

A Non-Orthogonal Variational Quantum Eigensolver

William J. Huggins,^{1,2,3,*} Joonho Lee,^{1,2,3,†} Unpil Baek,^{4,2,3} Bryan O’Gorman,^{2,5,3} and K. Birgitta Whaley^{1,2,3}

¹*Department of Chemistry, University of California, Berkeley, CA 94720, USA*

²*Berkeley Quantum Information and Computation Center,
University of California, Berkeley, CA 94720, USA*

³*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

⁴*Department of Physics, University of California, Berkeley, CA 94720, USA*

⁵*Department of Electrical Engineering and Computer Sciences,
University of California, Berkeley, CA 94720, USA*

Variational algorithms for strongly correlated chemical and materials systems are one of the most promising applications of near-term quantum computers. We present an extension to the variational quantum eigensolver that approximates the ground state of a system by solving a generalized eigenvalue problem in a subspace spanned by a collection of parametrized quantum states. This allows for the systematic improvement of a logical wavefunction ansatz without a significant increase in circuit complexity. To minimize the circuit complexity of this approach, we propose a strategy for efficiently measuring the Hamiltonian and overlap matrix elements between states parametrized by circuits that commute with the total particle number operator. We also propose a classical Monte Carlo scheme to estimate the uncertainty in the ground state energy caused by a finite number of measurements of the matrix elements. We explain how this Monte Carlo procedure can be extended to adaptively schedule the required measurements, reducing the number of circuit executions necessary for a given accuracy. We apply these ideas to two model strongly correlated systems, a square configuration of H_4 and the π -system of Hexatriene (C_6H_8). We show that it is possible to accurately capture the ground state of both systems using a linear combination of simpler ansatz circuits for which it is impossible individually.

I. INTRODUCTION

Large, error-corrected quantum computers are expected to provide powerful new tools for understanding quantum many-body physics. For example, such devices will be able to efficiently simulate long-time dynamics [1], and through phase estimation, measure the energy of a trial wavefunction while projecting it into the eigenbasis of the Hamiltonian [2]. Prior to the availability of such devices, it is natural to ask how today’s noisy, intermediate-scale quantum (NISQ) platforms may be used for similar ends. One appealing strategy, the variational quantum eigensolver (VQE) [3, 4], uses a potentially noisy quantum computer as a black box to prepare parametrized wavefunctions and measure their energy. By optimizing over the wavefunction parameters in a classical outer loop, one obtains a variational upper bound on the true ground state energy.

While it is believed that even a noisy, modestly-sized quantum computer can prepare and measure states that are out of reach for a classical computer [5], it will still likely be difficult to take advantage of this fact to surpass the capabilities of classical variational methods [6–9]. One serious challenge is the decoherence caused by noise, which is particularly damaging for quantum chemical calculations that demand a high degree of precision [8–10]. Recent works have presented a variety of approaches to

overcoming this difficulty, including combining error detection schemes with postselection [10–12], extrapolating to the zero-noise limit [13–15], and using additional measurements and post-processing to construct better energy estimators [12, 16–18]. A complementary body of research has focused on developing new variational ansätze that use fewer gates and thus offer less opportunity for errors to occur [19–21]. In this work, we shall present a new approach in this latter direction which allows for a systematic increase in wavefunction complexity without a growing circuit depth.

The standard VQE approach uses a quantum computer to measure the expectation value of the Hamiltonian for some parametrized wavefunction, $|\psi(\boldsymbol{\theta})\rangle$, in conjunction with a classical coprocessor that interprets the measurement outcomes, suggesting new values for the parameters to minimize $\boldsymbol{\theta}$ the energy [4]. In our approach, we define instead a logical ansatz

$$|\psi(\mathbf{c}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(M)})\rangle = \sum_{i=1}^M c_i |\phi_i(\boldsymbol{\theta}^{(i)})\rangle, \quad (1)$$

where each $|\phi_i(\boldsymbol{\theta}^{(i)})\rangle$ is a parametrized wavefunction with a compact quantum circuit description. Rather than preparing the state $|\psi\rangle$ directly on our device and measuring its energy, we use our quantum computer to prepare simpler superpositions of the states $\{|\phi_i\rangle\}$ and measure the matrix elements of the Hamiltonian and overlap matrices,

$$\begin{aligned} H_{ij} &= \langle \phi_i | \hat{H} | \phi_j \rangle, \\ S_{ij} &= \langle \phi_i | \phi_j \rangle. \end{aligned} \quad (2)$$

* wjhuggins@gmail.com

† current address: Department of Chemistry, Columbia University, New York, New York 10027, USA

This allows us to classically solve a generalized eigenvalue problem,

$$H\mathbf{c} = E\mathbf{S}\mathbf{c}, \quad (3)$$

thereby finding the optimal \mathbf{c} parameters and minimizing the energy in the subspace spanned by the set of states $|\phi_i\rangle$. The $\theta^{(i)}$ values that parametrize each basis function $|\phi_i(\theta^{(i)})\rangle$ can then be optimized by a classical outer loop to lower the energy further, solving a new generalized eigenvalue problem at each step.

Our approach shares similarities with recent work in NISQ-era variational algorithms that involve solving generalized eigenvalue problems [16, 22, 23], as well as a technique from classical quantum chemistry known as non-orthogonal configuration interaction [24–26] (NOCI). However, our approach also differs from this work in some key respects. Most importantly, we make no assumptions about the form of the component wavefunctions $|\phi_i\rangle$, other than that they have efficient quantum circuit implementations. In the context of quantum algorithms, prior work has assumed that these wavefunctions are generated by excitations from a fixed reference state [16], by imaginary time evolution [22], or by the simultaneous rotation of a set of orthogonal reference wavefunctions [23].

The classical non-orthogonal configuration interaction (NOCI) method, which is the closest relative to our approach, demands each of the $|\phi_i\rangle$ wavefunctions to be a Slater determinant (not necessarily in the same single particle basis) [24–26]. It is worthwhile to note the utility of NOCI in the context of classical simulations. Unlike orthogonal CI expansions, NOCI covers a larger Hilbert space for a fixed number of determinants. Furthermore, it avoids the notorious orbital optimization problem of multiconfigurational self-consistent field theory (MCSCF) while obtaining comparable energies. Instead of directly optimizing the orbitals of the CI expansion, it separately optimizes individual determinants by finding a collection of different solutions to the Hartree-Fock equations. Once these non-orthogonal determinants are obtained, a single diagonalization of the Hamiltonian in this basis defines the NOCI energy and wavefunction. Just as NOCI can mitigate difficult orbital optimization problems to some extent, we hope to use this approach in the context of VQE to help reduce the optimization difficulty in general.

By allowing each of the basis functions $|\phi_i\rangle$ to be any wavefunction which can be represented by a parametrized quantum circuit we obtain an extremely flexible form for our logical ansatz, $|\psi\rangle = \sum c_i |\phi_i\rangle$. We shall show that, for a wide variety of ansatz circuits, the required matrix element measurements between any $|\phi_i\rangle$ and $|\phi_j\rangle$ pair can be implemented using a number of quantum gates that is equal to the sum of the gates required to prepare $|\phi_i\rangle$ and $|\phi_j\rangle$, plus a small factor that scales linearly with the system size. Notably, the per-run gate complexity is independent of the number of wavefunctions in the logical ansatz, making it possible to systematically add flexibil-

ity to $|\psi\rangle$ without increasing the required gate fidelity or coherence times of the quantum hardware.

This flexibility, however, comes at the cost of demanding more matrix element measurements. To ameliorate this cost we propose using a Monte Carlo technique to estimate the uncertainty in the ground state energy and to adaptively allocate our measurements of the matrix elements. Essentially, this scheme involves sampling from the distributions representing the uncertainty in the estimates of the Hamiltonian and overlap matrices and solving a small generalized eigenvalue problem for each sampled matrix pair. We then characterize the resulting distribution of ground state energy values by a sample variance. We suggest a heuristic that repeatedly determines which measurement to perform by calculating the sensitivity of this sample variance to additional measurements of each of the matrix elements.

We apply these ideas to two model chemical systems, a square configuration of H_4 and the π -system of hexatriene (C_6H_8), which exhibit mixed strong correlation and dynamical correlation effects. In terms of strong correlation, we shall focus on a pair of strongly entangled electrons. Specifically, there can be two exactly degenerate determinants for certain geometries of these systems while the rest of the electrons contribute to dynamical correlation. We present two types of numerical experiments. In the first, we explore how well the ground state of these systems can be represented by an NOVQE logical ansatz, varying both the complexity of the constituent basis functions and the size of the subspace. In the second, we take a fixed set of basis wavefunctions and compare our adaptive protocol for scheduling measurements with a simpler alternative.

