the algorithm minimizing the unweighted average of individual state energies for a similar reason. If the relative energies between states are not accurately predicted, then the Boltzmann distribution will disproportionately populate states that are lower in energy, decreasing U and increasing S for the thermal ensemble, a phenomenon observed here. Variationality of the internal and free energies is guaranteed for both methods at $T = 0$.

The performance of MORE-ADAPT-VQE could be improved by introducing unequal weights to accelerate convergence or by using fewer references, albeit at the cost of reduced accuracy at convergence. Both of these strategies, however, would mean taking advantage of knowledge about the true Gibbs state, information that is not typically available a priori. Based on the iron cation results, HOT-ADAPT-VQE seems to be extremely promising for preparing approximate thermal ensembles at low temperatures. In addition to the choice of weights, HOT-ADAPT-VQE removes the ambiguity of whether to spin-adapt the references. This is because HOT-ADAPT-VQE can break or restore total spin symmetry during

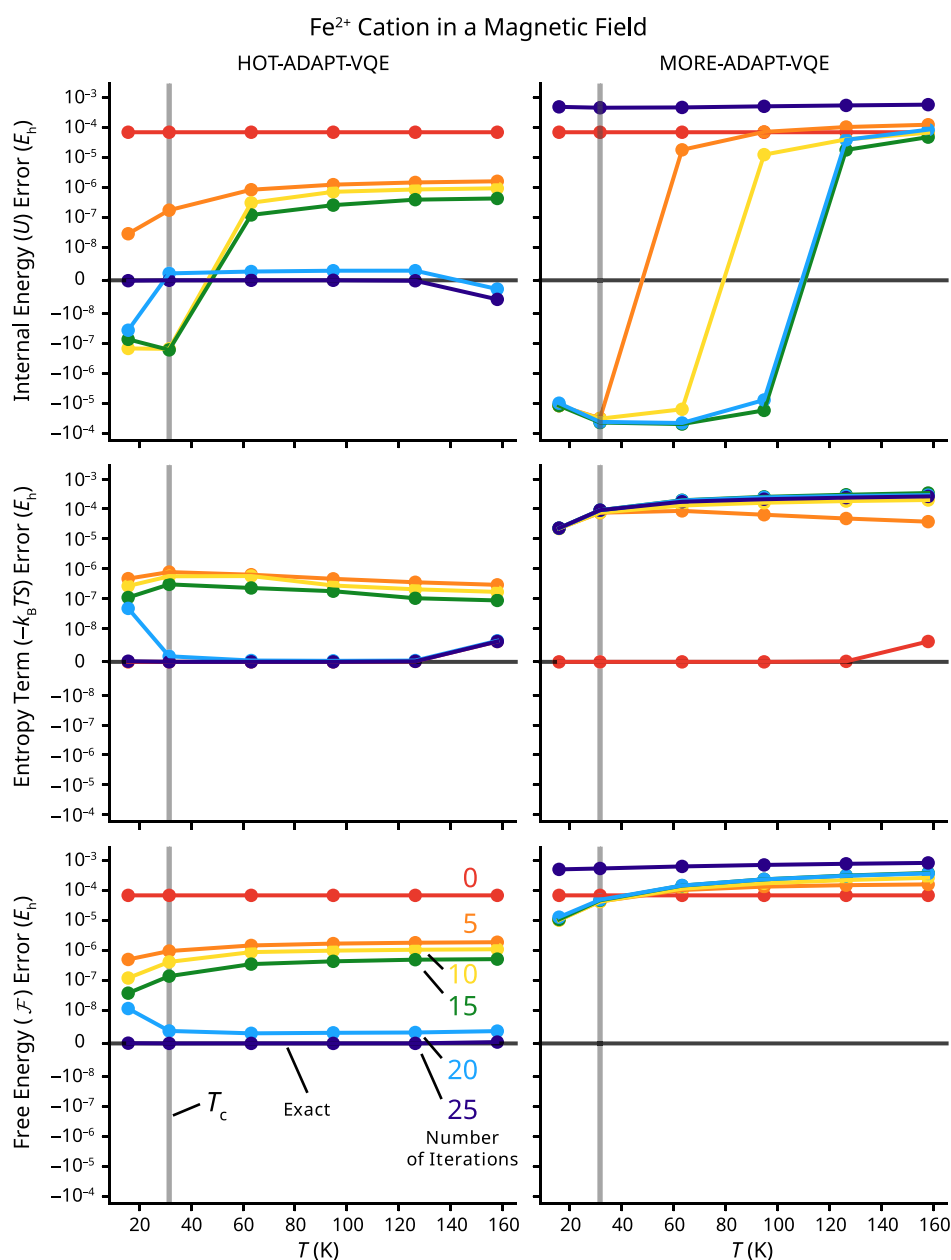


Figure 3. Plots of the error of energetic values from HOT-ADAPT-VQE (left) and MORE-ADAPT-VQE (right) at different iterations as a function of temperature for Fe^{2+} in a magnetic field. The x -axis is the temperature used for the two algorithms. From top to bottom, errors are reported in the internal energy U , the entropy term $-k_{\text{B}}TS$, and the free energy $\mathcal{F}$. The different line colors correspond to different ADAPT-VQE iterations. The iteration 0 and 25-parameter values of the entropy contribution overlap for HOT-ADAPT-VQE, making the red curve difficult to see. Note that the red curve will be the same for both HOT- and MORE-ADAPT-VQE since no operators have been chosen, and the energies in the Boltzmann distribution come from diagonalizing $\hat{H}$ within the same set of references.

the diagonalization of $\bar{H}$, ensuring that the best multi-determinantal basis is always used.

4.2. Cuprate Model. After examining the Fe^{2+} cation, we next applied HOT-ADAPT-VQE to a more challenging chemical system. A major potential application of finite-temperature methods is the study of electronic phase transitions, particularly in superconducting materials. An interesting example is the cuprate family, a class of high-temperature superconductors discovered in the late 1980s.^{72,73} Unlike conventional superconductors, the behavior of cuprates cannot be adequately described by the standard Bardeen–Cooper–Schrieffer (BCS)⁷⁴ theory of superconductivity.⁷⁵ These unconventional superconducting properties and a strong

superexchange mechanism in the CuO_2 planes⁷⁶ make these materials interesting from an electronic structure point of view. Characterizing the associated antiferromagnetic interaction represents a challenging problem for quantum chemistry methods, due to strong correlation and the need for a large active space.^{77,78} However, recent works have extended ab initio quantum chemistry methods to compute the properties of cuprates, shedding light on some properties of these materials not captured by effective low-energy models.^{79–81}

As a representative starting point, we consider a $[\text{Cu}_2\text{O}_7]^{10-}$ planar fragment of D_{2h} symmetry, pictured in Figure 5. This fragment models the essential local electronic structure within the CuO_2 planes of La_2CuO_4 , a classic parent material for high-

