RESEARCH ARTICLE

International Journal of QUANTUM CHEMISTRY WILEY

Exploring the scaling limitations of the variational quantum eigensolver with the bond dissociation of hydride diatomic molecules

Jacob M. Clary[1] | Eric B. Jones[2] | Derek Vigil-Fowler[1] | Christopher Chang[3] | Peter Graf[1]

[1]National Renewable Energy Laboratory, Golden, Colorado, USA

[2]ColdQuanta, Boulder, Colorado, USA

[3]Amazon Web Services, Seattle, Washington, USA

Correspondence
Peter Graf, Computational Science Center, National Renewable Energy Laboratory (NREL), Golden, CO, 80401, USA.
Email: peter.graf@nrel.gov

Abstract

Materials simulations involving strongly correlated electrons pose fundamental challenges to state-of-the-art electronic structure methods but are hypothesized to be the ideal use case for quantum computing algorithms. To date, no quantum computer has simulated a molecule of a size and complexity relevant to real-world applications, despite the fact that the variational quantum eigensolver (VQE) algorithm can predict chemically accurate total energies. Nevertheless, because of the many applications of moderately sized, strongly correlated systems, such as molecular catalysts, the successful use of the VQE stands as an important waypoint in the advancement toward useful chemical modeling on near-term quantum processors. In this paper, we take a significant step in this direction. We lay out the steps, write, and run parallel code for an (emulated) quantum computer to compute the bond dissociation curves of the TiH, LiH, NaH, and KH diatomic hydride molecules using the VQE. TiH was chosen as a relatively simple chemical system that incorporates d orbitals and strong electron correlation. Because current VQE implementations on existing quantum hardware are limited by qubit error rates, the number of qubits available, and the allowable gate depth, recent studies using it have focused on chemical systems involving s and p block elements. Through VQE + UCCSD calculations of TiH, we evaluate the near-term feasibility of modeling a molecule with d-orbitals on real quantum hardware. We demonstrate that the inclusion of d-orbitals and the use of the UCCSD ansatz, which are both necessary to capture the correct TiH physics, dramatically increase the cost of this problem. We estimate the approximate error rates necessary to model TiH on current quantum computing hardware using VQE + UCCSD and show them to likely be prohibitive until significant improvements in hardware and error correction algorithms are available.

KEYWORDS
computational catalysis, quantum computing, TiH, variational quantum eigensolver

1 | INTRODUCTION

To accelerate widespread decarbonization it is generally acknowledged that improved materials and chemicals have a large role to play and that exploration and design via simulation can be extremely valuable. This is the case for batteries, photovoltaics, carbon capture, utilization and storage, catalysis, etc. Because many of the systems in question require high-fidelity electronic structure calculations that can become extremely computationally expensive for classical computers, quantum chemists are increasingly interested in the prospects for quantum computing to accomplish these simulations.

A standard example and unsolved challenge is understanding the naturally occurring nitrogenase enzyme, which allows for nitrogen fixation under ambient conditions. In contrast, the current industrial equivalent to this enzyme is the highly energy intensive Haber-Bosch process, which alone accounts for between 1% and 2% of global carbon emissions and energy usage [1, 2]. Consequently, an improved understanding of the nitrogenase enzyme and other naturally occurring catalysts may allow for significant developments in carbon free, energy efficient industrial processes.

Unfortunately, a detailed understanding of the nitrogenase catalytic site (FeMoco) is complicated by the high degree of electron correlation present in its electronic structure. This characteristic necessitates the use of the most computationally expensive electronic structure methods and greatly limits studies of FeMoco to only the smallest system sizes. Note, too, that FeMoco is just an examplar; it is one of many difficult challenges in computational catalysis specifically and quantum chemistry generally. To achieve decarbonization and other societal goals relying on advanced chemistry we may require a fundamentally more efficient means of computing important materials and chemistry properties.

Quantum computing, in principle, by overcoming the scaling limitations of classical computing, offers just such a game-changing paradigm shift. But while it is now well established that real quantum computing hardware can be used to simulate relatively simple chemical systems, there are still no known use cases where a quantum computer has simulated something that could not be simulated on a classical computer, much less so for a chemistry with practical applications. In this work, we discuss how this divide might be bridged by focusing on quantum computing simulations of a molecule that begins to capture the complexity present in transition metal systems.

1.1 | Classical quantum chemistry

The ability to predict the physical properties of molecular and extended systems with chemical accuracy (resolution of 1 kcal/mol) using quantum chemical techniques has long been a driving goal for computational research. Although many methods have been developed, few entirely ab-initio approaches are able to make consistently chemically accurate predictions for metallic and semiconducting molecules, bulk phases, and surfaces. Coupled cluster accounting for connected single, double, and triple excitations (CCSD(T), where the triple excitations are accounted for perturbatively) and full configuration-interaction (FCI) are among the highest accuracy approaches. However, both of these approaches scale steeply with system size and can only currently be used to model small molecular systems. For a system of n electrons, CCSD(T) scales as n^7 while FCI scales as $n!$, meaning FCI is prohibitive beyond calculations with about 20 electrons in 20 molecular orbitals [3–5].

1.2 | Quantum computing and VQE

Quantum computers can potentially overcome this severe scaling due to their ability to simultaneously represent and manipulate a linear combination of 2^n states on n qubits [6]. More specifically, the variational quantum eigensolver (VQE) framework uses a quantum computer to prepare a parameterized wavefunction and measure fixed-accuracy expectation values of the many-body Hamiltonian while a classical optimizer iteratively updates the wavefunction parameters [6, 7]. The VQE is variational, thus in principle it allows for iterative improvement in the prediction of a chemically accurate energy and other system observables while simultaneously exhibiting only linear scaling with respect to the number of qubits. Despite the significant advantages of this approach, the VQE has currently only been used on relatively small chemical systems to predict properties such as bond lengths or reaction barriers, primarily due to the limitations of current quantum hardware [8]. Among the largest systems currently modeled on a quantum computer without using embedding techniques are H_x chains (where x is the number of H atoms), alkali hydride diatomic molecules, and N_2H_2 [8–11]. Additionally, current hardware limitations and/or the use of embedding techniques frequently necessitate orbital down-selection in which only orbitals near the Fermi level are represented on a quantum computer, with the contribution of the rest determined classically using Hartree-Fock theory. Ideally, quantum computers will become large enough for orbital-down-selection to be rendered unnecessary. Even if orbital down-selection is used, however, the significantly improved scaling behavior exhibited by quantum computers will still allow for larger active spaces to be chosen. In general, the VQE is highly flexible and allows one to account for the specifications of quantum hardware, thus seeing its application in a wide variety of studies [7–15].