II. THEORY

A. Matrix Element Measurement

The off-diagonal matrix elements of the Hamiltonian, $H_{ij} = \langle\phi_i|\hat{H}|\phi_j\rangle$, do not correspond to physical observables and therefore cannot be measured directly in the usual manner. Nevertheless, it is possible to construct circuits that allow us to estimate them, for example, by using the Hadamard test [27]. In this section we present a simple strategy for measuring these matrix elements. We combine ideas from recent proposals for measuring off-diagonal matrix elements that appear in other contexts [28, 29] with a trick inspired by the literature on the impossibility of black box protocols for adding controls to arbitrary unitaries [30]. Our strategy offers several benefits over a naive application of the Hadamard test. Namely, it doesn't require implementing controlled versions of the ansatz preparation circuits, and it enables the simultaneous measurement of matrix elements of multiple commuting observables while also yielding information about the overlap matrix elements, $S_{ij} = \langle\phi_i|\phi_j\rangle$.

For simplicity, we will describe below the case where

\hat{H} is a sum of commuting operators which can easily be simultaneously measured. In the more general case, the usual Hamiltonian averaging approach of grouping the terms into multiple sets that are each simultaneously measurable and measuring the sets separately can be applied without modification [4, 8, 31].

We begin by preparing the state

$$|+_{ij}\rangle := \frac{1}{\sqrt{2}}(|\phi_i\rangle|0\rangle + |\phi_j\rangle|1\rangle), \quad (4)$$

where the second register is an ancilla qubit. This task can be accomplished by using controlled versions of the unitaries \hat{U}_i and \hat{U}_j that prepare $|\phi_i\rangle$ and $|\phi_j\rangle$ from a fixed reference state. Given some quantum circuit that implements the unitaries \hat{U}_i and \hat{U}_j , it is possible to construct circuits that implement the controlled version of \hat{U}_i and \hat{U}_j , by replacing each gate in the original circuits with its controlled form. Even setting aside the difficulty of compiling such a circuit on a physical device with limited connectivity, the cost of implementing such a circuit on a near-term device (quantified by counting the number of two-qubit gates) will be substantially increased. For example, it is known that the decomposition of the Toffoli gate (the controlled-controlled-NOT gate) into a collection of single qubit and CNOT gates requires the use of six CNOT gates [32]. Given the limited coherence times and two-qubit gate fidelities of near-term hardware, we must ask if there are alternatives for implementing controlled versions of \hat{U}_i and \hat{U}_j .

An ideal protocol might allow us to implement a controlled version of an arbitrary \hat{U} using a single execution of the original, unmodified circuit that implements \hat{U} . Unfortunately, a single use of oracle (blackbox) access to a general \hat{U} is insufficient for implementing a controlled version of \hat{U} in the quantum circuit model [30]. However, if \hat{U}_i and \hat{U}_j preserve fermionic (or bosonic) excitation number and act trivially on the vacuum state, then we can circumvent this no-go result. We now show how this can be accomplished in the construction of a controlled unitary operator,

$$\hat{U}_i, \hat{U}_j \rightarrow \hat{U}_i \otimes |0\rangle\langle 0| + \hat{U}_j \otimes |1\rangle\langle 1|. \quad (5)$$

We begin with a generic input state $|\psi_0\rangle|0\rangle + |\psi_1\rangle|1\rangle$, subject to the restriction that $|\psi_0\rangle$ and $|\psi_1\rangle$ are both states that are orthogonal to the state with zero particles, $|\text{vac}\rangle$.

1. First, we adjoin an ancilla system register in the vacuum state to obtain

$$|\psi_0\rangle \otimes |\text{vac}\rangle \otimes |0\rangle + |\psi_1\rangle \otimes |\text{vac}\rangle \otimes |1\rangle.$$

2. Treating the final qubit as the control, we apply a controlled-SWAP operation between the two system registers, resulting in

$$|\psi_0\rangle \otimes |\text{vac}\rangle \otimes |0\rangle + |\text{vac}\rangle \otimes |\psi_1\rangle \otimes |1\rangle.$$

3. Next, we execute the unmodified circuit for \hat{U}_i on the first system register, while doing the same with \hat{U}_j on the second system register, yielding

$$\hat{U}_i |\psi_0\rangle \otimes |\text{vac}\rangle \otimes |0\rangle + |\text{vac}\rangle \otimes \hat{U}_j |\psi_1\rangle \otimes |1\rangle.$$

4. We follow this with a second controlled-SWAP operation to produce the state,

$$\hat{U}_i |\psi_0\rangle \otimes |\text{vac}\rangle \otimes |0\rangle + \hat{U}_j |\psi_1\rangle \otimes |\text{vac}\rangle \otimes |1\rangle.$$

5. Finally, we discard the now unentangled second system register to show completion of the action of the controlled unitary gate and obtain the desired result,

$$\hat{U}_i |\psi_0\rangle \otimes |0\rangle + \hat{U}_j |\psi_1\rangle \otimes |1\rangle.$$

For our purposes, we can take $|\psi_0\rangle$ and $|\psi_1\rangle$ to be the same fixed reference state, usually a Hartree-Fock state $|\psi_{\text{HF}}\rangle$. Then $|\phi_i\rangle = \hat{U}_i |\psi_{\text{HF}}\rangle$ and $|\phi_j\rangle = \hat{U}_j |\psi_{\text{HF}}\rangle$ and we see that with the last step we have successfully prepared the desired state, $|+_{ij}\rangle := \frac{1}{\sqrt{2}}(|\phi_i\rangle|0\rangle + |\phi_j\rangle|1\rangle)$. We then apply a Hadamard gate on the ancilla qubit and perform a \hat{Z} measurement. It is easy to see that the expectation value of \hat{Z} for the ancilla qubit will be $\langle \hat{Z}_{\text{anc}} \rangle = \text{Re} \langle \phi_i | \phi_j \rangle$. Furthermore, the post-measurement state of the system register is either

$$\frac{|\phi_i\rangle + |\phi_j\rangle}{\sqrt{2 + 2\text{Re} \langle \phi_i | \phi_j \rangle}}, \quad (6)$$

if the ancilla qubit was found to be in the +1 eigenstate, or

$$\frac{|\phi_i\rangle - |\phi_j\rangle}{\sqrt{2 - 2\text{Re} \langle \phi_i | \phi_j \rangle}}, \quad (7)$$

if the measurement outcome was -1 . These outcomes occur with probabilities $\frac{1 + \text{Re} \langle \phi_i | \phi_j \rangle}{2}$ and $\frac{1 - \text{Re} \langle \phi_i | \phi_j \rangle}{2}$ respectively.

In both cases, we proceed to measure the Hamiltonian \hat{H} on the system register. Depending on the result of the ancilla qubit measurement, the resulting expectation values will be either

$$\langle \hat{H} \rangle^{(+1)} = \frac{\langle \hat{H} \rangle_i + \langle \hat{H} \rangle_j + 2\text{Re} \langle \phi_i | \hat{H} | \phi_j \rangle}{2 + 2\text{Re} \langle \phi_i | \phi_j \rangle}, \quad (8)$$

or

$$\langle \hat{H} \rangle^{(-1)} = \frac{\langle \hat{H} \rangle_i + \langle \hat{H} \rangle_j - 2\text{Re} \langle \phi_i | \hat{H} | \phi_j \rangle}{2 - 2\text{Re} \langle \phi_i | \phi_j \rangle}. \quad (9)$$

Now we consider the expectation value of the operator $\hat{H}\hat{Z}_{\text{anc}}$. By multiplying each of the conditional expectation values of \hat{H} by the corresponding eigenvalue of \hat{Z}_{anc} and taking the appropriate weighted average, we find that

$$\langle \hat{H} \hat{Z}_{\text{anc}} \rangle = \text{Re} \langle \phi_i | \hat{H} | \phi_j \rangle. \quad (10)$$

Furthermore, if \hat{H} is a sum of Pauli operators, then the usual Hamiltonian averaging approach and upper

bounds on the variance of a VQE observable apply to Eq. 10 [31]. Therefore, by repeated measurement we can estimate $\text{Re} \langle \phi_i | \hat{H} | \phi_j \rangle$ to a fixed precision ϵ using approximately the same number of measurements that we would need to measure a diagonal matrix element to the same accuracy. A similar approach allows us to estimate $\text{Im} \langle \phi_i | \phi_j \rangle$ and $\text{Im} \langle \phi_i | \hat{H} | \phi_j \rangle$ by starting with the state $\frac{1}{\sqrt{2}}(|0\rangle |\phi_i\rangle + i |1\rangle |\phi_j\rangle)$.

B. Diagonalization With Uncertainty

Given a collection of states $\{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_n\rangle\}$, we are interested in determining the minimum energy state in the subspace that they span. To do this, we use our protocol described above to measure the matrix elements of the Hamiltonian and overlap matrices (Eq. 2), and solve the generalized eigenvalue problem (Eq. 3). Each of the matrix elements is only determined by a finite number of measurements, however, leading to some level of statistical uncertainty. In this section, we shall propose a simple Monte Carlo strategy to estimate the resulting uncertainty in the minimum eigenvalue of Eq. 3.

We model the experimentally determined values of each matrix element using a normal distribution. In practice, the experimental measurements of the matrix elements are individually described by draws from Bernoulli random variables, but the average of many such experimental outcomes is well-captured by a normal distribution [4]. When implementing the procedure described in this section on a near-term quantum computer, one could approximately determine the variance of these distributions from the experimental measurement record of the Hamiltonian and overlap matrix elements.

