Nonvariational ADAPT algorithm for quantum simulations

Ho Lun Tang,[1,2] Yanzhu Chen,[1,2,3,*] Prakriti Biswas,[4] Alicia B. Magann,[5] Christian Arenz,[4] and Sophia E. Economou[1,2]

[1]*Department of Physics, Virginia Tech, Blacksburg, Virginia 24061, USA*

[2]*Virginia Tech Center for Quantum Information Science and Engineering, Blacksburg, Virginia 24061, USA*

[3]*Department of Physics, Florida State University, Tallahassee, Florida 32306, USA*

[4]*School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona 85287, USA*

[5]*Quantum Algorithms and Applications Collaboratory, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*

(Received 1 December 2024; accepted 12 May 2025; published 17 June 2025)

We explore a nonvariational quantum state preparation approach combined with the ADAPT operator selection strategy in the application of preparing the ground state of a desired target Hamiltonian. In this algorithm, energy gradient measurements determine both the operators and the gate parameters in the quantum circuit construction. We compare this nonvariational algorithm with ADAPT-VQE and with feedback-based quantum algorithms in terms of the rate of energy reduction, the circuit depth, and the measurement cost in molecular simulation. We find that, despite using deeper circuits, this new algorithm reaches chemical accuracy at a similar measurement cost to ADAPT-VQE. Since it does not rely on a classical optimization subroutine, it may provide robustness against circuit parameter errors due to imperfect control or gate synthesis.

DOI: 10.1103/x8g1-7h1k

I. INTRODUCTION

Preparing the ground state of many-body Hamiltonians and finding the corresponding ground state energy is an important and challenging problem in physics and chemistry. Classical simulation of such systems can require enormous computational resources due to the exponential scaling of the Hilbert space. A potential alternative approach is to simulate the target problem using a quantum processor [1]. Here, the quantum phase estimation algorithm [2,3] is a promising candidate for estimating ground state energies on quantum computers [4,5]. However, to solve problems of a nontrivial size, its implementation requires deep quantum circuits and is expected to need fault-tolerant quantum computers [6]. In the Noisy-Intermediate-Scale-Quantum (NISQ) [6] era, quantum-classical hybrid algorithms provide a promising way to utilize more limited quantum computing resources, enabled by the power of classical computing [7–11]. One family of such algorithms is the variational quantum eigensolver (VQE) [7,12], which targets the problem of preparing the ground state of a desired Hamiltonian. VQEs operate by preparing an ansatz state on a quantum device using a parameterized quantum circuit. Then, measurements of this state allow for estimating the expectation of the Hamiltonian under consideration. A classical computer is then used to variationally find the values of the quantum circuit parameters in order to minimize this Hamiltonian expectation value. Various VQE algorithms have been demonstrated in experiments [7,13–16].

The performance of VQE algorithms relies on the quality of the variational ansatz. An ideal ansatz can explore the Hilbert space containing the ground state with a minimal number of variational parameters and quantum gates. The Adaptive Derivative-Assembled Problem-Tailored Variational Quantum Eigensolver (ADAPT-VQE) [17] was introduced to construct ansätze with this property. Instead of using a fixed ansatz, ADAPT-VQE grows the ansatz iteratively by adding the most effective operators. In various numerical simulations, it has demonstrated promising improvement compared to fixed-circuit VQEs in terms of the number of variational parameters, the circuit depth, and the measurement cost [17–19].

Despite such efforts, the energy landscape for physical systems is typically complicated [20], containing sub-optimal solutions due to the nonconvex nature of the optimization problem, which hinders the classical search for the global optimum. Although there exist many well-developed classical optimization algorithms, the optimization complexity inevitably increases with the system size and the number of variational parameters, and hence becomes a potential bottleneck of VQEs [21]. This bottleneck can manifest as the need for a large number of function evaluations in the classical optimization procedure, which amounts to a large number of quantum circuits to be run. Furthermore, the accuracy of Hamiltonian expectation value estimates, limited by the inherent noise in quantum devices, has a significant impact on the classical optimization. For VQEs to offer practical utility, the classical and quantum resources associated with the variational procedure have to be kept at a reasonable level.

*Contact author: yanzhu.chen@fsu.edu

In this work, we introduce an adaptive quantum algorithm, referred to as the Non-Variational ADAPT (NoVa-ADAPT) algorithm, for preparing the ground state of a target Hamiltonian that does not require classical optimization, Here, the name refers to the fact that within each adaptive step, for the state that takes a certain set of generators in the unitary circuit, we do not vary the parameters in front of those generators. We update the state once rather than carrying out a variational procedure at each step. The overall procedure of finding the final state still follows the variational principle. In the ansatz construction procedure of ADAPT-VQE, the energy derivatives with respect to a set of operators are measured, based on which the optimal operator is selected to be a generator of the ansatz. In our NoVa-ADAPT algorithm, we select operators in the same way but do not optimize over the previous parameters that enter the state. Instead, we estimate the coefficient associated with the selected generator using the measured information from the operator selection process. This allows us to save the measurement cost associated with the classical optimization process, and can potentially reduce the impact of erroneous measurement results. At the same time, selecting operators based on the energy gradient information allows us to save quantum and classical resources.