1.3 | TiH and paper outline

Previous quantum computing resource assessments suggest that molecules such as FeMoco are well out of reach of current quantum computing hardware [16]. As our goal is not just resource assessment but construction and execution of actual VQE calculations, we selected an intermediate level of chemical difficulty that we suggest a priori is likely out of reach for current quantum computing hardware and algorithms, but is also possibly attainable in the not too distant future. We will probe the scaling of the VQE approach on several hydride diatomic molecule systems, specifically LiH, NaH, KH, and TiH. While these are all of interest in order to understanding the scaling of the VQE, TiH is the main target. TiH was selected because it provides both an approximate model for a bond that may form during a variety of catalytic reactions and renewable energy technologies [17, 18] while also being one of the simplest chemical systems containing d-electrons. The partially filled d-orbitals allow for multiple electron configurations, which is a common feature for systems difficult to model with a classical computer. The presence of multiple configurations allows us to study how this complexity propagates within the VQE as well.

The main contributions of this paper are as follows:

- We develop a software framework to compute electronic structure and bond dissociation curves using VQE for relatively large systems that benefit from, for exmaple, more efficient Hamiltonian measurement using relationships between commuting Pauli strings, optimization of Pauli string term consolidation, and measurement parallelism. We demonstrate this pipeline using emulated error-free quantum devices on a classical supercomputer and compute the properties of several hydride diatomic molecules. These calculations provide practical insights into the difficulty in scaling the VQE algorithm into larger chemical systems.
- We compute the fidelity that would result on a real quantum computer with a range of error rates and confirm that although simple molecules can already be simulated on real quantum hardware, the possible system sizes are still small enough to also be modeled using classical computers. For molecules of more practical interest for catalysis, there is still a wide gap between current hardware and required error rates.
- We discuss several subtleties that are encountered when moving from proof-of-concept studies to real-world applications. For example, we describe in detail how basis set choice can affect not only the ground state energy but the ground state configuration itself, and that these assessments increase in difficulty when the corresponding classical calculations are not feasible.

Most generally, this paper seeks to begin building a bridge from the current state of VQE to realistic applications in catalysis. We show that the above additions to the VQE can together be leveraged to allow quantum chemical calculations for systems on the boundary of accessibility.

2 | RESULTS AND DISCUSSION

The following sections describe the VQE formulation, then what we learn about hydride electronic structure from classical calculations. Next, we discuss efficient parallel Pauli string measurement and optimization, followed by the results of error-free, emulated, quantum computations for the hydride systems. Finally, we estimate the fidelity of these calculations on real quantum hardware in order to predict the likelihood of successfully completing these calculations in the presence of noise. While we do not perform any computations on real quantum computing hardware, the primary focus of this study is instead to develop the code, perform relevant calculations on an error-free quantum computer emulated on a classical supercomputer, and then discuss, given realistic hardware error rates, what *would* have happened had we used real quantum computing hardware. As a result, this work aims to provide a realistic assessment of running these calculations on existing devices while developing a workflow that can be deployed on real quantum computing hardware as algorithmic advances stabilize and hardware error rates decrease.

2.1 | VQE formalism

Within the Born-Oppenheimer approximation, a chemical system is described as electrons interacting in the potential produced by the atomic nuclei at fixed positions. The Hamiltonian, H, of each system can be written in various forms. In this work, we use the so-called second-quantized Hamiltonian form, where systems are described using empty or occupied single-particle spin orbitals and interactions between electrons are represented using creation and annihilation operators. This form of the electronic Hamiltonian is written as Equation (1):

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \tag{1}$$

where

$$h_{pq} = \int dx \phi_p^*(x) \left(-\frac{\nabla^2}{2} - \sum_I \frac{Z_I}{|r - R_I|} \right) \phi_q(x),$$

and

$$h_{pqrs} = \int dx_1 dx_2 \frac{\phi_p^*(x_1)\phi_q^*(x_2)\phi_r(x_2)\phi_s(x_1)}{|r_1 - r_2|}.$$

Because the Coulomb interaction between electrons is a two-body interaction, the terms of this Hamiltonian contain up to two creation and two annihilation operators. The integral for h_{pq} describes the kinetic energy terms of electrons and their Coulomb interaction with the nuclei while the integral for h_{pqrs} describes the electron–electron Coulomb repulsion. Next, the second quantized Hamiltonian with operators acting on indistinguishable Fermions is mapped to operators acting on distinguishable qubits. The result of this mapping is a linear combination of products of single-qubit Pauli operators, with each product called a Pauli string. Although various encoding schemes exist, we used the parity encoding scheme because it mapped the Hamiltonians of the hydrides using the fewest number of Pauli strings as compared to the Jordan-Wigner and Bravyi-Kitaev schemes [19]. To calculate expectation values of the mapped Hamiltonians, we used the chemically inspired unitary coupled cluster ansatz which is truncated to either single excitations (UCCS) or singles and doubles excitations (UCCSD), to represent our trial wavefunction. Note that UCCS is a cheaper but fundamentally less accurate method than UCCSD, as the doubles amplitudes capture the electron correlation and the singles amplitudes mostly account for relaxation effects. The ground state can be calculated from the Hartree-Fock reference state, $|\Psi_{HF}\rangle$, using excitation operators. The UCCSD ansatz can thus be written as:

$$|\Psi_{UCCSD}(\Theta)\rangle = \exp^{T(\Theta) - T^\dagger(\Theta)} |\Psi_{HF}\rangle \tag{2}$$

where $T(\Theta) = T_1(\Theta_1) + T_2(\Theta_2)$ is the cluster operator, which is expanded using the connected operators $T_1(\Theta_1)$ and $T_2(\Theta_2)$ in order to introduce singles and doubles excitations into the wavefunction, respectively, and Θ is a vector of parameters needed to specify the single- or two-qubit unitary gates in the quantum circuit [7].

Finally, the ground state of each Hamiltonian is found by determining the set of parameters that minimize energy expectation value, as described by the variational principle

$$E_0 \leq \frac{\langle \Psi(\Theta)|H|\Psi(\Theta)\rangle}{\langle \Psi(\Theta)|\Psi(\Theta)\rangle}, \tag{3}$$

where E_0 is the true ground state energy of H. The state preparation and measurement of the quantum circuit is performed on the quantum computer. The transformation of the quantum circuit statevector into an expectation value and the subsequent parameter optimization are both performed on a classical computer.

2.2 | Hydride electronic structure

LiH, NaH, and KH are all diatomic molecules composed of an alkali earth metal atom and a hydrogen atom. These molecules were selected because of their relatively few valence electrons and lack of d-orbitals participating in their bonding. All three molecules adopt similar orbital occupations with the highest energy s orbital on Li, Na, or K bonding with the H 1s orbital to form a fully filled valence shell. As these diatomic molecules dissociate, the two asymptotic limits of the diatomic molecules (bound and dissociated) will mix. Despite this, the VQE is still able to predict chemically accurate total energies for these molecules [8, 10].