For the purposes of this work, we determine the variance of the Hamiltonian matrix element measurements using the upper bounds described in Refs. 4 and 31. Similarly, we observe that our scheme for measuring the overlap matrix elements will have a variance that is at most $\frac{1}{m}$, where m is the number of measurements performed, and we use this upper bound as an approximation to the true variance. We use these approximations both in our simulation of the experimental measurement record and in our subsequent protocol to determine the uncertainty in the ground state energy. Throughout this section, we use a notation which separates the intrinsic component of the variance, which we denote by σ^2 , from the scaling with the number of measurements, m .

Experimentally, we only have access to estimates of $\langle \phi_i | \hat{H} | \phi_j \rangle$ and $\langle \phi_i | \phi_j \rangle$ from our measurement record, which we denote by \tilde{h}_{ij} and \tilde{s}_{ij} . Taken together with our estimates of the variances, $\tilde{\sigma}_{H_{ij}}^2$ and $\tilde{\sigma}_{S_{ij}}^2$, we can

define the random variables

$$\tilde{H}'_{ij} = \tilde{h}_{ij} + \frac{\tilde{\sigma}_{H_{ij}}}{\sqrt{m_{H_{ij}}}} \mathcal{N}(0, 1), \quad (11)$$

$$\tilde{S}'_{ij} = \tilde{s}_{ij} + \frac{\tilde{\sigma}_{S_{ij}}}{\sqrt{m_{S_{ij}}}} \mathcal{N}(0, 1). \quad (12)$$

These distributions represent our uncertainty about the true value of the matrix elements given the limited information provided by our experimental data.

To quantify the corresponding uncertainty in the ground state energy in the NOVQE subspace, we use a Monte Carlo sampling procedure. We accomplish this by repeatedly drawing from the distributions \tilde{H}'_{ij} and \tilde{S}'_{ij} , and solving the resulting generalized eigenvalue problems. It is possible that the noise in our matrix element measurements and subsequent sampling destroys the positive semi-definite character of the overlap matrix. To deal with this, we follow the canonical orthogonalization procedure described in Ref. 33, discarding the eigenvalues of the sampled overlap matrices that are less than some numerical cutoff (and their associated eigenvectors). Each sampled pair of matrices yields a sample from the unknown distribution over possible NOVQE ground state energies. We then quantify our uncertainty in our estimate of this lowest eigenvalue by calculating the sample variance, σ_{MC}^2 .

1. Experiment Design Heuristic

In the previous section, we proposed a Monte Carlo scheme for estimating the uncertainty in the NOVQE ground state energy caused by a finite number of measurements of the individual matrix elements. Here we build on this proposal to determine the relative impact of performing additional measurements. Ultimately, our goal is to create a reasonable heuristic for adaptively scheduling measurements to most efficiently use a limited amount of device time. By repeatedly sampling from \tilde{H}'_{ij} and \tilde{S}'_{ij} and solving the resulting generalized eigenvalue problems, we obtain a distribution over NOVQE ground state energies with some mean μ_{MC} and standard deviation σ_{MC} .

We determine the impact of additional measurements of the matrix elements, H_{ij} and S_{ij} , on the uncertainty in the ground state energy by calculating the derivatives of the sample standard deviation, σ_{MC} , with respect to the number of measurements performed, $m_{H_{ij}}$ and $m_{S_{ij}}$. The resultant quantities, $\frac{d\sigma_{\text{MC}}}{dm_{H_{ij}}}$ and $\frac{d\sigma_{\text{MC}}}{dm_{S_{ij}}}$, estimate how much we expect the sample deviation to shrink if we perform additional measurements of H_{ij} or S_{ij} . Note that we take these derivatives only with respect to $m_{H_{ij}}$ and $m_{S_{ij}}$ in the Monte Carlo sampling procedure of Eqs. 11 and 12, not the original measurements on the device. Therefore, no additional quantum resources are required. We use the TensorFlow software package to perform the

Monte Carlo sampling of \tilde{H}'_{ij} and \tilde{S}'_{ij} , calculate of the ground state energies, and estimate σ_{MC} [34]. This enables us to evaluate the analytical expressions for each of $\frac{d\sigma_{\text{MC}}}{dm_{H_{ij}}}$ and $\frac{d\sigma_{\text{MC}}}{dm_{S_{ij}}}$ (for a fixed set of samples drawn from \tilde{H}'_{ij} and \tilde{S}'_{ij}) without explicitly deriving the equations.

To optimally allocate our experimental measurements, we begin by performing a small number of measurements of each matrix element. We then estimate the derivatives $\frac{d\sigma_{\text{MC}}}{dm_{H_{ij}}}$ and $\frac{d\sigma_{\text{MC}}}{dm_{S_{ij}}}$. Using these estimates, we simply choose to perform additional measurements on the matrix element whose corresponding derivative is the most negative. In practice, we perform these measurements in small batches so that the time taken by the classical processing of the measurement results is small compared to the time performing the measurements. By repeating this process for many steps, until we either achieve the desired accuracy or exhaust a pre-defined measurement budget, we will approximately optimize allocation of measurements between the different terms.

C. Implementation

The tools presented above are applicable for use with a variety of different ansätze, and subject only to the constraint that the circuits act on a common reference state and conserve fermionic excitation number in order to benefit from the efficient implementation of the effective controlled unitaries. For our numerical experiments, we shall focus on a particular class of wavefunctions known as ***k*-fold products of unitary paired coupled cluster with generalized single and double excitations** [20] (*k*-UpCCGSD). These wavefunctions have the appealing properties that 1) the required circuit depth scales only linearly in the size of the system, and 2) they can be systematically improved by increasing the refinement parameter *k*. We briefly review this ansatz below and then describe in more detail the implementation details of our numerical experiments.

1. The *k*-UpCCGSD Ansatz

The essential idea behind the *k*-UpCCGSD ansatz is to act on a reference state, Hartree-Fock in the case of this paper, with a product of *k* elementary blocks. Each block is an independently parametrized approximation to a unitary coupled cluster circuit generated by a sparse cluster operator containing only single and paired double excitations [35]. To this end, the wavefunction (before the Trotter approximation involved in compiling the circuits) is defined as follows.

$$|\psi\rangle = \prod_{x=1}^k \left(e^{\hat{T}^{(x)} - \hat{T}^{(x)\dagger}} \right) |\phi_0\rangle, \quad (13)$$

where each cluster operator

$$\hat{T} = \sum_{ia} t_{ii}^{aa} \hat{a}_{a\alpha}^\dagger \hat{a}_{a\beta}^\dagger \hat{a}_{i\beta} \hat{a}_{i\alpha} + t_i^a (\hat{a}_{a\alpha}^\dagger \hat{a}_{i\alpha} + \hat{a}_{a\beta}^\dagger \hat{a}_{i\beta}). \quad (14)$$

possesses an independent collection of variational parameters. (We omit the (*x*) superscript for simplicity and use Latin and Greek letters for spatial and spin indices respectively.)

In contrast with the standard unitary coupled cluster single and doubles (UCCSD), *k*-UpCCGSD only includes doubles excitations which collectively move a pair of electrons from one spatial orbital to another. The resulting loss of flexibility is ameliorated by the use of generalized excitations that do not distinguish between occupied and unoccupied orbitals [36, 37], and the *k*-fold repetition of the elementary circuit block. As a result, the number of free parameters in the ansatz scales as $O(kN^2)$. We make use of the generalized swap networks of Ref. 38 to implement a single Trotter step approximation to the *k*-UpCCGSD ansatz with the open source Cirq and OpenFermion-Cirq libraries [39]. The resulting circuits have a gate depth of $O(kN)$ and use $O(kN^2)$ two qubit gates.

2. Computational Details

The quantum chemical calculations of FCI ground state and HF reference were performed using the open source packages Psi4 and OpenFermion [40, 41]. We optimized the ground state energy in the NOVQE subspace by varying the parameters of the most recently added ansatz wavefunction and diagonalizing the Hamiltonian and overlap matrices as described above. The variational parameters were randomly initialized from a normal distribution with mean 0 and variance σ_i^2 . After some brief experimentation, for H_4 we set $\sigma_i^2 = 10^{-4}$, while for Hexatriene we set $\sigma_i^2 = 9.0$. We found the larger variation in initial parameters for Hexatriene to be mildly helpful in avoiding repeated convergence to states with a high overlap when a large number of NOVQE states were added, an issue that we shall discuss below. For the numerical optimization, we used the Scipy implementation of the quasi-Newton limited-memory BFGS (L-BFGS-B) method [42, 43]. We calculated the gradient at each step using a finite difference method with a step size of $\delta = 10^{-6}$. For H_4 we used at most $C = 5000$ calls to our objective function for each state added, while for Hexatriene we allowed for $C = 10000$.

III. RESULTS

H_4 is often used as a small testbed for single-reference coupled-cluster methods [44–48]. We shall focus on the square (D_{4h}) geometry here. The system exhibits two exactly degenerate determinants at the D_{4h} geometry.