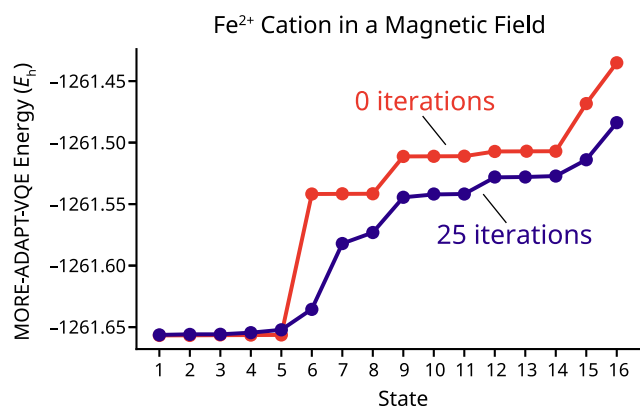


Figure 4. Plots of the energy of the MORE-ADAPT-VQE states for Fe^{2+} in a magnetic field at 0 and 25 iterations.

temperature cuprate superconductors.⁸² In the absence of doping, cuprate parent materials like La_2CuO_4 are Mott insulators;⁸³ i.e., conventional band-theory approaches incorrectly predict metallic behavior. Indeed, when modeling this system, we identify significant challenges for variational finite-temperature quantum algorithms when the reference states (the starting point for HOT-ADAPT-VQE) do not appropriately capture correlation effects. Specifically, the absence of references that contribute to the correlated singlet ground state leads to incorrectly biasing the optimization toward a triplet state. This qualitative discrepancy underscores one of the challenges of accurately capturing the physics of cuprates and motivates our choice of the $[\text{Cu}_2\text{O}_7]^{10-}$ fragment as a compelling benchmark test for quantum algorithms.

In our $[\text{Cu}_2\text{O}_7]^{10-}$ model, all Cu–O bond lengths were set to 1.90 Å, the experimental value for La_2CuO_4 .⁸⁴ An initial set of

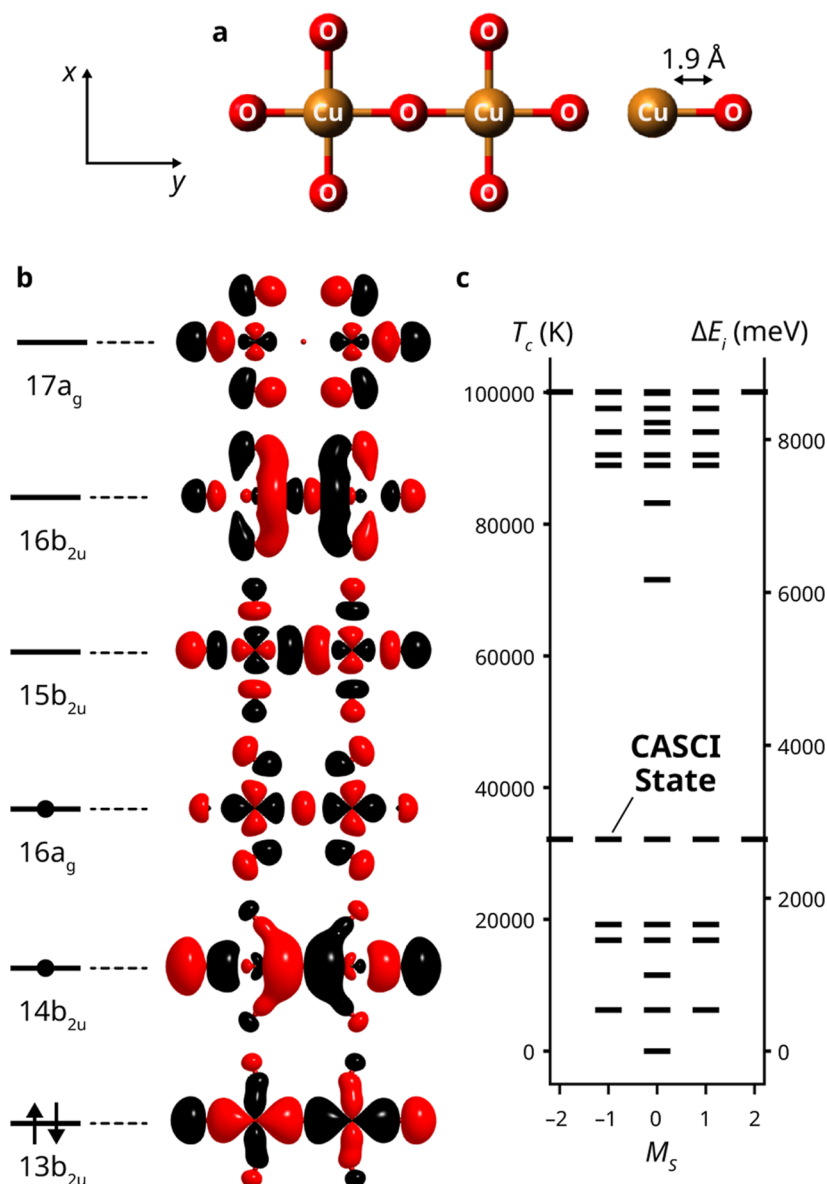


Figure 5. CASCI results for the $[\text{Cu}_2\text{O}_7]^{10-}$ anion using an active space containing four electrons in six orbitals. (a) Geometry of the planar $[\text{Cu}_2\text{O}_7]^{10-}$ anion studied in this work. All Cu–O distances are fixed to 1.90 Å. (b) State-specific CASSCF orbitals used for the active space computations and a representative occupation pattern for the leading determinants in the ground state. Orbital visualization was performed using the VMD software package.⁷¹ (c) The lowest 37 CASCI energies of the system.