In contrast to the bound alkali hydride diatomic molecules, bonding in TiH and other transition metal hydride diatomic molecules can arise from multiple occupations that have been studied both experimentally and computationally [17, 20, 21]. The two primary occupations to consider are the $3d^34s^1$ and $3d^24s^2$ occupations [17]. In the first occupation, a 4s–1s bond is formed between Ti and H, leading to the $^4\Phi(S=4)$ state with an occupation of $...6\sigma^27\sigma^13\pi^11\delta^1$, where the 6σ orbital is the Ti-H bonding orbital, the 3π and 1δ orbitals are the Ti 3d-like orbitals, and the 7σ orbital is a mixture of the 4s, 4p, and 3d orbitals. In the second occupation, hybridization of the 4s–4p orbitals and 4s–3d orbitals can occur, with one hybrid orbital bonding with H. The $^2\Delta(S=2)$ state has an occupation of $...6\sigma^27\sigma^21\delta^1$, where the 6σ orbital is the Ti—H bonding orbital and the 7σ orbitals is the nonbonding 4s–3d hybrid orbital. Because Ti has few d-electrons, the Ti 4s and 3d orbitals are spatially similar, and Ti—H bonding via the 4s–3d hybrid orbital competes with Ti—H bonding via the 4s–4p hybrid orbital such that the higher energy $^2\Delta(S=2)$ state lies
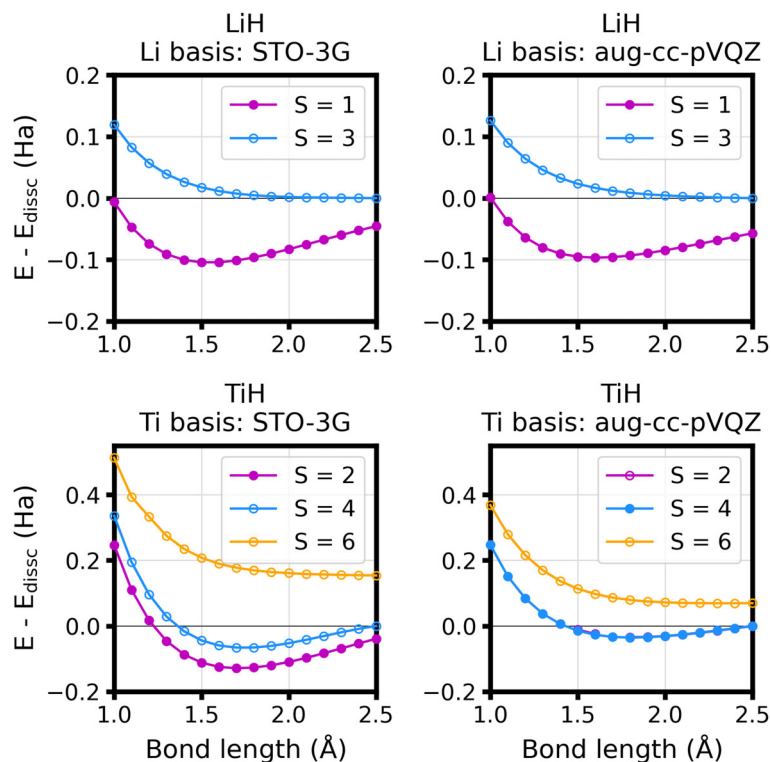
FIGURE 1 Predicted CCSD bond dissociation curves of the LiH and TiH diatomic molecules initialized to different spin multiplicities using different orbital basis sets on a classical computer. The predicted TiH ground state configuration changes depending on the orbital basis set chosen. The ground state configuration is denoted by filled markers while higher energy configurations are denoted by empty markers.

only 0.011 Ha above the $^4\Phi(S=4)$ state according to FCI calculations [17]. We note that the same asymptotic mixing that occurs during alkali hydride diatomic molecule dissociation is also present during TiH dissociation.

The choice of basis set used to model TiH can change the predicted ground state configuration. Classical CCSD quantum calculations were used to calculate the total energies of LiH and TiH initialized to different spin multiplicities (see Figure 1). These calculations show that using the STO-3G basis set on the Ti atom results in the low-spin configuration being the ground state configuration by 0.063 Ha while using the aug-cc-pVQZ basis set on the Ti atom predicts the correct opposite configuration ordering, with the high spin configuration being 0.001 Ha more stable. This issue is of course not present in LiH because the only other possible spin multiplicity does not have an energy minimum over the same range of bond lengths. As one would expect, these results demonstrate that the choice of orbital basis set can be an important consideration for systems with multiple possible spin configurations and may qualitatively affect the predicted system properties. Unfortunately, the capability of current quantum computers greatly limits the choice of orbital basis set because this choice can drastically impact the computational cost of the computations, as discussed further in Section 2.3. As a result, while basis set choice remains a key cost parameter in VQE calculations, care must be taken that a specific choice does not alter the fundamental properties of the system of interest. In the short term, results from existing classical calculations, including HF, post-HF, and DFT methods can provide guidance on the impact of basis set choice for classically studied chemical systems. In the long term, VQE calculations without corresponding classical results should be converged with respect to basis set choice, as is currently commonly done for classical calculations.

In addition to the correct energy ordering of different molecular spin configurations, basis set choice must also allow for an accurate description of a chemical system's electronic structure, with the frontier orbitals being of particular importance. For LiH, classical CCSD calculations at the experimental Li-H bond length using the STO-3G basis set predict frontier orbitals with similar degrees of Li s, Li p, and H s character as calculations using the aug-cc-pVQZ basis set. However, the STO-3G basis set still predicts a LiH HOMO–LUMO gap, E_g, of 9.90 eV, while the aug-cc-pVQZ basis set predicts a gap of only 7.99 eV (Figure S1). To further test the E_g and frontier orbital character dependence for these molecules on basis set choice, we repeated the above calculations using a larger variety of basis set choices (STO-3G, 3-21G, 6-31G, cc-pVDZ, aug-cc-pVDZ, cc-pVQZ, aug-cc-pVQZ). These calculations show that the 3-21G basis set generally allows for E_g and frontier orbital character predictions that approximate the results calculated using the aug-cc-pVQZ basis set for LiH, NaH, and KH (Figures S2 and S3). For $^4\Phi(S=4)$ TiH, E_g calculated using the 3-21G basis set was much closer to the aug-cc-pVQZ E_g than the STO-3G E_g was (Figure S2). However, Figure 2 shows that the STO-