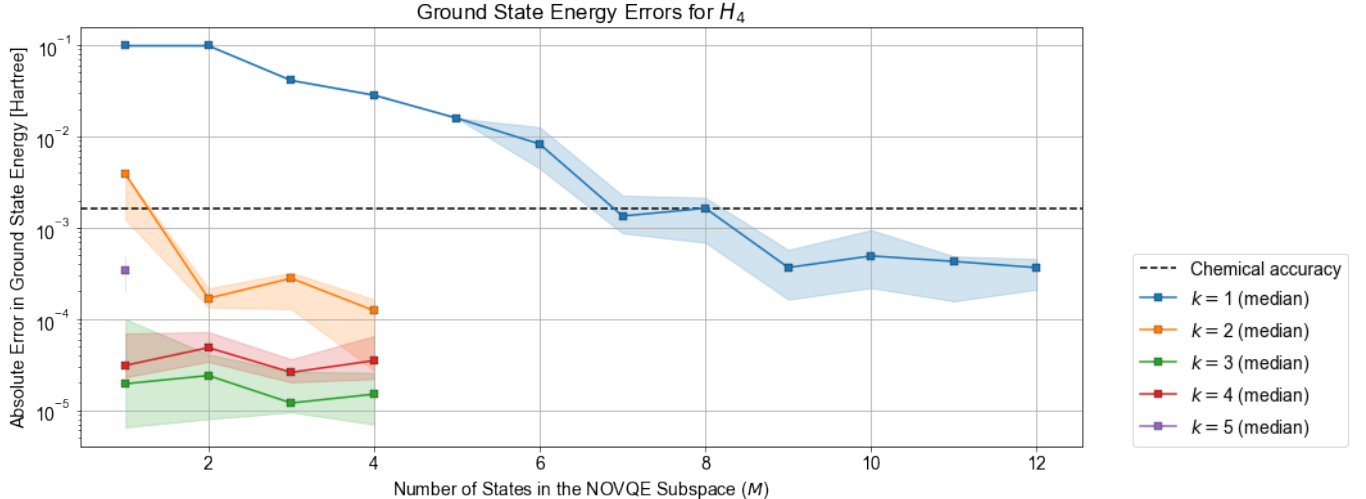


Figure 1. Difference between NOVQE energies and FCI energies for the ground state of H_4 for a variety of ansätze and sizes of the NOVQE subspace. Each independent calculation is plotted as a separate point, median values are connected by lines, and the 25th to 75th percentiles are shaded. The dotted horizontal line indicates 1 kcal/mol \approx 1.59 millihartree, a commonly accepted value for “chemical accuracy”. As more states are added to the NOVQE subspace, the error in the ground state energy declines substantially for the $k = 1$ and $k = 2$ versions of k -UpCCGSD.

Therefore, it will exhibit effects of both strong and weak correlation.

Another important class of chemical systems to investigate is hydrocarbons. In this work, we shall study a simple hydrocarbon, hexatriene (C_6H_8). The interesting aspect of this molecule is that the torsional PES of a double bond leads to a strong correlation problem. At $\theta = 90^\circ$, it exhibits two exactly degenerate determinants and therefore it is strongly correlated. To form the active space, we include the entire set of π electrons in the system along with both Π and Π^* orbitals. The resulting active space is then (6e, 6o), and this also possesses a good mixture of strong and weak correlation.

In the following section, we present the results of two types of experiments related to our proposed NOVQE approach on these chemical systems and discuss the potential utility of our approach for more general chemical problems. With the first class of experiment, we focus on understanding how effectively the ground state can be represented by a linear combination of parametrized wavefunctions, optimized using the gradient-based approach we described above. We vary both the complexity of the individual ansatz wavefunctions by adjusting the number of circuit blocks (k) in the k -UpCCGSD ansatz and the number of states (M) in the NOVQE subspace. For these calculations, we neglect the challenges posed by a finite number of measurements and the impact of circuit noise. In our second set of numerical experiments, we explore the extent to which our proposal for an adaptive measurement scheme is successful in reducing the number of circuit repetitions required to resolve the NOVQE ground state energy to a fixed precision.

A. NOVQE Ground State Energies

1. Hydrogen

Figure 1 presents data on the application of NOVQE to the square geometry of H_4 with fixed bond distance $R_{H-H} = 1.23 \text{ \AA}$ in a minimal STO-3G basis set, an $N = 8$ qubit problem. We consider the performance of the k -UpCCGSD ansatz for $k = 1$ to $k = 5$ with $M = 1$ up to $M = 12$ states in the NOVQE subspace, noting that $M = 1$ is equivalent to the regular VQE procedure. In all cases we consider the error in the median ground state energy found by the optimization procedure as a proxy for ansatz’s ability to describe the ground state.

Focusing first on understanding the behavior of the wavefunctions in the context of the standard VQE approach ($M = 1$), we note that while the ground state energy estimate dramatically improves with increasing k between $k = 1$ and $k = 3$, this trend fails and then reverses with higher values of k . Given that any state reachable by $k = 3$ is also trivially reachable by an ansatz with a larger value of k , we believe the difficulty in properly optimizing the richer ansatz is responsible for this behavior, particularly the challenges caused by the appearance of local minima or nearly vanishing gradients. Some preliminary experiments that used parameter values optimized with $k = 1$ to initialize calculations performed using richer, $k > 1$, circuits were not immediately successful. Exploration of more sophisticated initialization and optimization strategies constitutes a useful task for future work.

For $k = 1$ and $k = 2$, we observe that we can systematically improve the accuracy of the estimated ground

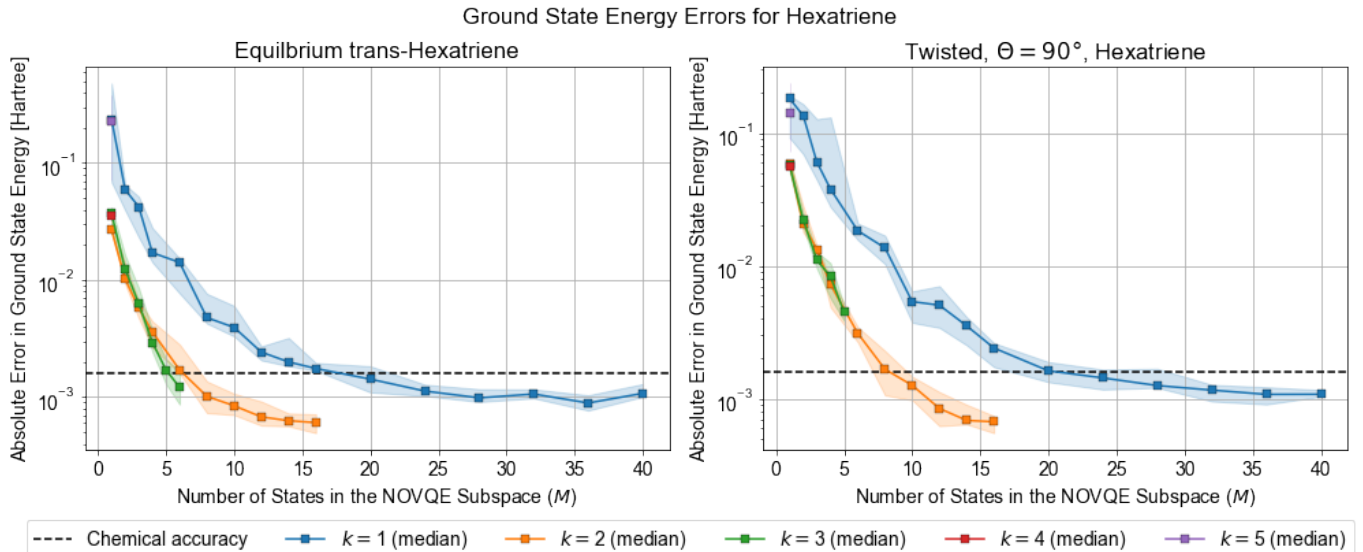


Figure 2. Difference between NOVQE energies and FCI energies for the ground states of the 90° and 180° twisted configurations of Hexatriene for a variety of ansätze and sizes of the NOVQE subspace. Each independent calculation is plotted as a separate point, median values are connected by lines, and the 25th to 75th percentiles are shaded. The dotted horizontal line indicates 1 kcal/mol \approx 1.59 millihartree, a commonly accepted value for “chemical accuracy”. By increasing the number of states in the NOVQE subspace, we are able to drive the error below the threshold for chemical accuracy even though the system is challenging to treat using any of the ansätze we consider here in the regular VQE framework ($M = 1$ in the plot).

state energy by increasing the number of states included in the NOVQE subspace (M). Given a sufficient value of M , all of the wavefunction ansätze we consider are able to achieve absolute errors well below chemical accuracy. This supports our thesis that a collection of ansatz which are individually insufficient to capture the desired state may be fruitfully combined to yield a logical ansatz of sufficient flexibility and optimizability. Interestingly, for this particular system, it appears to be more effective to add parameters to the individual ansatz by increasing k , than to add parameters to the logical ansatz by increasing M . This suggests that the most effective way to apply the NOVQE may be to work with the richest individual ansatz which can be afforded given the coherent quantum resources available, and to include additional states only when this limit is reached. The unsophisticated initialization and optimization strategy for subsequent states may also play a factor in the slow convergence with increasing M .