Previous work has considered the prospect of an optimization-free approach in both classical and quantum algorithms for quantum state preparation. Among them, the Anti-Hermitian Contracted Schrödinger Equation (ACSE) method uses pre-determined operators to iterate the state in a classical algorithm [22]. Another example is feedback-based quantum algorithms (FQAs), such as the Feedback-based ALgorithm for Quantum OptimizatioN (FALQON) [23] and related works [23–33], which utilize a fixed, repeating quantum circuit structure and operate without any classical optimization subroutine. Optimization-free quantum algorithms for ground state preparation have also been developed using the theory of Riemannian gradient flows [34–36]. These latter strategies aim to minimize a cost function by moving in each adaptive step into the direction of the Riemannian gradient. Efficient implementation of this approach on quantum computers can be obtained by approximating the Riemannian gradient through 1- and 2-local Pauli operators [35], or projecting the Riemannian gradient into a random direction [36]. In fact, moving in each step into a random direction provably gives convergence to the ground state (almost surely) despite the existence of sub-optimal solutions [37] (i.e., in the form of saddle points). In the present work, we draw on these previous results and compare our algorithm's performance against this prior art.

In Sec. II we briefly review ADAPT-VQE and its operator selection criterion, before describing the NoVa-ADAPT algorithm in Sec. III. We then review the feedback-based quantum algorithms that are most closely related in Sec. IV. In Sec. V, we demonstrate NoVa-ADAPT with numerical simulations, and compare it to ADAPT-VQE and the previous feedback-based algorithms. We also investigate multiple approaches for estimating the coefficients in the nonvariational approach and compare their performances in terms of the number of operators required in the state construction and the measurement cost. Then in Sec. VI, we explore the impact of the rotational error and the gradient measurement error on the performance, which are expected to be dominant in our algorithm. Finally, we propose a hybrid algorithm of ADAPT-VQE and the non-variational ADAPT method in Sec. VII for practical use in the near future.

II. ADAPT-VQE

ADAPT-VQE was introduced to dynamically construct a variational ansatz according to the problem Hamiltonian. It starts with an initial reference state and a predetermined operator pool $\{A_i\}$ consisting of a set of Hermitian operators A_i. At the beginning of the n-th iteration, if the exponential of the operator A_i is added to the ansatz, the energy expectation value will have the form

$$E_{A_i}(\theta) = \langle\psi^{(n-1)}|e^{-i\theta A_i} H e^{i\theta A_i}|\psi^{(n-1)}\rangle, \tag{1}$$

where $|\psi^{(n-1)}\rangle$ is the estimated ground state from the $(n-1)$-th optimization. ADAPT-VQE measures the first derivative of the energy with respect to the variational parameter associated with each candidate operator A_i in the pool

$$\left.\frac{\partial E_{A_i}}{\partial\theta}\right|_{\theta=0} = i\langle\psi^{(n-1)}|[H, A_i]|\psi^{(n-1)}\rangle. \tag{2}$$

The exponential of the operator $A^{(n)}$ with the largest derivative magnitude will be added to the previous ansatz $|\psi^{(n-1)}\rangle$ and form a new ansatz,

$$|\psi^{(n)}(\theta_n, \ldots, \theta_1)\rangle = e^{i\theta_n A^{(n)}}|\psi^{(n-1)}(\theta_{n-1}, \ldots, \theta_1)\rangle. \tag{3}$$

All the variational parameters $\{\theta_n, \ldots, \theta_1\}$ in the updated ansatz are then varied to minimize the energy $E(\theta_n, \ldots, \theta_1)$ by a classical computer with a chosen optimization algorithm. Starting this optimization from the previous optimized state $|\psi^{(n-1)}\rangle$ corresponds to initializing the newly added parameter θ_n from 0. This iteration is repeated until the norm of the measured gradient is smaller than a user-specified threshold.

The ADAPT-VQE algorithm has been studied extensively in numerical simulations [17–19,38–47]. It has shown promising performance in finding the ground state with high accuracy and leads to much shallower quantum circuits than VQEs with a fixed ansatz. This is because the operator selection procedure adds the locally optimal operator to the ansatz at each iteration, which allows us to explore the Hilbert space efficiently. One can also incorporate symmetries of the Hamiltonian into the operators in the pool, further restricting the Hilbert space for the optimization procedure. Even for such favorable ansätze, classically optimizing the variational parameters remains a challenge in the implementation of the algorithm, due to the large number of measurements required and the errors in the measurement results. In the next section, we investigate the strategy of abandoning the variational procedure and updating the state based on the information obtained in the operator selection procedure.