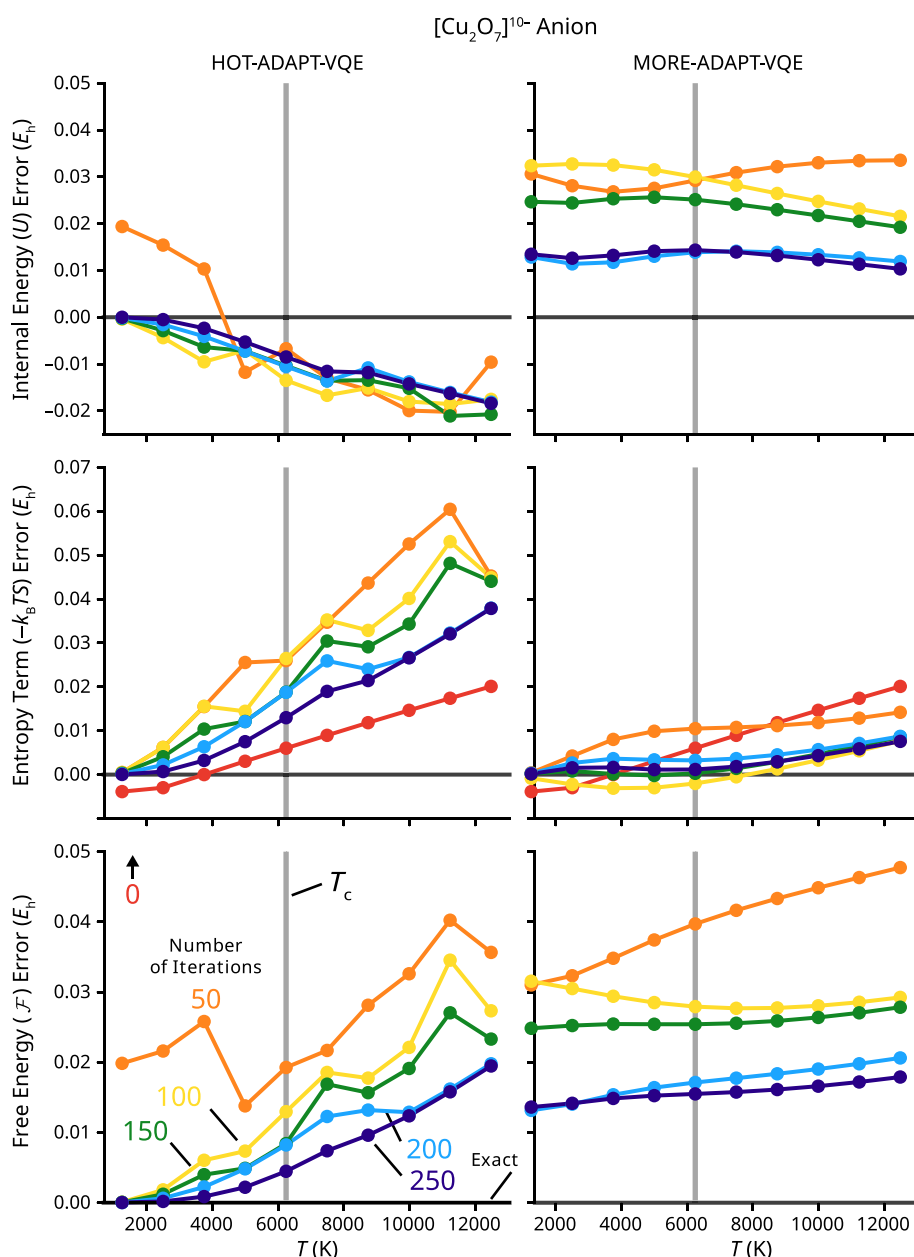


Figure 6. Plots of the error of energetic values from HOT-ADAPT-VQE (left) and MORE-ADAPT-VQE (right) at different iterations as a function of temperature for the $[\text{Cu}_2\text{O}_7]^{10-}$ model. Conventions are the same as in Figure 3 unless otherwise specified. The internal energy and free energy errors lie outside the range of the graph ($\approx 0.2 E_h$) at iteration 0 but do not depend on the method used. HOT-ADAPT-VQE data points for T equal to 0.2 and $0.4 T_c$ were obtained by interpolation from the $0.6 T_c$ solutions.

molecular orbitals was obtained by performing an RHF calculation in the cc-pVDZ basis.^{69,70,85} In designing our active space, we sought to include the $3d_{x^2-y^2}$ orbitals for each copper atom and the $2p_y$ orbital of the central oxygen atom, the minimum necessary active space to partially describe the superexchange mechanism. To this end, we performed state-specific CASSCF computations with a (4e,6o) active space, with starting orbitals selected to include two a_g and four b_{2u} orbitals, to account for orbitals that arise as a linear combination of the aforementioned atomic orbitals. CASSCF was used to target the lowest-energy $^1B_{2u}$ singlet state, with the resulting orbitals used in the quantum algorithm studies. It is important to note that due to the several simplifications made in our model (minimal fragment model, absence of counterions, etc.), this choice of

active orbitals does not lead to the lowest-energy CASSCF solution of the $[\text{Cu}_2\text{O}_7]^{10-}$ fragment.

The resulting CASSCF model space reasonably describes the superexchange mechanism, as our antiferromagnetic J is quite accurate relative to experiments (138 meV from our model vs 120–160 meV for experiment).^{76,86–88} J was estimated according to the singlet–triplet case of the Landé interval rule⁸⁹ as

$$J = \frac{E(T_1) - E(S_0)}{4} \quad (37)$$

where $E(S_0)$ and $E(T_1)$ are the energies of the lowest singlet and triplet states, respectively.

The closed-shell Aufbau determinant $|13b_{2u}^2 14b_{2u}^2\rangle$ and the set of singly excited determinants were used as the reference

manifold, giving a total space of dimension 33 for HOT- and MORE-ADAPT-VQE. The low-temperature thermodynamic character of this system is primarily defined by the singlet–triplet gap between the lowest CASCI states in Figure 5. The associated critical temperature for this gap in our model system is 6243 K (538.0 meV). We report the energetics of $[\text{Cu}_2\text{O}_7]^{10-}$ with different methods for temperatures ranging from 0.2 to $2 T_c$ in Figure 6. Our results include data for up to 250 iterations of each algorithm due to computational time limitations.

In our computations, we observe that for all temperatures, HOT-ADAPT-VQE eventually attempts to add the same operator twice in a row due to a failure to optimize it in the previous macroiteration, despite using a very tight convergence criterion in the BFGS algorithm. If the last operator in the ansatz has a parameter gradient that is too small for BFGS to continue, adding the same operator will just add another parameter with the same gradient to the end of the ansatz, making it pointless to continue the algorithm, as is the case with traditional ADAPT-VQE. We define this situation as premature convergence if the individual states $\{|\Psi_i\rangle\}_{i=1}^k$ do not have energy equal to the k lowest energies obtained with CASCI (only for symmetries accessible through the action of $\hat{U}(\theta)$ on the references $\{|\phi_i\rangle\}_{i=1}^k$). We were generally able, however, to overcome this premature convergence by reinitializing all the ansatz parameters to 0 (as opposed to using parameters from a previous ADAPT-VQE iteration as a guess) and performing a new HOT-VQE subroutine with the same ansatz. Using the resulting reoptimized parameters, HOT-ADAPT-VQE was allowed to proceed. Additional reboots of the algorithm were performed whenever the algorithm attempted to add the same operator twice consecutively. This technique was used to obtain the data reported in Figure 6. To illustrate this rebooting behavior, we include the iteration-by-iteration convergence of $\mathcal{F}$ at the critical temperature ($T = T_c$) in Figure 7. Rebooting was not used for MORE-ADAPT-VQE because it did not suffer from premature convergence to a local minimum in $\mathcal{F}$.