2. Hexatriene

Here we present our results for the ground state energy of two molecular configurations of Hexatriene (C_6H_8) in an STO-3G basis with an active space of 6 electrons in 6 π orbitals ($N = 12$ qubits). The geometries are given in A. As with our calculations on H_4 , we run the NOVQE experiments over circuit depths $k = 1$ to 5, but here we consider subspace sizes as large as $M = 40$, due to the system’s increased complexity. In Figure 2 we show

the calculations for the equilibrium geometry (obtained by performing geometry optimization using density functional theory) and a configuration with a 90° twist on the central Carbon-Carbon double bond respectively. We provide the geometries for these two configurations in the appendix in Table I and Table II.

Once again we notice that increasing the circuit complexity by taking larger values of k initially provides a substantial benefit but quickly saturates, with no benefit to taking k higher than 2 in this case. By contrast, adding additional states to the NOVQE subspace drives the error down below the threshold for chemical accuracy even with the most limited ansatz. We note that the marginal benefit of each additional state drops sharply with increasing values of M . This is reminiscent of the observation in conventional electronic structure calculations that a small number of determinants are sufficient to capture most of the wavefunction, but a long tail of dynamic correlation can result in a slow convergence to the true ground state as determinants are added to the variational space [49]. The classical intractability of calculating matrix elements between different coupled cluster wavefunctions means that relatively little work has been done on the representational power of wavefunctions like those used in NOVQE. In the future, it would be interesting to determine which of the general trends we have observed are specific to our choice of ansatz and optimization strategies and which would generically apply to any realization of NOVQE.

Figure 3 shows the relative energies obtained by pairing specific calculations from each configuration together

and taking the difference. In the low M regime we notice that the absolute errors are somewhat lower for the equilibrium configuration than for the twisted configuration, which is expected when using a linear combination of single-reference wavefunctions for a strongly correlated system. As we increase M , this difference largely disappears, a phenomenon that leads to some favorable cancellation of errors when calculating the relative energies between the two configurations.

B. NOVQE Matrix Element Measurements

1. H_4

In the previous subsection we presented data on the performance of NOVQE in the absence of noise during the circuit execution and measurement process. Now we consider the effects of statistical noise during measurement. Specifically, we determine how many circuit repetitions are necessary to evaluate the ground state energy within a target precision for a subspace defined by a fixed set of NOVQE states. For simplicity, we do not combine this analysis with an investigation of the optimization procedure. Instead, we take the optimized circuit parameters for a collection of M NOVQE states and compare the effectiveness of the adaptive protocol we described in [Section II B 1](#) to a simpler alternative for determining the ground state energy in the subspace spanned by the optimized states, which we shall explain below.

The simpler protocol, which we shall refer to as non-adaptive, consists of measuring each matrix element of the Hamiltonian and overlap matrices the same number of times. For the adaptive protocol, we repeatedly use the procedure described in [Section II B 1](#) to select a particular matrix element and perform measurements in batches of 10^5 . For the purpose of this comparison, we treat a ‘measurement’ of a particular Hamiltonian or overlap matrix element as a draw from a Gaussian random variable whose mean is the true value of the matrix element and whose variance is set by the upper bound described in [Ref. 4](#), scaled by the number of measurements performed. Note that in a real experiment, or a finer-grained simulation, the Hamiltonian has to be decomposed into groups of terms that can be simultaneously measured, and one could apply an adaptive scheme like the one we propose to schedule measurements between these groups as well. For both kinds of numerical experiments we calculate a 2σ error bar using a bootstrapping sample size of 200 using the techniques of [Section II B](#).

In [Figure 4](#) we plot the actual trajectories of the estimates for the ground state energies, together with their error bars for both the adaptive and non-adaptive approaches to measurement, for two realizations of this numerical experiment. In both cases, we see that the adaptive protocol converges more quickly towards the NOVQE ground state energy than the non-adaptive one. Panel A corresponds to a typical realization of this exper-

iment, one where the optimization procedure that generated the states was successful and the resulting generalized eigenvalue problem is well-conditioned. While the majority of the numerical experiments with H_4 were qualitatively similar to the results presented in panel A, we occasionally saw instances like those in panel B where the NOVQE procedure failed or the error bars were dramatically larger.

For the particular experiment shown in panel B, the large error bars and the 147.6 millihartree gap between the ground state energy in the NOVQE subspace and the FCI ground state derive from the same origin. In this instance, the optimization procedure which generated the fixed NOVQE states that were used to compare the measurement strategies became stuck and added several states that were nearly identical to each other. This resulted in an almost singular overlap matrix whose entries, therefore, had to be resolved to an extreme degree of precision before a good solution to the generalized eigenvalue problem could be found. This suggests that a practical implementation of NOVQE would benefit from including some regularization during the optimization that penalizes states whose overlap becomes too high.

Examining panel A, we find that the data qualitatively supports the assumption that the variance in the ground state energy estimate settles into an asymptotic regime where its behavior is well described by the relationship

$$\sigma^2(N) \approx \frac{\kappa}{N}, \quad (15)$$

where N indicates the number of measurements performed and κ is some constant. For this particular realization, we find κ to be approximately $600 E_h^2$ for the non-adaptive scheme, and approximately $40 E_h^2$ for the adaptive one. Using the same upper bounds to calculate the variance for a regular VQE calculation performed on the same system would yield $\kappa \approx 28 E_h^2$. Therefore, for these applications to H_4 , resolving the ground state energy to some target accuracy using our NOVQE approach requires only a modest increase in the number of measurements compared to the standard VQE.

2. Hexatriene

As in our analysis of H_4 , we compare the proposed adaptive approach to distributing measurements between the elements of the Hamiltonian and overlap matrices with a non-adaptive one, by choosing a collection of optimized NOVQE states and applying both methods to determine the ground state energy in the resulting subspace. In this case we choose to use $M = 16$ states, each of which is generated by a $k = 2$ k -UpCCGSD circuit, and focus on the equilibrium configuration of trans-Hexatriene. Examining [Figure 5](#), we see immediately that the increased difficulty of this problem compared to H_4 is reflected in a much larger gap between the FCI

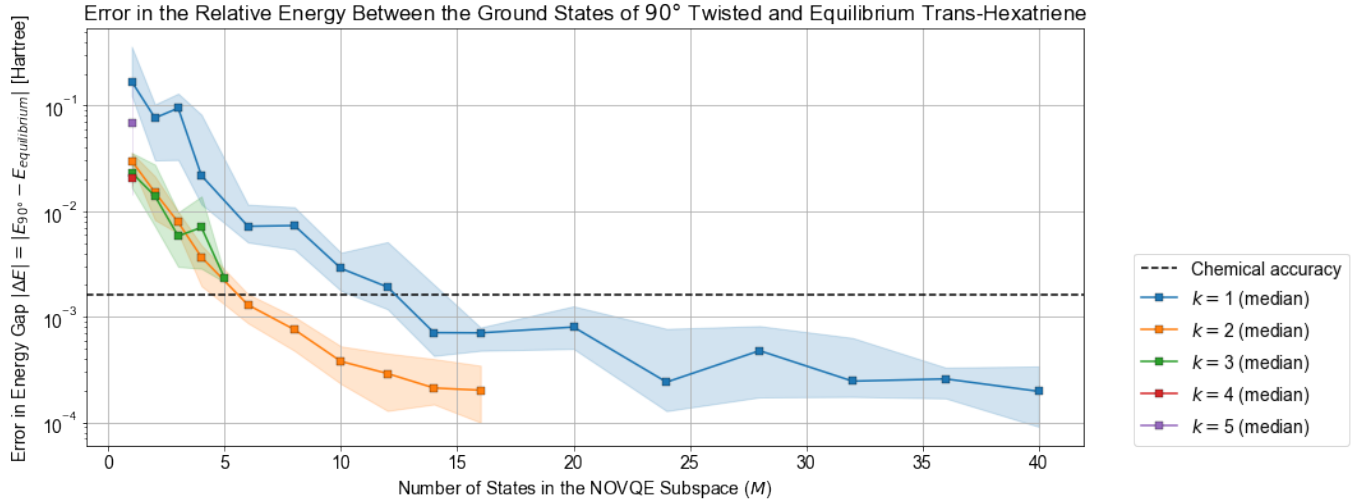


Figure 3. Error in the relative energies between the ground states of the two Hexatriene configurations calculated using NOVQE for a variety of ansätze and sizes of the NOVQE subspace. Each independent calculation of the relative energy is plotted as a separate point, median values are connected by lines, and the 25th to 75th percentiles are shaded. The dotted horizontal line indicates 1 kcal/mol \approx 1.59 millihartree, a commonly accepted value for “chemical accuracy”. We note that the errors here are lower than those for the absolute energies presented in Figure 2, indicating that we benefit from a significant cancellation of errors when calculating the relative energies.