TiH CCSD PDOS calculated with the STO-3G basis set on Ti

TiH CCSD PDOS calculated with the 3-21G basis set on Ti

TiH CCSD PDOS calculated with the 6-31G basis set on Ti

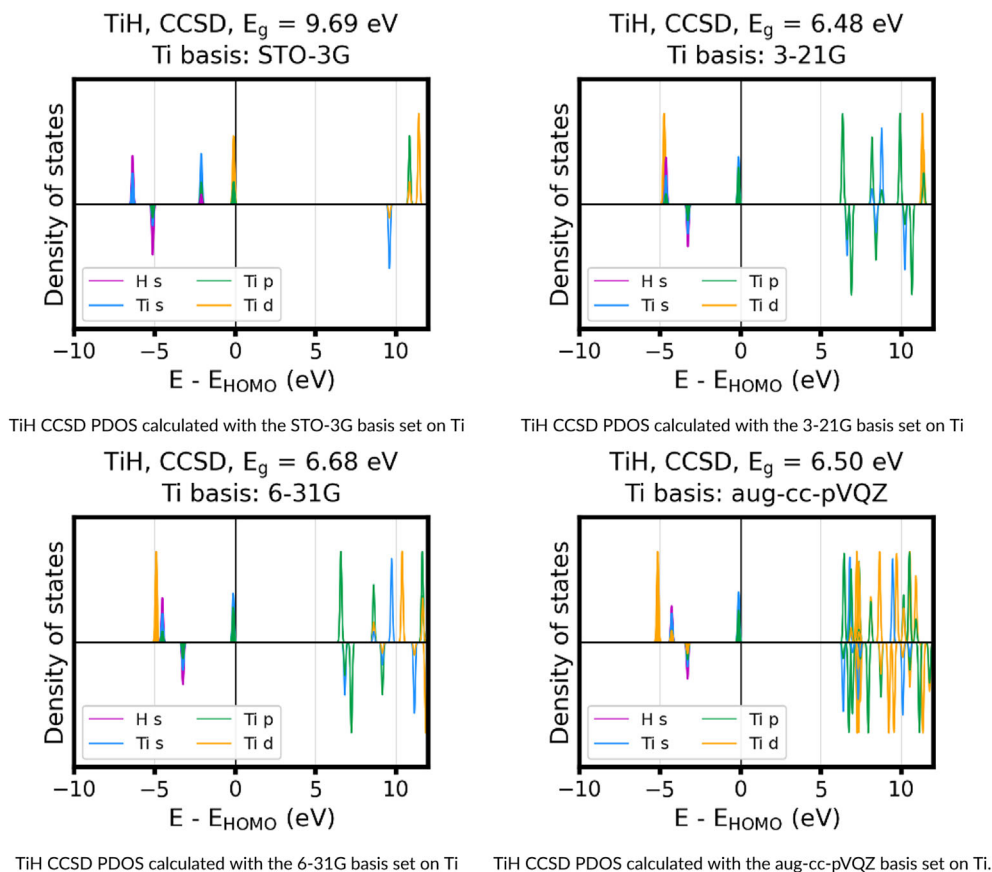TiH CCSD PDOS calculated with the aug-cc-pVQZ basis set on Ti.

FIGURE 2 Selected PDOS for the studied hydride diatomic molecules using CCSD on a classical computer. The STO-3G basis set was always used for H. The bonding/orbital hybridization mentioned previously is visible in the valence region of the PDOSs.

3G, 3-21G, and 6-31G basis sets all predicted that the minority spin LUMO orbitals were higher in energy than any of the correlation consistent basis sets, resulting in changes to the predicted LUMO characters.

As is well-known in classical calculations, the cost of total energy minimization using the VQE scales with the number of orbitals available for occupation and thus electron excitation. As a result, significant computational resource savings may be achieved by freezing core orbitals that negligibly contribute to the bonding and/or removing high energy/non-bonding virtual orbitals from the active space. Indeed, the frozen core approach is an approximation already commonly applied with great success in both molecular and periodic DFT calculations [22–24]. Classical CCSD calculations for LiH, NaH, and KH show that the highest energy occupied molecular orbital lies approximately 20 eV higher in energy than the next highest energy occupied orbital, justifying the selection of this orbital as the only valence orbital. $^4\Phi(S{=}4)$ TiH has five valence spin orbitals located within 10 eV of the Fermi level, with relative energies that all depend on basis set choice. These five occupied spin orbitals can be treated as unfrozen valence orbitals because the next highest energy orbital is approximately 40 eV lower in energy.

To summarize, (1) just as in classical quantum chemistry, basis set choice is critical, and validation by convergence with respect to basis set is desirable, and (2) the TiH system exhibits exactly the type of nuance that necessitates this type of validation.

2.3 | Efficient and parallel measurement of Hamiltonian Pauli strings

The number of Pauli strings in the second-quantized chemical Hamiltonian on n qubits grows as n^4 with the number of spin orbitals in the calculation (Figure 3) [25]. However, separate explicit measurements of the compiled quantum circuit in each set of measurement bases in the Hamiltonian is not necessary. For example, it is well known that the expectation value of the H_2 Hamiltonian within the Bravyi-Kitaev mapping can be measured much more efficiently than one would naively expect from the five term Hamiltonian [8]. This is because Pauli strings that share an eigenbasis can be measured simultaneously, and Pauli strings share an eigenbasis if and only if they commute [26]. As a result, the Hamiltonian can be partitioned into groups of qubit-wise commuting Pauli strings, each of which can be measured simultaneously on a quantum computer and later used to reconstruct the expectation value of the original Hamiltonian on a classical computer. Without accounting for any additional symmetries in the Hamiltonian, the general problem of determining the smallest set of unique measurements that can be used to determine the energy
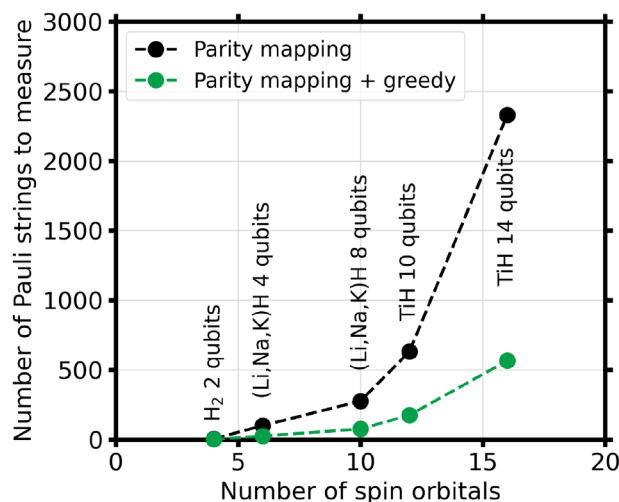
FIGURE 3   Scaling of number of Pauli strings in Hamiltonian for different molecules using the STO-3G basis set. The dashed line is included as a guide to the eye.