The fact that reinitializing the parameters to 0 can immediately improve the energy is somewhat surprising based on previous reports in the literature. In ref 63, the authors report

HOT-ADAPT-VQE Convergence for $[\text{Cu}_2\text{O}_7]^{10-}$ anion ($T = T_c$)

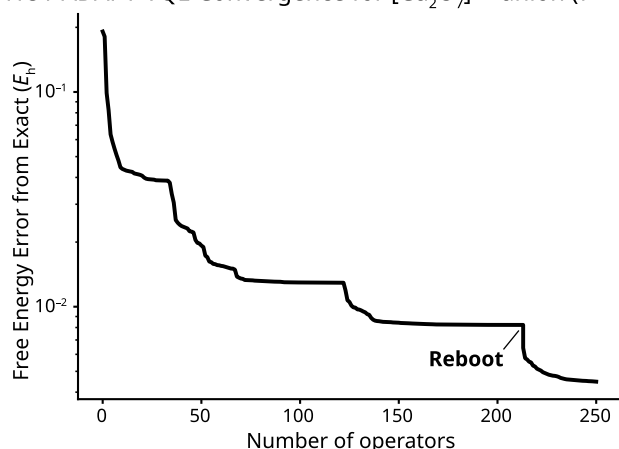


Figure 7. Plot showing the convergence of the free energy of $[\text{Cu}_2\text{O}_7]^{10-}$ using HOT-ADAPT-VQE with $T = T_c$. The large drop at iteration 213 corresponds to the first reboot of the algorithm, where the line connects the prematurely converged energy and the energy after reoptimizing the parameters for the same ansatz.

trying a large number of random initializations to escape a TEPID-ADAPT-VQE solution that failed to include a low-lying excited state for a system, capturing a higher-lying one instead. Using these random initializations did not improve the energy for the same ansatz. We defer a more involved comparison of the energy landscapes from the two methods to future work, but we note that this behavior of HOT-ADAPT-VQE is consistent with that of the traditional ADAPT-VQE ansatz.⁹⁰

Unfortunately, even with our rebooting procedure, HOT-ADAPT-VQE prematurely converged at 0.2 and $0.4 T_c$, essentially collapsing to form the lowest-lying triplet. This is particularly troubling, given that traditional ADAPT-VQE could at least obtain the correct ground state of the system. We tried performing such a traditional ADAPT-VQE calculation until convergence to the ground state singlet ($E_{\text{ADAPT}} - E_{\text{CASCI}} < 10^{-8} E_h$) to obtain an initial ansatz $\hat{U}(\theta)$ which had 111 operators. Using this unitary, a reboot was performed by reinitializing all of the parameters to 0 and switching to the HOT-ADAPT-VQE algorithm, using the same references as before. HOT-ADAPT-VQE was only allowed to perform 20 total reboots due to the long time required to reinitialize the parameters repeatedly. Even with these reboots, at $0.2 T_c$ HOT-ADAPT-VQE still had negligible contributions in the ensemble from states other than the lowest-lying triplet, despite beginning with the correct operator sequence associated with the exact ground state. Similarly, we also find that converging an ADAPT-VQE calculation and then switching to the HOT-ADAPT-VQE algorithm without reboots does not overcome issues of premature convergence.

This failure shows the degree of difficulty in treating a low-temperature system with poor initial guesses of the eigenstates of $\hat{H}$. In our model, the multideterminantal character of the singlet ground state is not captured well by the set of reference states, leading to a reversal of the predicted energy ordering, with the triplet state being incorrectly predicted to lie 0.802 eV lower in energy than the singlet. Consequently, the Boltzmann probabilities are inverted and HOT-ADAPT-VQE favors adding operators that primarily lower the energy of the triplet state. We ultimately used the interpolation strategy from ref 63 to obtain estimates for the Gibbs states at temperatures equal to 0.2 and $0.4 T_c$. For a given HOT-ADAPT-VQE iteration, we report 0.2 and $0.4 T_c$ based on the same $\hat{U}(\theta)$ and $\mathbf{C}$ obtained with HOT-ADAPT-VQE for T equal to $0.6 T_c$. The only difference from the $0.6 T_c$ ensemble is the set of ensemble weights $\{p_i\}$, which are obtained with the Boltzmann distributions for temperatures of 0.2 and $0.4 T_c$. It is important to note that while HOT-ADAPT-VQE has convergence problems, it is typically better than MORE-ADAPT-VQE at obtaining accurate free energies with the same number of operators, accomplishing the original goal of this work. It is notable that for this system, MORE-ADAPT-VQE maintains a significantly lower error in the entropy, despite having a higher error in the free energy. The low error in the entropy, combined with the high error in the internal energy, is due to the nature of MORE-ADAPT-VQE. In attempting to optimize the unweighted average of the energies, MORE-ADAPT-VQE can produce states that are correctly spaced in energy but have larger absolute energy errors.

Figure 6 clearly shows that HOT-ADAPT-VQE outperforms MORE-ADAPT-VQE in predicting free energies at lower temperatures. To quantitatively compare these two methods on equal footing in terms of ansatz length, in Table 2, we show the errors in the free energy $\mathcal{F}$ obtained after 250 iterations of both algorithms. The data in Table 2 show that HOT-ADAPT-

Table 2. Errors (in E_h) Relative to the Exact Free Energies of the $[\text{Cu}_2\text{O}_7]^{10-}$ Fragment at 250 Iterations for HOT- and MORE-ADAPT-VQE

T/T_c	HOT-ADAPT-VQE	MORE-ADAPT-VQE
0.2	2.847×10^{-5}	1.362×10^{-2}
0.4	1.844×10^{-4}	1.416×10^{-2}
0.6	8.724×10^{-4}	1.486×10^{-2}
0.8	2.204×10^{-3}	1.525×10^{-2}
1.0	4.477×10^{-3}	1.550×10^{-2}
1.2	7.399×10^{-3}	1.576×10^{-2}
1.4	9.617×10^{-3}	1.612×10^{-2}
1.6	1.237×10^{-2}	1.660×10^{-2}
1.8	1.579×10^{-2}	1.720×10^{-2}
2.0	1.947×10^{-2}	1.790×10^{-2}

VQE is generally more accurate, with its advantage diminishing as the temperature increases, eventually yielding a slightly less accurate free energy than MORE-ADAPT-VQE at $T = 2 T_c$. This decline in accuracy is unsurprising, as in the limit of $T \rightarrow \infty$, the Boltzmann distribution will be uniform for all eigenstates, removing any freedom in the ensemble weights $\{p_i\}$ and making MORE-ADAPT-VQE and HOT-ADAPT-VQE equivalent. Additionally, in this high-temperature limit, the coupling between references $\{|\phi_i\rangle\}$ through $\bar{H}$ (see eq 13) is also irrelevant since the internal energy simplifies to the unweighted trace of $1/k\bar{H}$, a quantity invariant under rotations among the references. We also remind the reader that the measurement requirements of HOT-ADAPT-VQE will be much greater than those for a MORE-ADAPT-VQE ansatz of the same length, so that if coherence time is not the primary problem for the hardware being used, this improved accuracy may be less meaningful. However, within the chemically relevant and challenging low-temperature regime, HOT-ADAPT-VQE demonstrates substantial promise due to its direct optimization of the Helmholtz free energy.