III. NOVA-ADAPT ALGORITHM

Here, we introduce the NoVa-ADAPT algorithm. Similar to the original ADAPT-VQE algorithm, NoVa-ADAPT starts with a reference state and selects operators from a predetermined operator pool based on the energy derivative. We still denote the state obtained from the $(n-1)$-th iteration

as $|\psi^{(n-1)}\rangle$, although no optimization is involved here. The energy derivatives will be of the same form as in Eq. (2). At the n-th iteration, the operator $A^{(n)}$ with the largest derivative magnitude is used to update the state from $|\psi^{(n-1)}\rangle$ to

$$|\psi^{(n)}\rangle = e^{i\eta_n A^{(n)}}|\psi^{(n-1)}\rangle, \tag{4}$$

where η_n is given by

$$\eta_n = -\gamma \left.\frac{\partial E_{A^{(n)}}}{\partial\theta}\right|_{\theta=0} \tag{5}$$

$$= -\gamma \frac{\partial}{\partial\theta}\langle\psi^{(n-1)}|e^{-i\theta A^{(n)}} H e^{i\theta A^{(n)}}|\psi^{(n-1)}|_{\theta=0}. \tag{6}$$

The factor γ is a real number controlling the magnitude of the update. Although it may look similar to the ADAPT-VQE ansatz, here $|\psi^{(n)}\rangle$ is *not* a variational ansatz since η_n has a fixed value. We now denote the energy expectation value obtained after the n-th iteration as $E^{(n)} \equiv \langle\psi^{(n)}|H|\psi^{(n)}\rangle$ for convenience.

As the gradient is evaluated at zero parameter value, a sufficiently small γ can ensure that the energy is lowered, i.e., $E^{(n)} - E^{(n-1)} \leqslant 0$. More specifically, based on a descent lemma for first-order optimization algorithms [36,48], it was shown that the energy reduction can be lower bounded by

$$E^{(n-1)} - E^{(n)} \geqslant \frac{1}{8\|H\|_2\|A^{(n)}\|_2^2}\left(\left.\frac{\partial E_{A^{(n)}}}{\partial\theta}\right|_{\theta=0}\right)^2 \tag{7}$$

by choosing

$$\gamma_n = \frac{1}{4\|H\|_2\|A^{(n)}\|_2^2}, \tag{8}$$

where $\|M\|_2$ is the spectral norm of operator M and $A^{(n)}$ is the operator added to the state at this iteration [36]. Although this choice ensures a nonzero energy reduction given a nonzero energy derivative, it limits the effect of each operator added to the state, thus increasing the number of operators required to reach the ground state.

To estimate the optimal value of γ to minimize the energy, we expand the energy reduction to the second order,

$$E^{(n)} - E^{(n-1)} \approx \eta_n \left.\frac{\partial E^{(n)}}{\partial\eta_n}\right|_{\eta_n=0} + \frac{\eta_n^2}{2}\left.\frac{\partial^2 E^{(n)}}{\partial\eta_n^2}\right|_{\eta_n=0}, \tag{9}$$

and find the stationary point,

$$\gamma^* = -\left(\left.\frac{\partial^2 E^{(n)}}{\partial\eta_n^2}\right|_{\eta_n=0}\right)^{-1}. \tag{10}$$

In practice, this requires additional measurements for the second derivative according to

$$\left.\frac{\partial^2 E^{(n)}}{\partial\eta_n^2}\right|_{\eta_n=0} = -\langle\psi^{(n-1)}|[A^{(n)},[A^{(n)},H]]|\psi^{(n-1)}\rangle. \tag{11}$$

At each iteration we only need to evaluate it for the selected operator $A^{(n)}$ so the cost does not grow with the size of the operator pool. Choosing γ^* is similar to applying Newton's method in optimizing one parameter, which leads to much faster update than the choice in Eq. (8) but there is no guarantee of energy reduction.

Alternatively, we can set γ to a constant value and treat it as a hyperparameter. This is similar to the learning rate in gradient descent optimization methods. We would like a high value to speed up convergence to a local minimum, but in general the optimization landscape is nonconvex and if γ is too high, the energy may oscillate or diverge. Between γ^* and a constant γ, it is unclear which one produces better results, since neither guarantees convergence. When they both lead to local minima, they may lead to different ones. Our state update rule is inspired by gradient descent optimization methods, which also face this issue. A line search may be adopted to avoid energy divergence in our algorithm at the cost of more measurements. We leave it to future work.

The nonvariational approach saves the sampling cost associated with the optimization process. On one hand, this can potentially reduce the number of measurement samples required by the algorithm as the measurement cost in ADAPT-VQE is dominated by the optimization process. On the other hand, to generate each measurement sample, each quantum circuit evaluation is subject to noise, and eliminating the optimization based on noisy measurements may reduce the impact of errors. We remark that the noise may have different effects on the operator selection process and the optimization process of ADAPT-VQE, which requires further studies in the future.