of any term in the Hamiltonian is equivalent to the clique cover and set cover combinatorics problems, both of which are known to be NP-hard [26–28]. Briefly, the set cover problem can be stated as asking for the smallest number, N_p, of given subsets whose union equals a universal set, {U}. Thus, the exact solution for the most efficient set of measurements to make for a system can quickly become intractable and, ironically, could itself likely benefit from quantum optimization heuristics. One approach to approximate a solution to the set cover problem involves the use of an iterative greedy algorithm. This algorithm iteratively adds the measurement basis (a Pauli string defining measurement bases for all qubits) to the final set of measurement bases {B}, that commutes with the greatest remaining number of Pauli strings in {U} (the Hamiltonian), that is, that covers the greatest number of Pauli strings that have yet to be covered. The commuting family of Pauli strings is then removed from {U} prior to the next iteration. We found that this algorithm reduces the scaling prefactor for this problem by approximately a factor of five for the systems studied here and thus significantly decreases the number of measurements required for a given Hamiltonian (Figure 3). Nevertheless, as the sophistication of the chosen basis set increases, the number of Pauli strings needed to describe the Hamiltonian and the number of possible orbitals that can be included in the active space grows. For example, TiH modeled using 14 qubits and only the 6-31G basis requires about 2.5× the original number of Pauli strings as the STO-3G basis set.

Although efficient measurement of Hamiltonian Pauli strings can decrease the number of measurements required for large Hamiltonians, a full optimization of the TiH wavefunction using this approach can still involve optimization of dozens to hundreds of parameters, Θ, primarily depending on the number of orbitals included in the calculation. However, further savings can be had because calculating the expectation value of the quantum circuit in each set of measurement bases can be performed independently. As a result, parallelization over P processors of the energy evaluation measurements of bases in {B} allows VQE minimization to become accessible for larger chemical systems. We emphasize that this approach is primarily relevant for the study of problems too large to reliably model on existing quantum hardware at existing error rates and thus necessitates the use of emulation on classical computers in a noiseless simulation environment. Secondarily, this approach is useful for the study of either general or system-specific VQE optimizations prior to calculations on real quantum devices.

In this scheme, the set of final measurement bases is distributed across all processors such that each processor collects statistics for $\lceil N_p/P \rceil$ measurement bases on a local quantum circuit (see Section 4). These statistics are then aggregated across all processors such that the expectation value of any Pauli string in the original Hamiltonian can be reconstructed from these statistics. On dual-socket nodes with two 3.0 GHz, 18-core Intel Xeon Gold 6154 Skylake processors, we find that the time required per energy evaluation can be decreased by at least an order of magnitude and can allow the optimization problem to become computationally feasible (Figure 4). These results demonstrate that qubit parallelization can help measurement throughput as is currently widely used in classical computers. Importantly however, the efficient selection of {B} and measurement parallelization only provide polynomial speed increases. The exponential scaling of the problem still dominates, indicating that Pauli string measurement basis parallelization and optimization must be married with high-fidelity quantum hardware in order to furnish a powerful computational paradigm for quantum chemistry generally.

2.4 | Choice of classical optimizer

As mentioned above, the size of the Hamiltonian, and thus the parameter space to optimize, grows exponentially with system size. As a result, the difficulty of the classical optimization increases significantly for larger chemical systems, both because of the higher dimensional parameter space and the
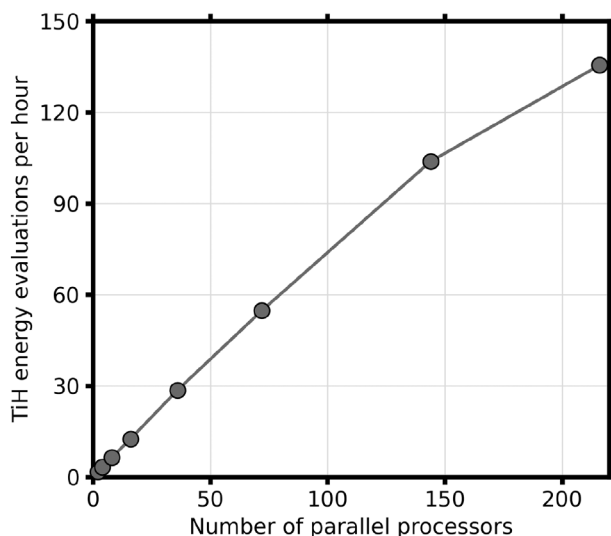
FIGURE 4 TiH energy evaluations per hour as a function of the number of processors the basis set measurements are parallelized over. Approximately linear scaling is observed until the number of processors exceeds the number of Pauli strings in the Hamiltonian.
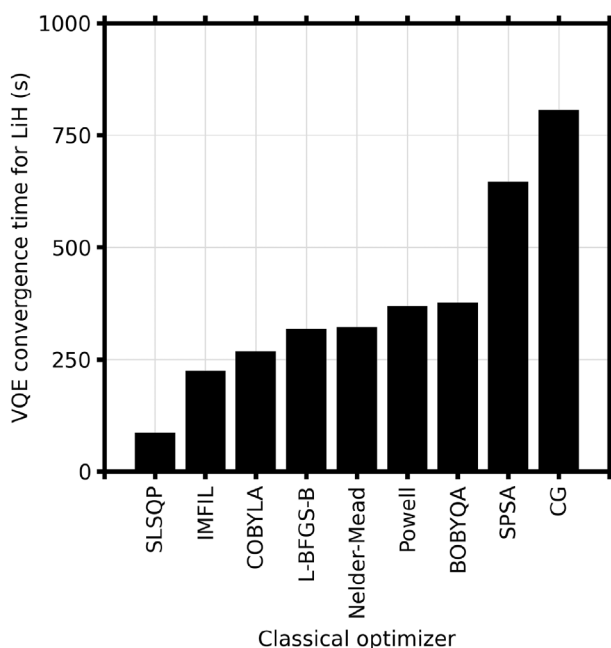

FIGURE 5 Comparison of optimizer performance to converge the total energy of the LiH diatomic molecule with a 6–31 g basis set to within 10 meV of the exact energy.

difficulty in quantifying the effects of noise. Gradient descent algorithms, in particular, are highly affected by noisy measurements of each Pauli string because noise can both increase the difficulty in converging to the global minimum rather than any local minima and cause estimates of the gradient to vary wildly, potentially making gradient descent impossible at all. Although noise will be a significant hurdle in modeling systems such as TiH on near term quantum devices, here we use the noiseless statevector simulator to decouple the effects of noise from the difficulty of the optimization.