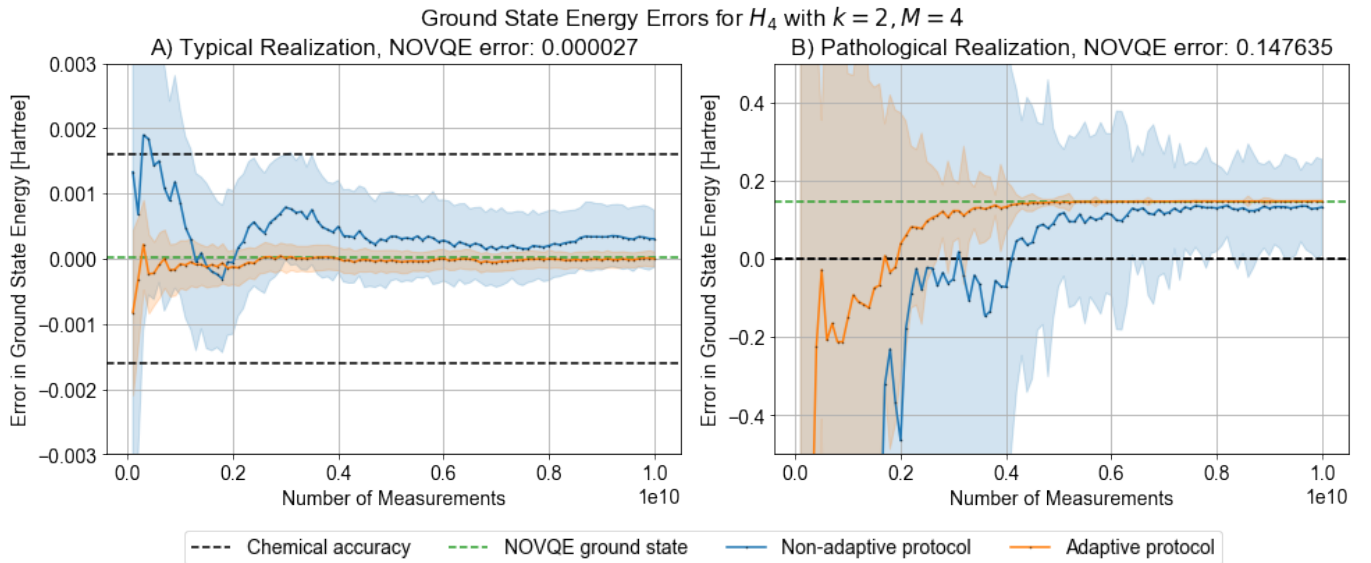


Figure 4. We compare the ability of the adaptive and non-adaptive schemes for scheduling measurements to resolve the ground state energy of H_4 in two different NOVQE subspaces of $M = 4$ optimized $k = 2$ k -UpCCGSD states. The trajectories of the estimated ground state energy are plotted in solid lines together with 2σ error bars indicated by the shaded regions. The actual energies of the ground states in the NOVQE subspaces are indicated with dashed green lines. In panel A we feature a typical NOVQE subspace with a well-conditioned overlap matrix. In panel B we present an NOVQE subspace where the optimization procedure failed to find the ground state and included multiple nearly identical states, resulting in a poorly conditioned overlap matrix. In both cases, the adaptive protocol converges significantly more quickly than the non-adaptive one.

ground state and the ground state in the NOVQE subspace, as well in a larger number of measurements required for convergence.

Figure 5 shows the same substantial difference between the performances of the adaptive and non-adaptive ap-

proaches that was observed for H_4 . Here we see that the true ground state of the subspace lies outside of the error bars for the non-adaptive scheme during a portion of the measurement procedure. Additional bootstrap samples might address this behavior, but we note that the

adaptive scheme moves quickly to a regime where the uncertainty estimates are reliable even with a small number of samples. We once again observe that the variance qualitatively converges with the expected long-time $\frac{1}{N}$ behavior of Eq. 15 for most of the numerical experiment. Therefore, we can determine the ‘intrinsic variance’, κ , of each method and compare their statistical efficiencies.

For the non-adaptive scheme we observe $\kappa \approx 4 \times 10^6$, while for the adaptive scheme we see $\kappa \approx 1.3 \times 10^5$, with $\kappa \approx 1.6 \times 10^2$ being the reference value for a regular VQE calculation, determined using the same bounds assumed throughout this comparison. Comparing with the simpler H_4 example, we see that the adaptive scheme for measuring the NOVQE ground state energy of Hexatriene results in an even greater gain when compared to the non-adaptive scheme, but now falls short of the goal of reducing the cost to be comparable to that of a regular VQE approach. If the NOVQE approach is to be useful for such systems in the future, it will likely be necessary to develop further strategies to reduce the measurement overhead. One promising avenue is the adaptation of recently proposed strategies for measurement in the context of regular VQE to NOVQE [10]. These strategies have been shown to reduce the number of circuit executions by orders of magnitude when compared with the bounds used to derive the number of measurements in this work.

IV. DISCUSSION AND OUTLOOK

We have introduced an extension to the variational quantum eigensolver that calls for the ground state energy to be approximated by solving a generalized eigenvalue problem in the subspace spanned by a linear combination of M parametrized quantum wavefunctions. The resulting logical wavefunction ansatz is a linear combination of all M states in the subspace but can be determined by pairwise measurements of the Hamiltonian and overlap matrices. Therefore, it is possible to increase the flexibility of the ansatz without requiring additional coherent quantum resources. By analogy with the non-orthogonal configuration interaction method of classical quantum chemistry, we call our approach the non-orthogonal variational quantum eigensolver, NOVQE.

Our proposal necessitates off-diagonal measurements of the Hamiltonian and overlap matrices. We perform these using a modified Hadamard test. Naively, this would require us to implement controlled versions of the quantum circuits for state preparation. To avoid this cost, we demanded that the state preparation circuits all act on a common reference state and preserve fermionic excitation number. This allowed us to avoid the need to add controls to the ansatz circuits by instead performing controlled swap operations between two copies of the system register, a cost that scales linearly and modestly with the system size.

To determine the ground state energy in the subspace,

our approach requires that we measure all M^2 elements of the Hamiltonian and overlap matrices in the NOVQE subspace. We presented a statistical strategy for estimating the uncertainty in the resultant ground state energy estimate for a given uncertainty in the matrix elements. We also pointed out how the machinery that generates these estimates can be leveraged in a Monte Carlo sampling process to determine which matrix element should be chosen for additional measurements to optimally reduce the uncertainty. We proposed an iterative approach, in which small batches of measurements are repeatedly performed according to this Monte Carlo prescription, to minimize the overall number of circuit repetitions required by our NOVQE method.

We demonstrated an implementation of our approach using a collection of k -UpCCGSD wavefunctions to approximate the ground state of two model strongly-correlated systems, a square geometry of H_4 and the π -space of Hexatriene in two configurations. Growing the NOVQE subspace by adding and optimizing one state at a time, we showed how a collection of ansätze that are individually incapable of well approximating the ground state can be fruitfully combined to form a logical ansatz that does provide a good approximation. Interestingly, we observed that making the individual ansatz wavefunctions more flexible by increasing k from 1 to 2 yielded a large benefit, but that adding more parameters did not. In contrast, adding states to the NOVQE subspace yielded a consistent but slow improvement of the ground state estimate. While these observations are highly coupled to our particular implementation choices, they suggest that increasing the number of NOVQE states may be most useful when optimization challenges or coherence time limits the ability to improve the individual wavefunctions.

To characterize our proposal for adaptively scheduling measurements to minimize the number of circuit repetitions required by our approach, we focused on quantifying the number of measurements required to approximate the ground state energy in a fixed NOVQE subspace. For the purposes of this investigation we approximated the variance of the individual matrix element measurements using the bounds described in Refs. 4 and 31. For both our square H_4 and our equilibrium configuration of trans-Hexatriene, we optimized a set of NOVQE states and froze their parameters. We then applied our adaptive approach for scheduling measurements and compared it to a simpler non-adaptive scheme, in which each matrix element was measured the same number of times. We found that our adaptive approach used approximately 12 – 15 times fewer measurements for H_4 and approximately 30 times fewer measurements for Hexatriene than the non-adaptive alternative did. As a third point of comparison, we also considered the number of measurements required for a standard VQE approach and saw that our proposed method was within a factor of two of this baseline for most of our experiments with H_4 . In contrast with this optimistic result, our experiments with Hexatriene indi-