IV. GRADIENT-BASED FEEDBACK

The strategy of preparing the ground state based on the gradient estimate in an optimization-free way serves as the basis of FQAs [23,33] and the randomized adaptive quantum state preparation algorithm [36]. While the NoVa-ADAPT algorithm updates the parameter in the same way, it differs from the previous algorithms in the operator selection. FQAs were inspired by quantum control protocols and use the terms in the Hamiltonian as generators for each iteration in the circuit, whereas the randomized adaptive algorithm gains robustness from random operators as generators. A classical algorithm for calculating reduced density matrices in chemistry adopted a similar way of iterating the state, with a linear combination of fermionic operators as the generator [22]. A quantum-classical hybrid version of this algorithm turns the transformation on the state into a quantum circuit with Trotterization [49]. We briefly review these closely related approaches and compare the simulated performances in order to study the importance of operators.

A. Feedback-based quantum algorithms

The FQA procedure for preparing molecular ground states starts with separating the Hamiltonian into two parts,

$$H = H_1 + H_2, \tag{12}$$

where the first part

$$H_1 = \sum_{pq} h_{pq} a_p^\dagger a_q, \tag{13}$$

collects the single-body operators, and the second part

$$H_2 = \frac{1}{2}\sum_{pqrs} a_p^\dagger a_q^\dagger a_r a_s. \tag{14}$$

includes all the two-body operators in the Hamiltonian. The FQA circuit of n layers can be written as

$$U_{\mathrm{FQA}}^{(n)}(\vec{\beta}) = \tilde{U}_1(\beta_n \Delta t)\tilde{U}(\Delta t)\cdots\tilde{U}_1(\beta_1 \Delta t)\tilde{U}(\Delta t), \quad (15)$$

where $\tilde{U}(\Delta t)$ and $\tilde{U}_1(\beta_k \Delta t)$ are the Trotterized versions of $e^{-iH\Delta t}$ and $e^{-i\beta_k H_1 \Delta t}$, respectively. The parameter β_n is evaluated with the state obtained in the last iteration,

$$\beta_n = -i\langle\psi^{(n-1)}|[H_1, H_2]|\psi^{(n-1)}\rangle, \quad (16)$$

where $|\psi^{(n-1)}\rangle = U_{\mathrm{FQA}}^{(n-1)}(\vec{\beta})|\psi^{(0)}\rangle$. Conventionally, $|\psi^{(0)}\rangle$ is taken to be the ground state of H_1 in FQAs. In the material below, we deviate from this convention. We also remark that the time step Δt has to be chosen separately, which plays the same role as the γ parameter in Eq. (6). One option is to use the lower bounding value introduced in Eq. (8), substituting H_1 for $A^{(n)}$.

B. Randomized adaptive quantum algorithm

The randomized adaptive quantum algorithm [36] dynamically constructs the quantum circuit through randomly selecting an operator in each adaptive step, thereby achieving convergence to the ground state almost surely [37]. We consider here randomization by constructing the operators $\{A^{(n)}\}$ in two different ways:

Sampling a random operator from a pool of operators and measuring the gradient of the cost function. The sampled operator is added to the state as long as it gives a nonzero gradient.

Conjugating a fixed traceless operator A by a random unitary transformation V_n as

$$A^{(n)} = V_n^\dagger A V_n. \quad (17)$$

The random unitary transformation V_n can be sampled from a unitary 2-design that can be efficiently approximated by the construction in Ref. [50]. It has been shown that sampling from a 2-design, instead of the full Haar distribution, achieves convergence almost surely [37]. This is independent of the set of unitaries forming the 2-design.

C. Anti-Hermitian Contracted Schrödinger Equation

The ACSE method [22] approaches the ground state of a molecular Hamiltonian H iteratively, acting on the previous state $|\psi\rangle$ with $e^{\epsilon S}$, where ϵ is an infinitesimal step size and the anti-Hermitian operator S contains up to two-body fermionic operators

$$S = \sum_{p,r} {}^1S_r^p a_p^\dagger a_r + \sum_{p,q,r,s} {}^2S_{r,s}^{p,q} a_p^\dagger a_q^\dagger a_r a_s. \quad (18)$$

The matrix elements

$$\begin{aligned} {}^1S_r^p &= \langle\psi|[a_p^\dagger a_r, H]|\psi\rangle, \\ {}^2S_{r,s}^{p,q} &= \langle\psi|[a_p^\dagger a_q^\dagger a_r a_s, H]|\psi\rangle \end{aligned} \quad (19)$$

are selected to reduce the energy expectation value using the Euler's method. By Trotterizing $e^{\epsilon S}$, one can compile it as a quantum circuit and turn this classical algorithm into a quantum-classical hybrid algorithm, the quantum contracted-eigenvalue-equation solver [49,51]. This may produce a similar circuit structure to the NoVa-ADAPT algorithm with a pool containing 1-body and 2-body fermionic operators, as the matrix elements of S are the same as the energy gradients (up to a factor of i) and ϵ plays the same role as γ in the NoVa-ADAPT algorithm. However, the states constructed by the two algorithms are very different. In this algorithm, a sequence of exponentials $\prod_k e^{i\alpha_k O_k}$ of all 1-body and 2-body fermionic operators $\{O_k\}$ is appended to the state at each iteration, with $\{\alpha_k\}$ calculated for this iteration. The effect depends on the choice of ordering in the Trotterization. In contrast, the NoVa-ADAPT approach only selects the one term with the largest magnitude of the coefficient. The difference between the circuit structures is analogous to the difference between the unitary coupled-cluster single and double (UCCSD) ansatz [7] and the ADAPT-VQE ansatz [17]. Moreover, even if the selected operators coincide with the ordered sequence $\{O_k\}$, the update parameters are calculated from many iterations of the state and will be different from $\{\alpha_k\}$. Because each iteration of the quantum contracted-eigenvalue-equation solver involves a sequence of terms, we expect the NoVa-ADAPT algorithm to lead to more compact quantum circuits.