5. CONCLUSIONS

In this work, we introduced HOT-ADAPT-VQE, a novel hybrid quantum-classical algorithm designed to adaptively construct Gibbs states by explicitly minimizing the Helmholtz free energy through continuous mixing and optimization of an ensemble of reference states. Compared with a simpler strategy based on equally weighting each reference state (MORE-ADAPT-VQE), HOT-ADAPT-VQE can often accurately determine the free energy using significantly shallower operator products. This improved efficiency is achieved by actively adapting the ensemble weights p_i based on current state energies, thereby focusing the ansatz expressiveness on the optimization of those states that dominate the thermal ensemble. Despite these promising features, we have observed some troubling pathologies with variational thermal state algorithms. Specifically, HOT-ADAPT-VQE may fail to converge to the global minimum of $\mathcal{F}$ when important states have very low initial populations and thus they are not improved in the optimization process. This challenge is illustrated in our cuprate model system, where the strongly correlated singlet ground state is poorly represented by the initial set of reference determinants. In many situations, rebooting the algorithm by reinitializing the ansatz parameters to zero can allow the algorithm to continue, though even this approach can fail for sufficiently low temperatures. It is encouraging to observe that interpolation from higher temperatures⁶³ provides a simple and effective

solution to the problem of premature convergence, enhancing the practical utility of HOT-ADAPT-VQE. A more thorough exploration of the numerical parameter optimization details would be beneficial going forward, including the choice of the optimizer itself.

The choice of reference determinants is an open problem, in terms of both the number of references needed and the selection of references for any given value of k . In the case of our cuprate model, CASSCF was used to identify a classically tractable active space; then, a subset of the determinants in that subspace was used. In a real use case where larger active spaces could be simulated on quantum hardware, CASSCF could still be useful for identifying a small set of important determinants to use as references. To obtain the same number of references, the CASSCF active space would not need to grow larger with system size. As mentioned before, situations where thermal effects are localized could have constant requirements for k with the system size, which would make CASSCF reference selection viable. In any case, the results of HOT-ADAPT-VQE are extremely sensitive to the choice of references, making this a key area of future work.

A natural direction for future work is to generalize HOT-ADAPT-VQE to the grand canonical ensemble, relevant for systems exchanging electrons with their environment. In the grand canonical ensemble, the electron number is not fixed but controlled by the chemical potential μ , and the corresponding free energy functional is given by

$$\Omega[\hat{\rho}] = U - TS - \mu N \quad (38)$$

where N is the average number of electrons, expressible as $\text{Tr}(\hat{\rho}\hat{N})$, where $\hat{N}$ is the number operator. Extending HOT-ADAPT-VQE to this ensemble requires generalizing the selection of reference states (which will span different sectors of the Fock space) and considering nonparticle-conserving generalizations of the ADAPT-VQE operator pool. Another desirable extension of HOT-ADAPT-VQE is to model Hamiltonians describing Fermions coupled to phonons. In real finite-temperature chemical and solid-state systems, vibrational modes (phonons) often possess energies comparable to or lower than electronic excitation energies, resulting in significant coupling of these two degrees of freedom that profoundly affects thermal properties. Additionally, the flexibility in choosing the operator pool and reference states within HOT-ADAPT-VQE offers opportunities for future algorithmic improvements. In summary, the good performance demonstrated by HOT-ADAPT-VQE, combined with clear avenues for further generalization and optimization, makes it a promising quantum algorithm for accurate and practical thermal state calculations on quantum computers. Note added in proof: while finalizing this work, we became aware of another quantum algorithm for thermal simulation (see ref 91).

Algorithm 1 The HOT-ADAPT-VQE algorithm.

Require: k references $\{|\phi_i\rangle\}$, generator pool $\{\hat{A}_\mu\}$, inverse temperature β , convergence criteria

- 1: $N \leftarrow 0$
- 2: $\hat{U}^{(0)} \leftarrow \hat{I}$
- 3: $\theta \leftarrow \langle \rangle$
- 4: Solve $\text{HC}^{(0)} = E^{(0)}\text{C}^{(0)}$ $\triangleright \text{H}_{ij} = \langle \phi_i | \hat{H} | \phi_j \rangle$
- 5: $\{p_i\}^{(0)} \leftarrow \left\{ \frac{1}{Z} e^{-\beta(E_i^{(0)} - E_0^{(0)})} \right\}$
- 6: **while** convergence criteria not met **do**
- 7: $N \leftarrow N + 1$
- 8: Determine largest candidate gradient $|g_\mu|$ $\triangleright$ See eq 22
- 9: $\hat{U}^{(N)} \leftarrow e^{\theta_N \hat{A}_\mu} \hat{U}^{(N-1)}$ $\triangleright$ Extend $\hat{U}$ by the new operator
- 10: $\theta^{(N)} \leftarrow (\theta^{(N-1)}, 0)$ $\triangleright$ Extend θ by a new $\theta_N = 0$
- 11: Optimize $\theta^{(N)}$, $\{p_i\}^{(N)}$, and $\text{C}^{(N)}$ with HOT-VQE $\triangleright$ See Figure 1
- 12: **end while**

■ ASSOCIATED CONTENT

Data Availability Statement

Code used in this project is available in a local fork of the QFORTE package at https://github.com/hgrimsml/qforte/tree/new_thermal and in the FORTE package at https://github.com/evangelistalab/forte/tree/qforte_stuff. All data are available upon reasonable request.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.5c01313>.

(TXT)

Output from the FORTE package for Fe^{2+} containing electron integrals, orbital symmetries, etc.; output from the FORTE package for $[\text{Cu}_2\text{O}_7]^{10-}$ containing electron integrals, orbital symmetries, etc.; thermal properties of Fe^{2+} computed using CASCI; and thermal properties of $[\text{Cu}_2\text{O}_7]^{10-}$ computed using CASCI (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Harper R. Grimsley – Department of Chemistry and Cherry Emerson Center for Scientific Computation, Emory University, Atlanta, Georgia 30322, United States; orcid.org/0000-0001-5672-468X; Email: harper.r.grimsley@gmail.com

Francesco A. Evangelista – Department of Chemistry and Cherry Emerson Center for Scientific Computation, Emory University, Atlanta, Georgia 30322, United States; orcid.org/0000-0002-7917-6652; Email: francesco.evangelista@emory.edu

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.5c01313>

Author Contributions

Both authors contributed to project design. Theory was developed by H.R.G. The code was written by H.R.G. Data were collected by H.R.G. Both authors contributed to manuscript preparation.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by the U.S. Department of Energy under Award No. DE-SC0019374. The authors thank Dr. Ilias Magoulas for comments on the manuscript.