We study the effect of optimizer choice on the VQE time to convergence for a LiH diatomic molecule using the 6-31G basis set (Figure 5). The ideal optimizer will converge to chemical accuracy in the shortest time. We find that the sequential least squares programming (SLSQP) optimizer consistently produces the most accurate energy minimization in the fastest time [29]. However, we note that this optimizer is a local search algorithm that is not guaranteed to find the global minimum. This drawback is particularly important for systems such as TiH because the large parameter space requires an optimizer that can perform a broad-breadth search of the potential energy surface while still being able to descend into potentially narrow energy wells once they are found. Unfortunately, the classical optimization required for VQE is global optimization of
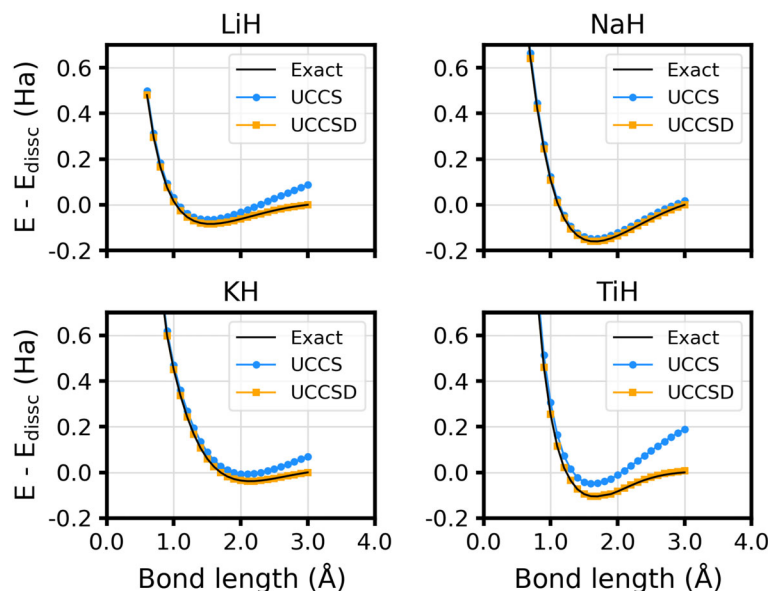
FIGURE 6 Bond dissociation curves of LiH, NaH, KH, and TiH using the STO-3G basis set using the VQE + UCCS ansatz, the VQE + UCCSD ansatz, and exact Hamiltonian diagonalization.

large, nonconvex, expensive, noisy functions and is one of the hardest types of optimization problems there is, meaning that such an ideal optimizer does not exist.

2.5 | Alkali hydride diatomic molecule bond dissociation

We find that the VQE can predict chemically accurate bond dissociation curves for LiH, NaH, and KH using the STO-3G basis set (Figure 6) as compared to the exact bond dissociation curve (direct diagonalization of the Hamiltonian). Because the valence electronic structures of LiH, NaH, and KH are all extremely similar, the scaling of the VQE cost for these systems is similar. Furthermore, despite the much more severe scaling with basis set choice, the STO-3G basis set coupled with the additional orbital freezing and reduction discussed in Section 2.2 also allows the TiH system to become accessible to quantum computer emulation schemes on reasonable timescales. We again emphasize that while important to report here, the physical relevance of this TiH calculation is primarily qualitative due to the inherent limitations in accuracy of the basis set choice, as described in Section 2.2. Nevertheless, the performance of the VQE + UCCS(D) algorithms can still be compared to the exact dissociation curve because the Hamiltonian is limited by the same restrictions in all cases. We find that the UCCSD ansatz accurately reproduces the exact bond dissociation curve for all of the diatomic molecules. In contrast, the UCCS ansatz begins to significantly deviate from both the UCCSD and exact energy curves at large bond distances, resulting in the energy of the dissociated limit being much higher in energy. We note that the bond dissociation curve for TiH is very sensitive to the choice of orbitals included as the active space. A set that insufficiently describes the bonding may exhibit an artificial kink at the Coulson-Fischer point [30, 31] (approximately 2.4 Å) where the bond dissociation curves for different spin configurations cross.

2.6 | Fidelity

Despite the usefulness of modeling chemical systems with the VQE using a noiseless emulator, ideally the continued improvement of real quantum device capabilities will eventually allow for the modeling of larger systems such as TiH. Although this work demonstrates that diatomic molecule calculations using the STO-3G basis set are currently accessible, TiH calculations using larger active spaces or more sophisticated basis sets can become dramatically more expensive.

In order to further understand the viability of modeling the above chemical systems on a real quantum device, it is useful to estimate the state preparation and measurement (SPAM), single qubit, and two-qubit error rates required to obtain various levels of calculation fidelity. Calculation fidelity, F, in the digital error model can be employed as a useful proxy to estimate the hypothetical VQE total energy calculation fidelity [32]. To this end, F was calculated using Equation (4):
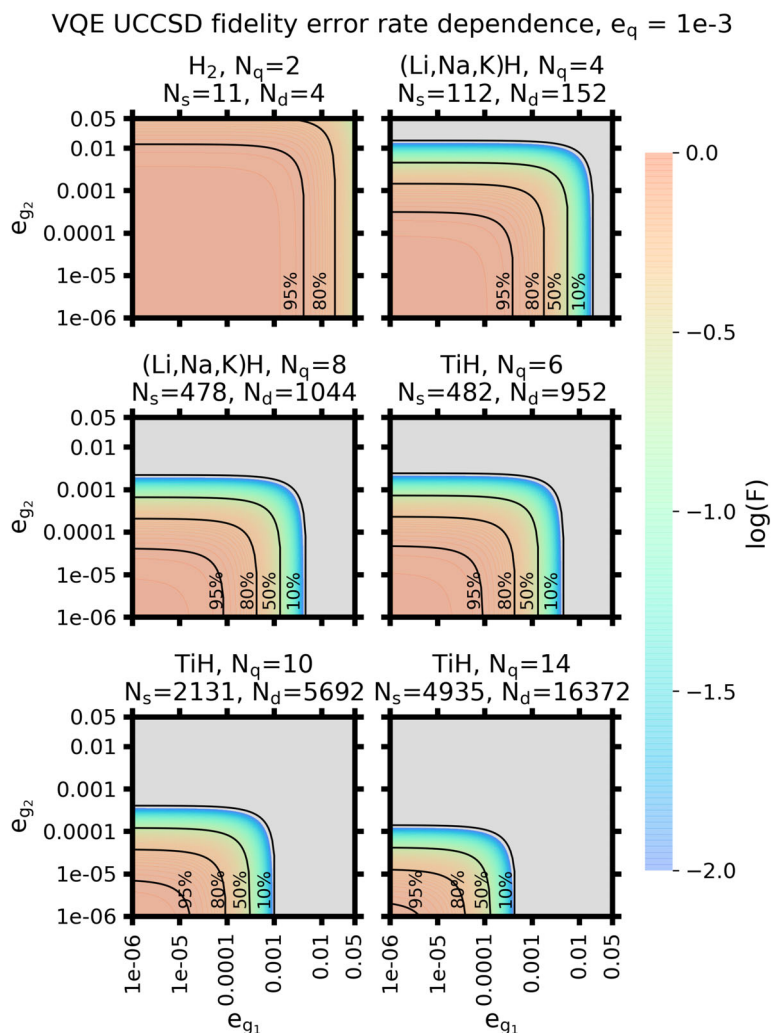
FIGURE 7 Fidelity estimates for the UCCSD ansatz for the hydride diatomic molecules studied in this work with different numbers of qubits, single qubit gates, and two-qubit gates in the quantum circuit. A fidelity of 1.0 corresponds to no fidelity loss due to the considered errors.