V. RESULTS

We simulate the performance of the proposed algorithm for finding the ground state of molecular Hamiltonians, with the Hartree-Fock state as the reference state. The Hartree-Fock state is a product state calculated with a mean-field theory approach to approximate the ground state. In our simulations, we take the linear H_4 molecule in the STO-3G basis with a bond distance of 1.5 Å as our main example. The Hamiltonian consists of fermionic creation and annihilation operators,

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2}\sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \quad (20)$$

where $\{p, q, r, s\}$ index the 8 spin-orbitals we consider. For ADAPT algorithms, we start from the restricted Hartree-Fock state and use an operator pool containing the spin-adapted fermionic operators, which preserve the number of electrons, spin polarization S_z, and total spin S^2 [17]. They consist of single excitation operators:

$$\tau_1 \propto |\uparrow\rangle_i\langle\uparrow|_j + |\downarrow\rangle_i\langle\downarrow|_j - \mathrm{H.c.}, \quad (21)$$

and double excitation operators:

$$\begin{aligned} \tau_{2,T} &\propto |T, 1\rangle_{ij}\langle T, 1|_{kl} + |T, -1\rangle_{ij}\langle T, -1|_{kl} \\ &\quad + |T, 0\rangle_{ij}\langle T, 0|_{kl} - \mathrm{H.c.} \\ \tau_{2,S} &\propto |S, 0\rangle_{ij}\langle S, 0|_{kl} - \mathrm{H.c.}, \end{aligned} \quad (22)$$

where $\{i, j, k, l\}$ are spatial orbitals and T and S refer to triplet and singlet states formed by (i, j) or (k, l). We write these operators in terms of wavefunctions on 1 or 2 particles to explicitly show that S^2 and S_z are conserved. They can also be expressed in terms of annihilation and creation operators.

In our simulations, we map these operators to Pauli operators using the Jordan-Wigner transformation, where 8 qubits are required. The expression of the Hamiltonian in terms of Pauli operators is given in Appendix B. Using the pool of spin-adapted fermionic operators, as both the Hamiltonian

and the state are represented using fermionic operators, the algorithm and our simulation results are independent of the qubit encoding.

We use the energy error, which is the difference between the energy produced by the algorithms and the energy from exact diagonalization, to measure the quality of the ground state. For resource costs, we consider the number of operators added to the circuit and the measurement budget. The number of operators is chosen because it directly affects the circuit complexity. Although 2-body fermionic operators generally lead to more complex 1-body fermionic operators, we do not examine the circuit depth or the number of entangling gates as the NoVa-ADAPT algorithm and ADAPT-VQE are not observed to choose predominantly one type of operators. Since measuring an observable in practice consists of measuring the Pauli terms in the observable, which generally do not all commute, the number of measurements required depends on how to group these Pauli terms into commuting sets [52]. Therefore, we use the number of function evaluations in each algorithm as a proxy for the measurement budget, where estimating either the energy or the gradient counts as one function evaluation. This allows us to examine the algorithmic performance without the effect of different strategies of grouping Pauli terms. The classical optimization in ADAPT-VQE simulations are carried out with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm and the Hessian approximated by BFGS is recycled for the next ADAPT iteration; this Hessian recycling scheme introduced in [45] can reduce the measurement cost in ADAPT-VQE.

A. Comparison to ADAPT-VQE and random operators

Figure 1 compares the NoVa-ADAPT algorithm with ADAPT-VQE and the randomized adaptive algorithm. For values of γ, we take three possibilities: the lower bounding value provided in Eq. (8) (magenta), γ^* as in Eq. (10) (red), and a constant value (blue). The constant $\gamma = 1$ is selected from a range of values to provide relatively fast energy reduction without oscillatory behavior. We include the simulated performances with different contant γ's in Appendix A. Note that for the randomized adaptive algorithm (green), we take the average of energy errors in 100 runs to show the performance in the upper panel, since different operators are selected at each run. In the lower panel, however, we show the 100 runs as individual curves because the numbers of function evaluations across different runs are generally different. At each iteration the algorithm randomly selects an operator from the pool, which may produce a zero-energy gradient. In this case, another operator is randomly selected until the gradient is nonzero. Consequently, for a given iteration, each run is associated with a different number of function evaluations and a different energy error.