■ REFERENCES

- (1) White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **1992**, *69*, 2863–2866.
- (2) Verstraete, F.; Cirac, J. I. Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions. *arXiv* **2004**, arXiv: cond-mat/0407066.
- (3) Stoudenmire, E. M.; White, S. R. Minimally entangled typical thermal state algorithms. *New J. Phys.* **2010**, *12*, 055026.
- (4) Anderson, J. B. A random-walk simulation of the Schrödinger equation: H_3^+ . *J. Chem. Phys.* **1975**, *63*, 1499–1503.
- (5) Liu, Y.; Cho, M.; Rubenstein, B. Ab Initio Finite Temperature Auxiliary Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2018**, *14*, 4722–4732.
- (6) Petras, H. R.; Ramadugu, S. K.; Malone, F. D.; Shepherd, J. J. Using Density Matrix Quantum Monte Carlo for Calculating Exact-on-Average Energies for ab Initio Hamiltonians in a Finite Basis Set. *J. Chem. Theory Comput.* **2020**, *16*, 1029–1038.
- (7) Metzner, W.; Vollhardt, D. Correlated Lattice Fermions in $d = \infty$ Dimensions. *Phys. Rev. Lett.* **1989**, *62*, 324–327.
- (8) Müller-Hartmann, E. Correlated fermions on a lattice in high dimensions. *Z. Phys. B* **1989**, *74*, 507–512.
- (9) Gonis, A.; Däne, M. Extension of the Kohn-Sham formulation of density functional theory to finite temperature. *J. Phys. Chem. Solids* **2018**, *116*, 86–99.
- (10) Yoshikawa, T.; Doi, T.; Nakai, H. Finite-temperature-based time-dependent density-functional theory method for static electron correlation systems. *J. Chem. Phys.* **2020**, *152*, 244111.
- (11) Hirata, S.; He, X. On the Kohn-Luttinger conundrum. *J. Chem. Phys.* **2013**, *138*, 204112.
- (12) Santra, R.; Schirmer, J. Finite-temperature second-order many-body perturbation theory revisited. *Chem. Phys.* **2017**, *482*, 355–361.
- (13) Hirata, S.; Jha, P. K. In *Annual Reports in Computational Chemistry*; Dixon, D. A., Ed.; Elsevier, 2019; Vol. 15, pp 17–37.
- (14) White, A. F.; Chan, G. K.-L. A Time-Dependent Formulation of Coupled-Cluster Theory for Many-Fermion Systems at Finite Temperature. *J. Chem. Theory Comput.* **2018**, *14*, 5690–5700.
- (15) Harsha, G.; Henderson, T. M.; Scuseria, G. E. Thermofield Theory for Finite-Temperature Coupled Cluster. *J. Chem. Theory Comput.* **2019**, *15*, 6127–6136.
- (16) Shushkov, P.; Miller, T. F., III Real-time density-matrix coupled-cluster approach for closed and open systems at finite temperature. *J. Chem. Phys.* **2019**, *151*, 134107.
- (17) Sanyal, G.; Mandal, S. H.; Mukherjee, D. Thermal averaging in quantum many-body systems: a non-perturbative thermal cluster cumulant approach. *Chem. Phys. Lett.* **1992**, *192*, 55–61.
- (18) Sanyal, G.; Mandal, S. H.; Guha, S.; Mukherjee, D. Systematic nonperturbative approach for thermal averages in quantum many-body systems: The thermal-cluster-cumulant method. *Phys. Rev. E* **1993**, *48*, 3373–3389.
- (19) Mandal, S. H.; Ghosh, R.; Mukherjee, D. A non-perturbative cumulant expansion method for the grand partition function of quantum systems. *Chem. Phys. Lett.* **2001**, *335*, 281–288.
- (20) Mandal, S. H.; Ghosh, R.; Sanyal, G.; Mukherjee, D. A non-perturbative path-integral based thermal cluster expansion approach for grand partition function of quantum systems. *Chem. Phys. Lett.* **2002**, *352*, 63–69.
- (21) Kou, Z.; Hirata, S. Finite-temperature full configuration interaction. *Theor. Chem. Acc.* **2014**, *133*, 1487.
- (22) Harsha, G.; Henderson, T. M.; Scuseria, G. E. Thermofield theory for finite-temperature quantum chemistry. *J. Chem. Phys.* **2019**, *150*, 154109.
- (23) Lipkowitz, K. B.; Cundari, T. R.; Boyd, D. B. *Reviews in Computational Chemistry*; John Wiley & Sons: Incorporated: Newark, United States, 2008; Vol. 26.
- (24) Kitaev, A. Y. Quantum measurements and the Abelian Stabilizer Problem. *arXiv* **1995**, arXiv: quant-ph/9511026.
- (25) Aspuru-Guzik, A.; Dutoi, A. D.; Love, P. J.; Head-Gordon, M. Simulated Quantum Computation of Molecular Energies. *Science* **2005**, *309*, 1704–1707.
- (26) 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.
- (27) 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.
- (28) Nakanishi, K. M.; Mitarai, K.; Fujii, K. Subspace-search variational quantum eigensolver for excited states. *Phys. Rev. Res.* **2019**, *1*, 033062.
- (29) Parrish, R. M.; Hohenstein, E. G.; McMahon, P. L.; Martínez, T. J. Quantum Computation of Electronic Transitions Using a Variational Quantum Eigensolver. *Phys. Rev. Lett.* **2019**, *122*, 230401.
- (30) Ollitrault, P. J.; Kandala, A.; Chen, C.-F.; Barkoutsos, P. K.; Mezzacapo, A.; Pistoia, M.; Sheldon, S.; Woerner, S.; Gambetta, J. M.; Tavernelli, I. Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor. *Phys. Rev. Res.* **2020**, *2*, 043140.