$$F = \left(1 - e_{g_1}\right)^{G_1} \left(1 - e_{g_2}\right)^{G_2} \left(1 - e_q\right)^{Q}, \tag{4}$$

where e_{g_1} is the single qubit gate error rate, e_{g_2} is the two-qubit gate error rate, e_q is the SPAM error rate, G_1 is the number of single qubit gates, G_2 is the number of two-qubit gates, and Q is the number of qubits in the circuit. The error rate for each single and two-qubit gate was assumed to be constant. Equation (4) is plotted in Figure 7 for the circuits of the different hydride diatomic molecule systems discussed above using the STO-3G basis set and the UCCSD ansatz. The same results for the UCCS ansatz are shown in Figure S4. Existing quantum computing devices have SPAM, single, and two-qubit error rates of approximately 1e−2, 1e−3, and 1e−2, respectively, [33–35]. As a result, Figure 7 (bottom) shows that this error model predicts that the larger UCCSD scale calculations are not yet reliably feasible on existing hardware. We note that the fidelity of an H_2 molecule is predicted to be approximately 0.95 at current error rates (Figure 7 top left). Both observations are consistent with the existing literature on chemical properties predicted with the VQE on real quantum computers [8–11]. They support the idea that our estimated gate counts and consequent fidelities provide reasonably accurate lower fidelity bounds for the more computationally demanding systems given the lack of hardware-specific circuit optimization. Without further advances in error correction schemes and circuit optimization for specific hardware, robust TiH models will require error rates approximately two orders of magnitude lower than exhibited by any existing quantum computer. In contrast, the UCCS ansatz might provide a much faster route toward experimental validation of the above results. In the near future, all of the studied hydride diatomic molecule calculations can likely be carried out on real quantum computers with reasonable fidelities with only an approximately factor of five improvement in existing error rates, however, they will still likely require the use of a qualitatively inaccurate wavefunction ansatz.

3 | CONCLUSIONS

In this work, we have examined the computational complexity of VQE against a series of hydride diatomic molecules from LiH to the d-orbital-containing TiH. We show that the complexity of the simulation drastically increases upon incorporation of d orbitals due to more Pauli strings in the Hamiltonian, a larger number of 1- and 2-electron excitation operators, and the presence of multiple orbital occupation configurations. Furthermore, we show that although a simpler basis set facilitates VQE estimation of the TiH bond dissociation curve on current (classically emulated) quantum devices, it comes at the cost of incorrect prediction of even the ground state occupation configuration and relies on the use of error free emulation. This tradeoff will likely remain relevant within the NISQ computing era, particularly for more complicated chemical systems involving multiple transition metal atoms.

Although the inclusion of d orbitals in the TiH VQE calculations clearly increases the model complexity, the exact impact of transition metals in a chemical system may not be as clear for more complicated chemical systems. These types of systems may lack data from prior high-fidelity classical calculations that inform future quantum computing calculations, and may instead rely on iterative VQE testing. This process, if needed, will be greatly facilitated by further developments in the VQE algorithm itself.

For example, combining the UCCSD formalism with an adaptive ansatz such as ADAPT-VQE [36] has the potential to reduce the complexity of the quantum circuit (number of optimization parameters, gate depth, etc.) significantly. Additionally, the use of so-called transcorrelated Hamiltonians has been shown to achieve quantitatively accurate ground and excited state energies using minimal basis sets [37]. Optimizers that are resilient to noise will also be particularly helpful for large system modeling on real hardware [38]. Quantum embedding theory [39] shows great potential to extend the reach of quantum computing to larger systems, particularly those with a small set of atoms that require high-fidelity methods surrounded by a less computationally complex environment. Finally, improvements in qubit error rates, error mitigation [40, 41], and postselection [42] could further improve accuracy for a given circuit depth.

It is currently unclear whether the combined improvements offered by these developments will allow for the full modeling of molecules like nitrogenase on a quantum computer. We nevertheless hope that this detailed study will provide a benchmark that can be revisited following further algorithmic and hardware developments on the way to modeling chemical systems that are both practically important and currently out of reach for classical computing.

4 | METHODS

All classical CCSD electronic structure calculations were performed using the Gaussian 16 software package [43]. The LiH, NaH, and KH diatomic molecules were modeled with either a spin multiplicity of 1 or 3 while TiH was modeled with spin multiplicities of 2, 4, and 6 (Section 2.2) [17]. The screened bond lengths were chosen to cover the bonding energy well present within the dissociation curve. The basis set screening was carried out by varying the basis set on Ti while using the STO-3G basis set on H to mimic the likely progression of future calculations on a quantum computer. The number of primitive gaussians used for each molecule calculation is shown in Figure S5. Altering the H basis set did not significantly change the qualitative energy ordering or electronic structure trends.

Quantum computer emulation was performed using IBM's Qiskit API and simulator [44]. Hamiltonian preparation and diagonalization and preparation of the wavefunction ansatz for each system modeled in this work were performed with the Qiskit code package. The gate counts used in Section 2.5 were obtained by summing the circuit occurrences of single-qubit rotation U1, U2, and U3 gates and 2-qubit CX gates. These gate counts could be further optimized for specific quantum hardware and thus are upper bound estimates for fidelity expectations. Parallelization of energy evaluations over multiple processes was performed by using the mpi4py and Qiskit code packages. Qiskit was used to construct the appropriate shared quantum circuit for each system while an mpi4py wrapper distributed the circuit and one or more needed measurement bases equally to the different available processors. Each processor then determines the bit string counts that result from each circuit measurement while mpi4py aggregates all resulting data in order to calculate the final energy evaluation for a given set of parameters. This code used in this work is available on request and will be released open-source.