The randomized adaptive algorithm (green), where the state is constructed with operators randomly selected from the pool, needs the largest number of operators to reach a given energy accuracy. This is expected since the operators are not selected based on the energy gradient information. In terms of the measurement cost, it performs better than the NoVa-ADAPT algorithm with the lower-bounding γ (magenta),

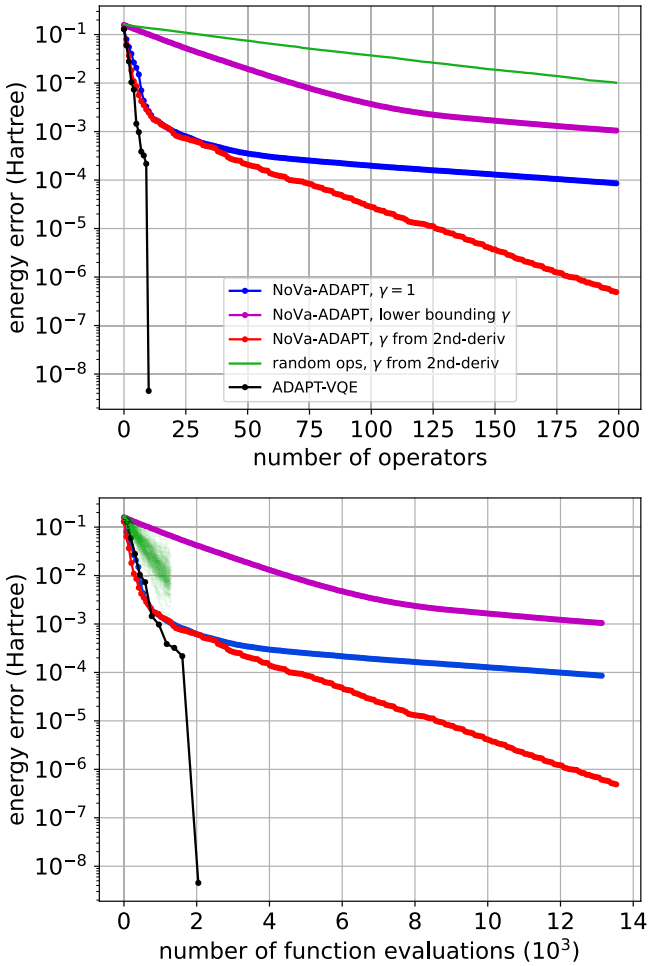

FIG. 1. Results for H_4 molecule of bond length 1.5 Å with the spin-adapted fermionic operator pool [17]. Energy error is plotted as a function of the number of operators added to the state (upper panel), and as a function of the number of function evaluations needed (lower panel). Blue curves show the results of the NoVa-ADAPT algorithm with constant $\gamma = 1$. Green curves show the results of constructing the state with operators randomly picked from the pool with γ calculated from the second derivative using Eq. (10), sampled from 100 runs. Magenta curves show the results of the NoVa-ADAPT algorithm with lower bounding γ from Eq. (8). Red curves show the results of the NoVa-ADAPT algorithm with γ calculated from the second derivative using Eq. (10). Black curves show the results of ADAPT-VQE, where classical optimization is implemented.

which provides the lowest update rate for the nonvariational state construction. The NoVa-ADAPT algorithm exhibits similar behaviors at early stages with constant $\gamma = 1$ (blue) and with $\gamma = \gamma^*$, as in Eqs. (10) and (11) (red). Eventually, estimating the value of γ based on the second derivative leads to a better energy accuracy, as this provides more information about the optimization landscape. The ADAPT-VQE simulation finds the ground state with 11 operators, which is much lower than the NoVa-ADAPT algorithm. However, if we compare the number of function evaluations required in the algorithm [Fig. 1(b)], the ADAPT-VQE performance is comparable to the NoVa-ADAPT algorithm with

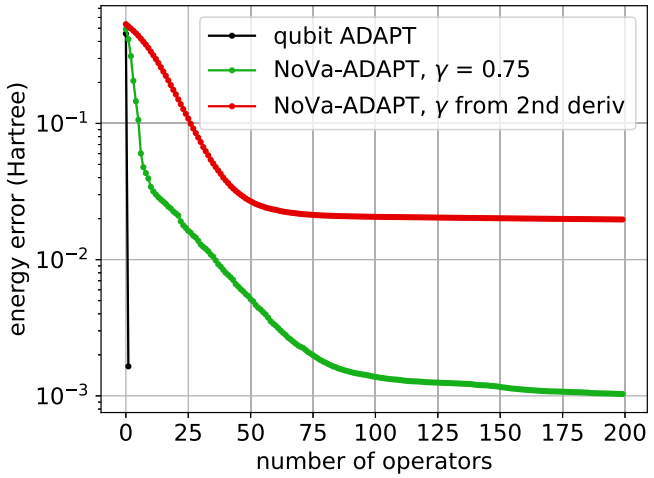

FIG. 2. Results for H_4 molecule of bond length 3 Å with the qubit operator pool [18]. Energy error as a function of the number of operators added to the state. The black curve shows the result of qubit-ADAPT-VQE, the green curve shows the result of the NoVa-ADAPT algorithm with constant $\gamma = 0.75$, and the red curve shows the result of the NoVa-ADAPT algorithm with γ calculated with the second derivative of energy using Eq. (10).

constant $\gamma = 1$ and with $\gamma = \gamma^*$ until both algorithms reach chemical accuracy (≈0.0016 Hartree). In Appendix C, we include a comparison between the NoVa-ADAPT algorithm and ADAPT-VQE for the LiH moledule of bond length 1.5 Å, which is slightly larger than the H_4 molecule and contains 12 spin orbitals. We observe qualitatively similar performances there, although more comprehensive studies are required for the scaling behavior.