- (31) Motta, M.; Sun, C.; Tan, A. T. K.; O'Rourke, M. J.; Ye, E.; Minnich, A. J.; Brandão, F. G. S. L.; Chan, G. K.-L. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nat. Phys.* **2020**, *16*, 205–210.
- (32) Kremenetski, V.; Hogg, T.; Hadfield, S.; Cotton, S. J.; Tubman, N. M. Quantum Alternating Operator Ansatz (QAOA) Phase Diagrams and Applications for Quantum Chemistry. *arXiv* **2024**, arXiv: 2108.13056.
- (33) Yalouz, S.; Senjean, B.; Günther, J.; Buda, F.; O'Brien, T. E.; Visscher, L. A state-averaged orbital-optimized hybrid quantum–classical algorithm for a democratic description of ground and excited states. *Quantum Sci. Technol.* **2021**, *6*, 024004.
- (34) Asthana, A.; Kumar, A.; Abraham, V.; Grimsley, H. R.; Zhang, Y.; Cincio, L.; Tretiak, S.; Dub, P. A.; Economou, S. E.; Barnes, E.; Mayhall, N. J. Quantum self-consistent equation-of-motion method for computing molecular excitation energies, ionization potentials, and electron affinities on a quantum computer. *Chem. Sci.* **2023**, *14*, 2405–2418.
- (35) Kumar, A.; Asthana, A.; Abraham, V.; Crawford, T. D.; Mayhall, N. J.; Zhang, Y.; Cincio, L.; Tretiak, S.; Dub, P. A. Quantum Simulation of Molecular Response Properties in the NISQ Era. *J. Chem. Theory Comput.* **2023**, *19*, 9136–9150.
- (36) Reinholdt, P.; Kjellgren, E. R.; Fuglsbjerg, J. H.; Ziem, K. M.; Coriani, S.; Sauer, S. P. A.; Kongsted, J. Subspace Methods for the Simulation of Molecular Response Properties on a Quantum Computer. *J. Chem. Theory Comput.* **2024**, *20*, 3729–3740.
- (37) Ziem, K. M.; Kjellgren, E. R.; Reinholdt, P.; Jensen, P. W. K.; Sauer, S. P. A.; Kongsted, J.; Coriani, S. Which Options Exist for NISQ-Friendly Linear Response Formulations? *J. Chem. Theory Comput.* **2024**, *20*, 3551–3565.
- (38) Martyn, J.; Swingle, B. Product spectrum ansatz and the simplicity of thermal states. *Phys. Rev. A* **2019**, *100*, 032107.
- (39) Warren, A.; Zhu, L.; Mayhall, N. J.; Barnes, E.; Economou, S. E. Adaptive variational algorithms for quantum Gibbs state preparation. *arXiv* **2022**, arXiv: 2203.12757.
- (40) Zhu, D.; Johri, S.; Linke, N. M.; Landsman, K. A.; Huerta Alderete, C.; Nguyen, N. H.; Matsuura, A. Y.; Hsieh, T. H.; Monroe, C. Generation of thermofield double states and critical ground states with a quantum computer. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 25402–25406.
- (41) Sun, S.-N.; Motta, M.; Tazhigulov, R. N.; Tan, A. T.; Chan, G. K.-L.; Minnich, A. J. Quantum Computation of Finite-Temperature Static and Dynamical Properties of Spin Systems Using Quantum Imaginary Time Evolution. *PRX Quantum* **2021**, *2*, 010317.
- (42) Getelina, J. C.; Gomes, N.; Iadecola, T.; Orth, P. P.; Yao, Y.-X. Adaptive variational quantum minimally entangled typical thermal states for finite temperature simulations. *SciPost Phys.* **2023**, *15*, 102.
- (43) Preskill, J. Quantum Computing in the NISQ era and beyond. *Quantum* **2018**, *2*, 79.
- (44) Kutzelnigg, W. In *Methods of Electronic Structure Theory*, 1st ed.; Schaefer, I. I. H. F., Ed.; *Modern Theoretical Chemistry*; Springer Science & Business Media: New York, 1977; Vol. 3; pp 170–172.
- (45) 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.
- (46) 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.
- (47) 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.
- (48) Greene-Diniz, G.; Muñoz Ramo, D. Generalized unitary coupled cluster excitations for multireference molecular states optimized by the variational quantum eigensolver. *Int. J. Quantum Chem.* **2021**, *121*, No. e26352.
- (49) Otten, M.; Hermes, M. R.; Pandharkar, R.; Alexeev, Y.; Gray, S. K.; Gagliardi, L. Localized Quantum Chemistry on Quantum Computers. *J. Chem. Theory Comput.* **2022**, *18*, 7205–7217.
- (50) Sugisaki, K.; Kato, T.; Minato, Y.; Okuwaki, K.; Mochizuki, Y. Variational quantum eigensolver simulations with the multireference unitary coupled cluster ansatz: a case study of the C_{2v} quasi-reaction pathway of beryllium insertion into a H_2 molecule. *Phys. Chem. Chem. Phys.* **2022**, *24*, 8439–8452.
- (51) Guo, S.; Sun, J.; Qian, H.; Gong, M.; Zhang, Y.; Chen, F.; Ye, Y.; Wu, Y.; Cao, S.; Liu, K.; Zha, C.; Ying, C.; Zhu, Q.; Huang, H.-L.; Zhao, Y.; Li, S.; Wang, S.; Yu, J.; Fan, D.; Wu, D.; Su, H.; Deng, H.; Rong, H.; Li, Y.; Zhang, K.; Chung, T.-H.; Liang, F.; Lin, J.; Xu, Y.; Sun, L.; Guo, C.; Li, N.; Huo, Y.-H.; Peng, C.-Z.; Lu, C.-Y.; Yuan, X.; Zhu, X.; Pan, J.-W. Experimental quantum computational chemistry with optimized unitary coupled cluster ansatz. *Nat. Phys.* **2024**, *20*, 1240–1246.
- (52) Chawla, P.; Shetty, D.; Tsem, P. B.; Sugisaki, K.; Riu, J.; Nogue, J.; Mukherjee, D.; Prasanna, V. S. Trapped ion quantum hardware demonstration of energy calculations using a multireference unitary coupled cluster ansatz: application to the BeH_2 insertion problem. *Eur. Phys. J. Plus* **2025**, *140*, 924.
- (53) Wu, D.; Bai, C. L.; Sagawa, H.; Zhang, H. Q. Multi-Reference UCCSD Variational Quantum Algorithm for Molecular Ground State Energies. *arXiv* **2025**, arXiv: 2408.16523.
- (54) Fitzpatrick, A.; Nykänen, A.; Talarico, N. W.; Lunghi, A.; Maniscalco, S.; García-Pérez, G.; Knecht, S. Self-Consistent Field Approach for the Variational Quantum Eigensolver: Orbital Optimization Goes Adaptive. *J. Chem. Phys. A* **2024**, *128*, 2843–2856.
- (55) Grimsley, H. R.; Evangelista, F. A. Challenging excited states from adaptive quantum eigensolvers: subspace expansions vs. state-averaged strategies. *Quantum Sci. Technol.* **2025**, *10*, 025003.
- (56) Ding, L.; Hong, C.-L.; Schilling, C. Ground and Excited States from Ensemble Variational Principles. *Quantum* **2024**, *8*, 1525.
- (57) Hong, C.-L.; Colmenarez, L.; Ding, L.; Benavides-Riveros, C. L.; Schilling, C. Refining the weighted subspace-search variational quantum eigensolver: compression of ansätze into a single pure state and optimization of weights. *arXiv* **2024**, arXiv: 2306.11844.
- (58) Olsen, S. Canonical-ensemble state-averaged complete active space self-consistent field (SA-CASSCF) strategy for problems with more diabatic than adiabatic states: Charge-bond resonance in monomethine cyanines. *J. Chem. Phys.* **2015**, *142*, 044116.
- (59) Ståhring, J.; Bernhardsson, A.; Lindh, R. Analytical gradients of a state average MCSCF state and a state average diagnostic. *Mol. Phys.* **2001**, *99*, 103–114.
- (60) Deskevich, M. P.; Nesbitt, D. J.; Werner, H.-J. Dynamically weighted multiconfiguration self-consistent field: Multistate calculations for $F+H_2O \rightarrow F+OH$ reaction paths. *J. Chem. Phys.* **2004**, *120*, 7281–7289.
- (61) Werner, H.-J.; Meyer, W. A quadratically convergent MCSCF method for the simultaneous optimization of several states. *J. Chem. Phys.* **1981**, *74*, 5794–5801.
- (62) Economou, S. E.; Warren, A.; Barnes, E. The Role of Initial Entanglement in Adaptive Gibbs State Preparation on Quantum Computers. In *ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp 1–5.
- (63) Sambasivam, B.; Sherbert, K.; Shirali, K.; Mayhall, N. J.; Barnes, E.; Economou, S. E. TEPID-ADAPT: Adaptive variational method for simultaneous preparation of low-temperature Gibbs and low-lying eigenstates. *arXiv* **2025**, arXiv: 2503.14490.
- (64) Nooijen, M. Can the Eigenstates of a Many-Body Hamiltonian Be Represented Exactly Using a General Two-Body Cluster Expansion? *Phys. Rev. Lett.* **2000**, *84*, 2108–2111.
- (65) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; James, A. M.; Lehtola, S.; Misiewicz, J. P.; Scheurer, M.; Shaw, R. A.; Schriber, J. B.; Xie, Y.; Glick, Z. L.; Sirianni, D. A.; O'Brien, J. S.; Waldrop, J. M.; Kumar, A.; Hohenstein, E. G.; Pritchard, B. P.; Brooks, B. R.; Schaefer, H. F., III; Sokolov, A. Y.; Patkowski, K.; DePrince, A. E., III; Bozkaya, U.; King, R. A.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.