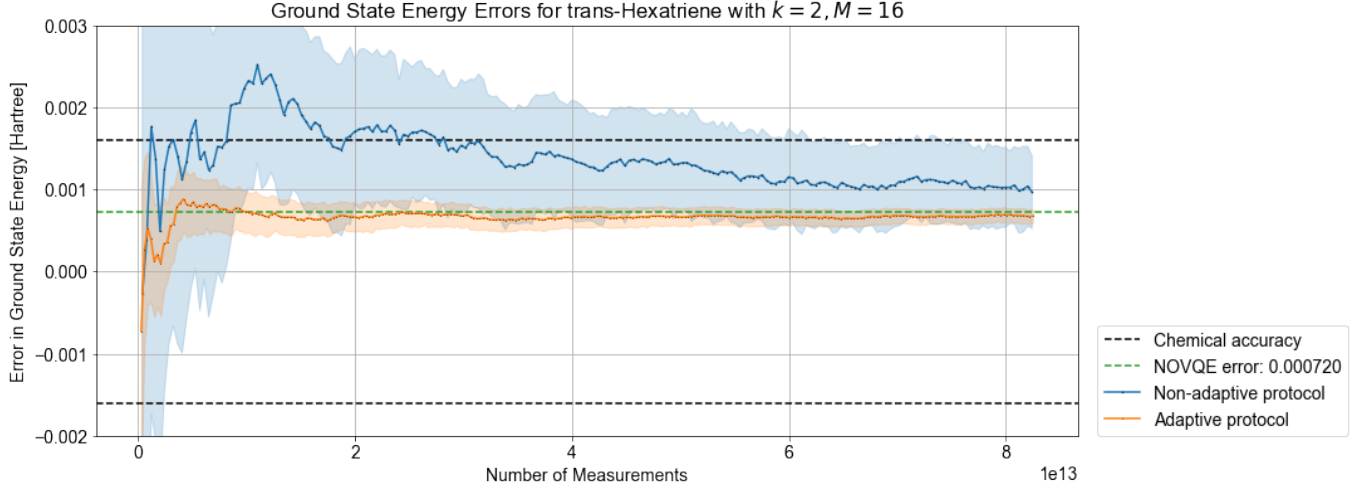


Figure 5. We compare the ability of the adaptive and non-adaptive schemes for scheduling measurements to resolve the ground state energy of trans-Hexatriene in a particular NOVQE subspace of $M = 16$ optimized $k = 2$ k -UpCCGSD states. The trajectories of the estimated ground state energy are plotted in solid lines together with 2σ error bars indicated by the shaded regions. The actual energy of the ground state in the NOVQE subspace is indicated with a dashed green line. The adaptive protocol converges much more quickly towards the correct value but still requires a large number of circuit repetitions to do so.

cated the need for almost three orders of magnitude more circuit repetitions than a normal VQE calculation, even with the benefit of our adaptive scheme for scheduling measurements.

We can imagine several routes towards ameliorating this challenge and developing NOVQE further. First of all, having states that are nearly linearly dependent in the NOVQE subspace can dramatically increase the cost of measurement. Developing an optimization strategy that regularizes this behavior away would be useful. Related to this is the possibility of extending the tools for measuring analytical gradients of parametrized quantum circuits to work with the NOVQE formalism. Another avenue for future work would be the development of good initialization strategies for NOVQE, potentially using reference states derived from a classical NOCI calculation. Finally, recent work has shown that a measurement strategy based on factorizations of the two-electron integral tensor can dramatically reduce the cost of the standard VQE approach [10], especially when compared to the type of bounds used throughout this paper [4, 31]. Adapting this approach for use with NOVQE is likely to offer a significant improvement.

Beyond these modifications to the NOVQE approach outlined in this paper, it is also conceivable that the tools we have presented might be usefully employed in other ways. For example, we have focused on the optimization of a logical ansatz that is a superposition of individual parametrized wavefunctions. An alternative is to take inspiration from Ref. 22, and from the classical NOCI method [24–26]. Following these approaches, one could optimize the individual wavefunctions separately, and solve the generalized eigenvalue problem only once with the final collection of states. Another possible di-

rection to pursue is the inclusion of one or more states in the NOVQE subspace that can be classically optimized. Along these lines, one could envisage using tensor network techniques to efficiently optimize an approximation to the ground state with a classical computer and only turn to the use of parametrized quantum circuits to prepare small corrections to the state, thereby reducing the number of measurements required.

In summary, this work has presented a promising new extension to the VQE formalism and highlighted both its advantages and its drawbacks. We have also presented a strategy for compiling off-diagonal matrix element measurements and promoted a general approach to Monte Carlo estimation of uncertainty, which may be of independent interest. Moreover, our circuit simulations of the k -UpCCGSD ansatz deepen the analysis of Ref. 20. We believe that the ability of our NOVQE to trade off coherent quantum resources for additional measurements may prove to be a useful tool in making use of NISQ-era quantum hardware for studying challenging strongly correlated systems.

In the final stages of preparing this manuscript a related work was posted which independently developed a similar approach in the context of variational quantum algorithms for solving linear systems of equations [50].

V. ACKNOWLEDGEMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Quantum Algorithm Teams Program, under contract number DE-AC02-05CH11231. Additionally, this work was supported by a Quantum AI-

gorithms Focused Award from Google LLC. B.O. was

supported by a NASA Space Technology Research Fellowship.

-
- [1] S Lloyd, “Universal quantum simulators,” *Science* **273**, 1073–1078 (1996).
 - [2] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon, “Simulated quantum computation of molecular energies,” *Science* **309**, 1704–1707 (2005).
 - [3] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” *Nat. Commun.* **5**, 4213 (2014).
 - [4] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik, “The theory of variational hybrid quantum-classical algorithms,” *New J. Phys.* **18**, 023023 (2016).
 - [5] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven, “Characterizing quantum supremacy in near-term devices,” *Nat. Phys.* **14**, 595–600 (2018).
 - [6] Dave Wecker, Matthew B Hastings, and Matthias Troyer, “Progress towards practical quantum variational algorithms,” *Phys. Rev. A* **92**, 042303 (2015).
 - [7] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven, “Barren plateaus in quantum neural network training landscapes,” *Nat. Commun.* **9**, 4812 (2018).
 - [8] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta, “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets,” *Nature* **549**, 242–246 (2017).
 - [9] J I Colless, V V Ramasesh, D Dahlen, M S Blok, M E Kimchi-Schwartz, J R McClean, J Carter, W A de Jong, and I Siddiqi, “Computation of molecular spectra on a quantum processor with an Error-Resilient algorithm,” *Phys. Rev. X* **8**, 011021 (2018).
 - [10] William J Huggins, Jarrod McClean, Nicholas Rubin, Zhang Jiang, Nathan Wiebe, K Birgitta Whaley, and Ryan Babbush, “Efficient and noise resilient measurements for quantum chemistry on Near-Term quantum computers,” *arXiv:1907.13117* (2019).
 - [11] Sam McArdle, Xiao Yuan, and Simon Benjamin, “Error-Mitigated digital quantum simulation,” *Phys. Rev. Lett.* **122**, 180501 (2019).
 - [12] X Bonet-Monroig, R Sagastizabal, M Singh, and T E O’Brien, “Low-cost error mitigation by symmetry verification,” *Phys. Rev. A* **98**, 062339 (2018).
 - [13] Kristan Temme, Sergey Bravyi, and Jay M Gambetta, “Error mitigation for Short-Depth quantum circuits,” *Phys. Rev. Lett.* **119**, 180509 (2017).
 - [14] Matthew Otten and Stephen K Gray, “Accounting for errors in quantum algorithms via individual error reduction,” *npj Quantum Information* **5**, 11 (2019).
 - [15] Abhinav Kandala, Kristan Temme, Antonio D Córcoles, Antonio Mezzacapo, Jerry M Chow, and Jay M Gambetta, “Error mitigation extends the computational reach of a noisy quantum processor,” *Nature* **567**, 491–495 (2019).
 - [16] Jarrod R McClean, Mollie E Kimchi-Schwartz, Jonathan Carter, and Wibe A de Jong, “Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states,” *Phys. Rev. A* **95**, 042308 (2017).
 - [17] Jarrod R McClean, Zhang Jiang, Nicholas C Rubin, Ryan Babbush, and Hartmut Neven, “Decoding quantum errors with subspace expansions,” *arXiv:1903.05786* (2019).
 - [18] R Sagastizabal, X Bonet-Monroig, M Singh, M A Rol, C C Bultink, X Fu, C H Price, V P Ostroukh, N Muthusubramanian, A Bruno, M Beekman, N Haider, T E O’Brien, and L DiCarlo, “Error mitigation by symmetry verification on a variational quantum eigensolver,” *arXiv:1902.11258* (2019).
 - [19] Ian D Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush, “Quantum simulation of electronic structure with linear depth and connectivity,” *Phys. Rev. Lett.* **120**, 110501 (2018).
 - [20] Joonho Lee, William J Huggins, Martin Head-Gordon, and K Birgitta Whaley, “Generalized unitary coupled cluster wave functions for quantum computation,” *J. Chem. Theory Comput.* **15**, 311–324 (2019), <https://doi.org/10.1021/acs.jctc.8b01004>.
 - [21] Harper R Grimsley, Sophia E Economou, Edwin Barnes, and Nicholas J Mayhall, “An adaptive variational algorithm for exact molecular simulations on a quantum computer,” *Nat. Commun.* **10**, 3007 (2019).
 - [22] Mario Motta, Chong Sun, Adrian Teck Keng Tan, Matthew J O’Rourke, Erika Ye, Austin J Minnich, Fernando G S, and Garnet Kin-Lic Chan, “Quantum imaginary time evolution, quantum lanczos, and quantum thermal averaging,” *arXiv:1901.07653* (2019).
 - [23] Robert M Parrish, Edward G Hohenstein, Peter L McMahon, and Todd J Martínez, “Quantum computation of electronic transitions using a variational quantum eigensolver,” *Phys. Rev. Lett.* **122**, 230401 (2019).
 - [24] Per Åke Malmqvist, “Calculation of transition density matrices by nonunitary orbital transformations,” *Int. J. Quantum Chem. Lecture Notes*, **30**, 479–494 (1986).
 - [25] Alex J W Thom and Martin Head-Gordon, “Hartree-Fock solutions as a quasidiabatic basis for nonorthogonal configuration interaction,” *J. Chem. Phys.* **131**, 124113 (2009).
 - [26] Eric J Sundstrom and Martin Head-Gordon, “Non-orthogonal configuration interaction for the calculation of multielectron excited states,” *J. Chem. Phys.* **140**, 114103 (2014).
 - [27] A Yu. Kitaev, “Quantum measurements and the abelian stabilizer problem,” *arXiv:quant-ph/9511026* (1995).
 - [28] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran, “Evaluating analytic gradients on quantum hardware,” *Phys. Rev. A* **99**, 032331 (2019).