On the other hand, ADAPT-VQE can terminate prematurely after converging to a local minimum in the energy landscape, in which case the NoVa-ADAPT algorithm may provide an alternative approach to avoid this situation. We pick one such case in Fig. 2, the H_4 molecule with bond distance 3 Å which is a strongly correlated Hamiltonian as the electrons are highly nonlocal. Using the qubit operator pool [18], which consists of the simplified individual Pauli operators in the Jordan-Wigner mapping of one-body and two-body fermionic operators, ADAPT-VQE finds a local minimum after adding the second operator. For this Hamiltonian, the NoVa-ADAPT algorithm with γ based on the second derivative also reaches a region where the energy reduction slows down exponentially. With a constant $\gamma = 0.75$, however, the NoVa-ADAPT algorithm can lower the energy down to chemical accuracy, beyond the point where ADAPT-VQE was trapped. In this particular problem, the fact that ADAPT-VQE quickly converges to a sub-optimal trap suggests that the reference state is close to the trap. The NoVa-ADAPT algorithm with γ based on the second derivative likely converges to another sub-optimal trap. With a constant γ, the NoVa-ADAPT algorithm can overshoot past these two local minima, allowing it to avoid getting trapped there, despite converging more slowly. We remark that this is problem-dependent. In general it is unclear whether γ based on the second derivative or a constant value will lead to a lower minimum. For the rest of this work, the pool of spin-adapted fermionic operators is used in the NoVa-ADAPT algorithm and ADAPT-VQE.

B. Random initial states

The performance of classical optimization subroutine can strongly depend on the initial parameter values. For some Hamiltonians, the classically calculated reference state has a very small overlap with the ground state, which may hinder VQE's performance, requiring many more iterations to achieve convergence. In this case, state preparation with randomly selected operators may be favorable as the initial state cannot introduce bias in this case. Here, we compare ADAPT-VQE, the NoVa-ADAPT algorithm, and the randomized adaptive algorithm, as reviewed in Sec. IV B on the same H_4 molecule with a bond distance of 1.5 Å and a spin-adapted fermionic operator pool, but from randomly sampled initial states. In the problem of finding the molecular ground state and energy, the Hartree-Fock state usually serves as a good reference state with the correct symmetries. We use the simulation with randomly sampled initial states to imitate problems where no such good reference states are known and probe the performance of the algorithms.

The random initial states are obtained by sampling from a unitary V from a 2-design and applying V to the Hartree-Fock state. We adopt the method in Ref. [50] and approximate the 2-design using random unitaries diagonal in the Pauli bases. For all the methods except the original ADAPT-VQE, the parameters are estimated using the second derivative with Eq. (10). We remark that in the lower panel of Fig. 3 we show individual curves corresponding to different initial states for both ADAPT-VQE (black) and the algorithm with operators randomly selected from the pool (green). This is because the numbers of function evaluations across runs differ from each other in these two algorithms. For the algorithm randomly selecting operators, the reason is explained in Sec. V A. In contrast, the random adaptive algorithm that constructs generators with 2-design (blue), the selected generators generally do not lead to a zero energy gradient so the number of function evaluations at each step is the same across runs. For ADAPT-VQE, the selected operator for each run may be different, so the number of function evaluations in the optimization subroutine can be drastically different.

From Fig. 3, we see that none of the methods can reach the ground state within a reasonable number of function evaluations. For ADAPT-VQE (black), even if the simulation stops at the 30th iteration, the number of function evaluations has already exceeded that for the NoVa-ADAPT algorithm (red), which contains 500 operators. Using randomized operators as generators to update the state (blue and green) lowers the energy by a smaller amount for a given number of operators, but the number of measurements associated with a given energy reduction for some initial states can be lower than that for ADAPT-VQE and the NoVa-ADAPT algorithm. In this case, the random initial state breaks the symmetries in the problem, whereas the operators in the ADAPT pool preserve the symmetries, which are therefore inefficient in approaching the symmetric ground state.