- (66) Evangelista, F. A.; Li, C.; Verma, P.; Hannon, K. P.; Schriber, J. B.; Zhang, T.; Cai, C.; Wang, S.; He, N.; Stair, N. H.; Huang, M.; Huang, R.; Misiewicz, J. P.; Li, S.; Marin, K.; Zhao, Z.; Burns, L. A. Forte: A suite of advanced multireference quantum chemistry methods. *J. Chem. Phys.* **2024**, *161*, 062502.
- (67) Stair, N. H.; Evangelista, F. A. QForte: An Efficient State-Vector Emulator and Quantum Algorithms Library for Molecular Electronic Structure. *J. Chem. Theory Comput.* **2022**, *18*, 1555–1568.
- (68) Fletcher, R. *Practical Methods of Optimization*, 2nd ed.; Wiley: Chichester, 2000.
- (69) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.
- (70) Balabanov, N. B.; Peterson, K. A. Basis set limit electronic excitation energies, ionization potentials, and electron affinities for the 3d transition metal atoms: Coupled cluster and multireference methods. *J. Chem. Phys.* **2006**, *125*, 074110.
- (71) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (72) Bednorz, J. G.; Müller, K. A. Possible high T_c superconductivity in the Ba–La–Cu–O system. *Z. Phys. B: Condens. Matter* **1986**, *64*, 189–193.
- (73) Wu, M. K.; Ashburn, J. R.; Torng, C. J.; Hor, P. H.; Meng, R. L.; Gao, L.; Huang, Z. J.; Wang, Y. Q.; Chu, C. W. Superconductivity at 93 K in a new mixed-phase Y–Ba–Cu–O compound system at ambient pressure. *Phys. Rev. Lett.* **1987**, *58*, 908–910.
- (74) Bardeen, J.; Cooper, L. N.; Schrieffer, J. R. Theory of Superconductivity. *Phys. Rev.* **1957**, *108*, 1175–1204.
- (75) Božović, I.; Wu, J.; He, X.; Bollinger, A. T. What is *really* extraordinary in cuprate superconductors? *Phys. C* **2019**, *558*, 30–37.
- (76) Coldea, R.; Hayden, S. M.; Aeppli, G.; Perring, T. G.; Frost, C. D.; Mason, T. E.; Cheong, S.-W.; Fisk, Z. Spin Waves and Electronic Interactions in La_2CuO_4 . *Phys. Rev. Lett.* **2001**, *86*, 5377–5380.
- (77) Malrieu, J. P.; Caballol, R.; Calzado, C. J.; de Graaf, C.; Guichéry, N. Magnetic Interactions in Molecules and Highly Correlated Materials: Physical Content, Analytical Derivation, and Rigorous Extraction of Magnetic Hamiltonians. *Chem. Rev.* **2013**, *114*, 429–492.
- (78) Schurkus, H.; Chen, D.-T.; Cheng, H.-P.; Chan, G.; Stanton, J. Theoretical prediction of magnetic exchange coupling constants from broken-symmetry coupled cluster calculations. *J. Chem. Phys.* **2020**, *152*, 234115.
- (79) Bogdanov, N. A.; Li Manni, G.; Sharma, S.; Gunnarsson, O.; Alavi, A. Enhancement of superexchange due to synergetic breathing and hopping in corner-sharing cuprates. *Nat. Phys.* **2022**, *18*, 190–195.
- (80) Cui, Z.-H.; Zhai, H.; Zhang, X.; Chan, G. K.-L. Systematic electronic structure in the cuprate parent state from quantum many-body simulations. *Science* **2022**, *377*, 1192–1198.
- (81) Cui, Z.-H.; Yang, J.; Tölle, J.; Ye, H.-Z.; Yuan, S.; Zhai, H.; Park, G.; Kim, R.; Zhang, X.; Lin, L.; Berkelbach, T. C.; Chan, G. K.-L. Ab initio quantum many-body description of superconducting trends in the cuprates. *Nat. Commun.* **2025**, *16*, 1845.
- (82) Cheong, S.-W.; Thompson, J. D.; Fisk, Z. Properties of La_2CuO_4 and related compounds. *Phys. C* **1989**, *158*, 109–126.
- (83) Shen, Z. X.; Spicer, W. E.; King, D. M.; Dessau, D. S.; Wells, B. O. Photoemission Studies of High- T_c Superconductors: The Superconducting Gap. *Science* **1995**, *267*, 343–350.
- (84) Longo, J. M.; Raccach, P. M. The structure of La_2CuO_4 and LaSrVO_4 . *J. Solid State Chem.* **1973**, *6*, 526–531.
- (85) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (86) Aeppli, G.; Hayden, S. M.; Mook, H. A.; Fisk, Z.; Cheong, S.-W.; Rytz, D.; Remeika, J. P.; Espinosa, G. P.; Cooper, A. S. Magnetic Dynamics of La_2CuO_4 and $\text{La}_2\text{Ba}_x\text{CuO}_4$. *Phys. Rev. Lett.* **1989**, *63*, 1120.
- (87) Singh, R. R. P.; Fleury, P. A.; Lyons, K. B.; Sulewski, P. E. Quantitative Determination of Quantum Fluctuations in the Spin-Planar Antiferromagnet. *Phys. Rev. Lett.* **1989**, *62*, 2736–2739.
- (88) Braicovich, L.; Ament, L. J. P.; Bisogni, V.; Forte, F.; Aruta, C.; Balestrino, G.; Brookes, N. B.; De Luca, G. M.; Medaglia, P. G.; Granozio, F. M.; Radovic, M.; Salluzzo, M.; Van Den Brink, J.; Ghiringhelli, G. Dispersion of Magnetic Excitations in the Cuprate La_2CuO_4 and CaCuO_2 Compounds Measured Using Resonant X-Ray Scattering. *Phys. Rev. Lett.* **2009**, *102*, 167401.
- (89) Landé, A. Termstruktur und Zeemaneffekt der Multipletts. *Z. Phys.* **1923**, *15*, 189–205.
- (90) Grimsley, H. R.; Barron, G. S.; Barnes, E.; Economou, S. E.; Mayhall, N. J. Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus. *npj Quantum Inf.* **2023**, *9*, 19.



CAS INSIGHTS™

EXPLORE THE INNOVATIONS SHAPING TOMORROW

Discover the latest scientific research and trends with CAS Insights. Subscribe for email updates on new articles, reports, and webinars at the intersection of science and innovation.

[Subscribe today](#)

CAS
A division of the
American Chemical Society