- [29] Artur F Izmaylov, Tzu-Ching Yen, Robert A Lang, and Vladyslav Verteletskyi, “Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method,” [arXiv:1907.09040](#) (2019).
- [30] Mateus Araújo, Adrien Feix, Fabio Costa, and Časlav Brukner, “Quantum circuits cannot control unknown operations,” *New J. Phys.* **16**, 093026 (2014).
- [31] Nicholas C Rubin, Ryan Babbush, and Jarrod McClean, “Application of fermionic marginal constraints to hybrid quantum algorithms,” *New J. Phys.* **20**, 053020 (2018).
- [32] Vivek V Shende and Igor L Markov, “On the CNOT-cost of TOFFOLI gates,” [arXiv:0803.2316](#) (2008).
- [33] Attila Szabo and Neil S Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Courier Corporation, 2012).
- [34] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, and Others, “Tensorflow: A system for large-scale machine learning,” in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)* (usenix.org, 2016) pp. 265–283.
- [35] Tamar Stein, Thomas M Henderson, and Gustavo E Scuseria, “Seniority zero pair coupled cluster doubles theory,” *J. Chem. Phys.* **140**, 214113 (2014).
- [36] Hiroshi Nakatsuji, “Structure of the exact wave function,” *J. Chem. Phys.* **113**, 2949–2956 (2000).
- [37] M Nooijen, “Can the eigenstates of a many-body hamiltonian be represented exactly using a general two-body cluster expansion?” *Phys. Rev. Lett.* **84**, 2108–2111 (2000).
- [38] Bryan O’Gorman, William J Huggins, Eleanor G Rieffel, and K Birgitta Whaley, “Generalized swap networks for near-term quantum computing,” [arXiv:1905.05118](#) (2019).
- [39] The Cirq Developers, “Cirq,” (2019).
- [40] Robert M Parrish, Lori A Burns, Daniel G A Smith, Andrew C Simmonett, A Eugene DePrince, 3rd, Edward G Hohenstein, Uğur Bozkaya, Alexander Yu Sokolov, Roberto Di Remigio, Ryan M Richard, Jérôme F Gonthier, Andrew M James, Harley R McAlexander, Ashutosh Kumar, Masaaki Saitow, Xiao Wang, Benjamin P Pritchard, Prakash Verma, Henry F Schaefer, 3rd, Konrad Patkowski, Rollin A King, Edward F Valeev, Francesco A Evangelista, Justin M Turney, T Daniel Crawford, and C David Sherrill, “Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability,” *J. Chem. Theory Comput.* **13**, 3185–3197 (2017).
- [41] Jarrod R McClean, Kevin J Sung, Ian D Kivlichan, Yudong Cao, Chengyu Dai, E Schuyler Fried, Craig Gidney, Brendan Gimby, Pranav Gokhale, Thomas Häner, Tarini Hardikar, Vojtěch Havlíček, Oscar Higgott, Cupjin Huang, Josh Izaac, Zhang Jiang, Xinle Liu, Sam McArdle, Matthew Neeley, Thomas O’Brien, Bryan O’Gorman, Isil Ozfidan, Maxwell D Radin, Jhonathan Romero, Nicholas Rubin, Nicolas P D Sawaya, Kanav Setia, Sukin Sim, Damian S Steiger, Mark Steudtner, Qiming Sun, Wei Sun, Daochen Wang, Fang Zhang, and Ryan Babbush, “OpenFermion: The Electronic Structure Package for Quantum Computers,” [arXiv:1710.07629](#) (2017).
- [42] Eric Jones, Travis Oliphant, and Pearu Peterson, “SciPy: Open source scientific tools for Python,” .
- [43] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal, “Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization,” *ACM Trans. Math. Softw.* **23**, 550–560 (1997).
- [44] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, “Application of Hilbert-space coupled-cluster theory to simple (H₂)₂ model systems: Planar models,” *Phys. Rev. A* **47**, 2738–2782 (1993).
- [45] Uttam Sinha Mahapatra, Barnali Datta, and Debashis Mukherjee, “A size-consistent state-specific multireference coupled cluster theory: Formal developments and molecular applications,” *J. Chem. Phys.* **110**, 6171 (1999).
- [46] Karol Kowalski and Piotr Piecuch, “Complete set of solutions of multireference coupled-cluster equations: The state-universal formalism,” *Phys. Rev. A* **61**, 052506 (2000).
- [47] K. Jankowski and J. Paldus, “Applicability of coupled-pair theories to quasidegenerate electronic states: A model study,” *Int. J. Quantum Chem.* **18**, 1243–1269 (1980).
- [48] Francesco A. Evangelista, Wesley D. Allen, and Henry F. Schaefer, “High-order excitations in state-universal and state-specific multireference coupled cluster theories: Model systems,” *J. Chem. Phys.* **125**, 154113 (2006).
- [49] Joonho Lee, Fionn D Malone, and Miguel A Morales, “An auxiliary-field quantum monte carlo perspective on the ground state of the dense uniform electron gas: An investigation with hartree-fock trial wavefunctions,” *J. Chem. Phys.* **151**, 064122 (2019).
- [50] Hsin-Yuan Huang, Kishor Bharti, and Patrick Rebentrost, “Near-term quantum algorithms for linear systems of equations,” [arXiv:1909.07344](#) (2019).
- [51] Jeng-Da Chai and Martin Head-Gordon, “Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections,” *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).
- [52] Thom H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen,” *J. Chem. Phys.* **90**, 1007–1023 (1989).
- [53] Yihan Shao, Zhengting Gan, Evgeny Epifanovsky, Andrew T.B. Gilbert, Michael Wormit, Joerg Kussmann, Adrian W. Lange, Andrew Behn, Jia Deng, Xintian Feng, Debashree Ghosh, Matthew Goldey, and Paul R. et al. Horn, “Advances in molecular quantum chemistry contained in the Q-Chem 4 program package,” *Mol. Phys.* **113**, 184–215 (2015).

Appendix A: Hexatriene Geometries

The equilibrium geometry was obtained from the geometry optimization with ω B97X-D [51] and cc-pVTZ [52] using a development version of Q-Chem[53]. The 90° twisted configuration was obtained by rotating the middle C-C double bond out-of-plane.

Atom	X	Y	Z
C	(0.5987833, 0.2969975, 0.0000000)		
H	(0.6520887, 1.3822812, 0.0000000)		
C	(-0.5987843, -0.2970141, 0.0000000)		
H	(-0.6520904, -1.3822967, 0.0000000)		
C	(-1.8607210, 0.4195548, 0.0000000)		
H	(-1.8010551, 1.5036080, 0.0000000)		
C	(-3.0531867, -0.1693136, 0.0000000)		
H	(-3.9685470, 0.4053361, 0.0000000)		
H	(-3.1479810, -1.2485605, 0.0000000)		
C	(1.8607264, -0.4195599, 0.0000000)		
H	(1.8010777, -1.5036141, 0.0000000)		
C	(3.0531816, 0.1693296, 0.0000000)		
H	(3.9685551, -0.4052992, 0.0000000)		
H	(3.1479561, 1.2485793, 0.0000000)		

Table I. The geometry of the equilibrium configuration of trans-Hexatriene.

Atom	X	Y	Z
C	(0.5987833, 0.2969975, 0.0000000)		
H	(1.3716346, -0.0683717, 0.6707370)		
C	(-0.5987843, -0.2970141, 0.0000000)		
H	(-1.3716354, 0.0683544, 0.6707361)		
C	(-0.9484080, -1.4197297, -0.8504282)		
H	(-0.1721763, -1.7803215, -1.5183873)		
C	(-2.1390983, -2.0121775, -0.8520831)		
H	(-2.3554088, -2.8468591, -1.5037144)		
H	(-2.9353514, -1.6772360, -0.1982062)		
C	(0.9484189, 1.4197134, -0.8504230)		
H	(0.1721980, 1.7803171, -1.5183881)		
C	(2.1391167, 2.0121462, -0.8520613)		
H	(2.3554502, 2.8468291, -1.5036834)		
H	(2.9353585, 1.6771903, -0.1981764)		

Table II. The geometry of the 90° twisted configuration of Hexatriene.