AUTHOR CONTRIBUTIONS

Jacob Clary: Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft. **Eric Jones:** Conceptualization; funding acquisition; investigation; methodology; supervision; validation. **Derek Vigil-Fowler:** Conceptualization; methodology; supervision; writing – review and editing. **Christopher Chang:** Conceptualization; methodology; supervision. **Peter Graf:** Conceptualization; funding acquisition; methodology; project administration; supervision; writing – review and editing.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Peter Graf https://orcid.org/0000-0003-3150-3404

REFERENCES

[1] S. Ghavam, M. Vahdati, I. Wilson, P. Styring, *Front. Energy Res.* **2021**, *9*, 34.

[2] M. Capdevila-Cortada, *Nat. Catal.* **2019**, *2*(12), 1055.

[3] S. Lehtola, N. M. Tubman, K. B. Whaley, M. Head-Gordon, *J. Chem. Phys.* **2017**, *147*(15), 154105.

[4] G. H. Booth, A. Alavi, *J. Chem. Phys.* **2010**, *132*(17), 174104.

[5] K. D. Vogiatzis, D. Ma, J. Olsen, L. Gagliardi, W. A. De Jong, *J. Chem. Phys.* **2017**, *147*(18), 184111.

[6] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, X. Yuan, *Rev. Mod. Phys.* **2020**, *92*(1), 15003.

[7] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, A. Aspuru-Guzik, *Chem. Rev.* **2019**, *119*(19), 10856.

[8] P. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, et al., *Phys. Rev. X* **2016**, *6*(3), 031007.

[9] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, J. M. Gambetta, *Nature* **2017**, *549*(7671), 242.

[10] A. J. McCaskey, Z. P. Parks, J. Jakowski, S. V. Moore, T. D. Morris, T. S. Humble, et al., *NPJ Quantum Inf.* **2019**, *5*(1), 1.

[11] G. A. Quantum, Collaborators, F. Arute, K. Arya, R. Babbush, D. Bacon, et al., *Science* **2020**, *369*(6507), 1084.

[12] I. G. Ryabinkin, S. N. Genin, A. F. Izmaylov, *J. Chem. Theory Comput.* **2018**, *15*(1), 249.

[13] R. M. Parrish, E. G. Hohenstein, P. L. McMahon, T. J. Martínez, *Phys. Rev. Lett.* **2019**, *122*(23), 230401.

[14] K. M. Nakanishi, K. Mitarai, K. Fujii, *Phys. Rev. Res.* **2019**, *1*(3), 33062.

[15] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, P. Zoller, *Nature* **2019**, *569*(7756), 355.

[16] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, M. Troyer, *Proc. Natl. Acad. Sci.* **2017**, *114*(29), 7555.

[17] C. W. Bauschlicher Jr., *J. Phys. Chem.* **1988**, *92*(11), 3020.

[18] D. A. Panayotov, A. I. Frenkel, J. R. Morris, *ACS Energy Lett.* **2017**, *2*(5), 1223.

[19] J. T. Seeley, M. J. Richard, P. J. Love, *J. Chem. Phys.* **2012**, *137*(22), 224109.

[20] S. P. Walch, C. W. Bauschlicher Jr., *J. Chem. Phys.* **1983**, *78*(7), 4597.

[21] D. P. Chong, S. R. Langhoff, C. W. Bauschlicher Jr., S. P. Walch, H. Partridge, *J. Chem. Phys.* **1986**, *85*(5), 2850.

[22] K. Patkowski, K. Szalewicz, *J. Chem. Phys.* **2007**, *127*(16), 164103.

[23] P. E. Blöchl, *Phys. Rev. B* **1994**, *50*(24), 17953.

[24] G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*(3), 1758.

[25] O. Crawford, B. van Straaten, D. Wang, T. Parks, E. Campbell, S. Brierley, *Quantum* **2021**, *5*, 385.

[26] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, F. T. Chong, *arXiv Preprint* arXiv:190713623 **2019**.

[27] V. Verteletskyi, T. C. Yen, A. F. Izmaylov, *J. Chem. Phys.* **2020**, *152*(12), 124114.

[28] A. Jena, S. Genin, M. Mosca, *arXiv Preprint* arXiv:190707859 **2019**.

[29] D. Kraft, *Forschungsbericht-Deutsche Forschungs-Und Versuchsanstalt Fur Luft-Und Raumfahrt* **1988**, *88*.

[30] C. A. Coulson, I. Fischer XXXIV, *Lond. Edinburgh Dublin Philos. Mag. J. Sci.* **1949**, *40*(303), 386.

[31] G. S. Kedziora, S. A. Barr, R. Berry, J. C. Moller, T. D. Breitzman, *Theor. Chem. Acc.* **2016**, *135*(3), 1.

[32] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, J. M. Martinis, *Nature* **2019**, *574*(7779), 505.

[33] K. Wright, K. M. Beck, S. Debnath, J. Amini, Y. Nam, N. Grzesiak, J. S. Chen, N. C. Pisenti, M. Chmielewski, C. Collins, K. M. Hudek, *Nat. Commun.* **2019**, *10*(1), 1.

[34] Error rate of Rigetti Aspen-M-1 quantum computer; **2022**. https://www.rigetti.com/what-we-build.

[35] R. Finsterhoelzl, G. Burkard, *Quantum Sci. Technol.* **2022**, *8*(1), 015013.

[36] H. R. Grimsley, S. E. Economou, E. Barnes, N. J. Mayhall, *Nat. Commun.* **2019**, *10*(1), 1.

[37] A. Kumar, A. Asthana, C. Masteran, E. F. Valeev, Y. Zhang, L. Cincio, S. Tretiak, P. A. Dub, *J. Chem. Theory Comput.* **2022**, *18*(9), 5312.

[38] J. Müller, W. Lavrijsen, C. Iancu, W. de Jong. Accelerating Noisy VQE Optimization with Gaussian Processes. In: *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, IEEE; **2022**. p. 215–225.

[39] Q. Sun, G. K. L. Chan, *Acc. Chem. Res.* **2016**, *49*(12), 2705.

[40] S. Endo, S. C. Benjamin, Y. Li, *Phys. Rev. X* **2018**, *8*(3), 31027.

[41] E. Kapit, *Quantum Sci. Technol.* **2017**, *2*(3), 33002.

[42] W. J. Huggins, J. R. McClean, N. C. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, R. Babbush, *NPJ Quantum Inf.* **2021**, *7*(1), 1.

[43] M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, *Gaussian 16 Rev. C. 01*, Gaussian, Inc., Wallingford, CT **2016**.

[44] M. S. Anis, H. Abraham, R. A. AduOffei, G. Agliardi, M. Aharoni, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, S. Anagolum, *Qiskit: An Open-source Framework for Quantum Computing*, Qiskit/qiskit, **2021**. https://doi.org/10.5281/ZENODO.2562111

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.