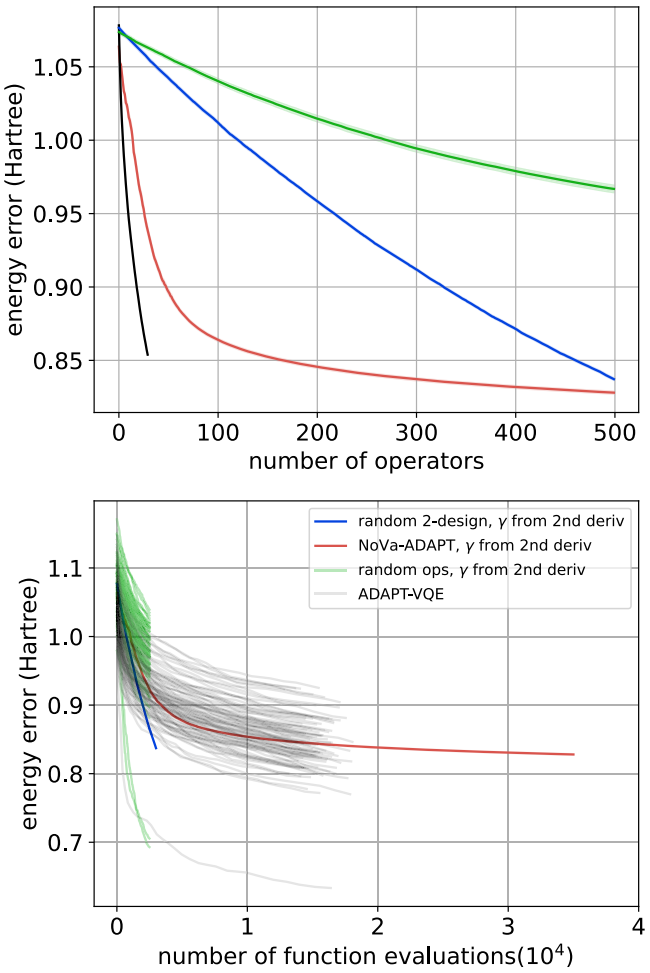

FIG. 3. Results for H_4 molecule of bond length 1.5 Å with 100 random initial states, obtained by ADAPT-VQE (black), the NoVa-ADAPT algorithm (red), randomly sampling operators from the pool (green), and the 2-design randomized adaptive method (blue). In the latter three nonvariational approaches, γ is computed with the second derivative using Eq. (10). (Upper panel) Energy error is plotted as a function of the number of operators added to the state, averaged over random initial states. (Lower panel) Energy is plotted as a function of the number of function evaluations. Black curves and green curves show the results of 100 random initial states. Green curves include the measurements associated with the operators rejected by the algorithm (i.e., with zero energy derivative). The red curve shows the average over 100 random initial states. The blue curve shows the average of 100 samples for different initial states and different random 2-design circuits.

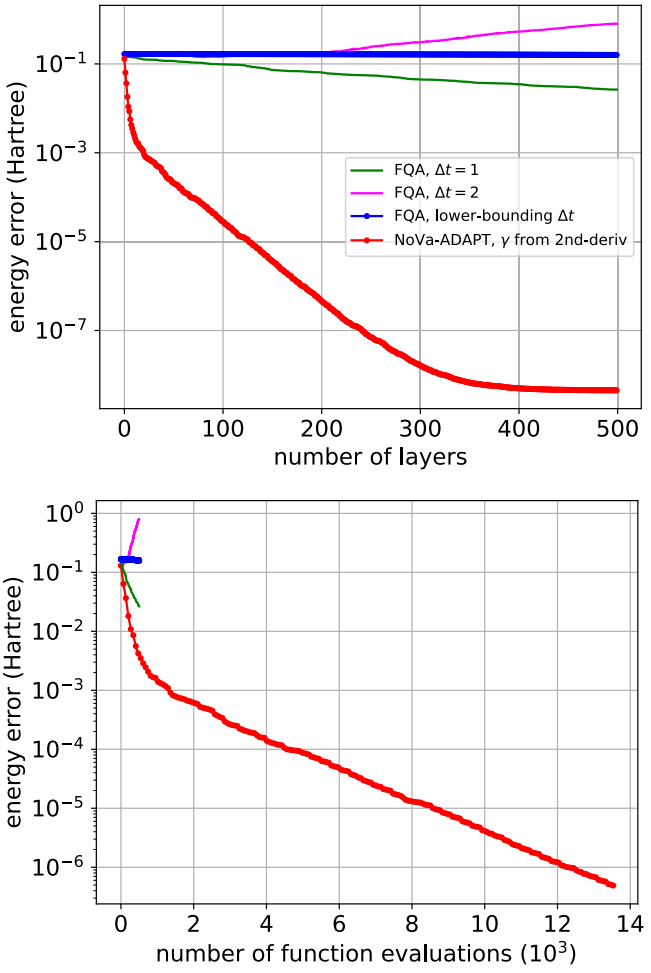

FIG. 4. Results for H_4 molecule of bond length 1.5 Å using the NoVa-ADAPT algorithm and FQA. Energy error as a function of the number of iterations/layers in the state (upper panel) and as a function of the number of function evaluations needed (lower panel). In each iteration/layer, the NoVa-ADAPT algorithm adds one operator, while FQA adds two unitaries containing multiple operators. The red curve shows the result of the NoVa-ADAPT algorithm with γ computed with second derivative using Eq. (10). The blue, green, and pink curves show the results of FQA with the lower-bounding Δt, $\Delta t = 1$, and $\Delta t = 2$, respectively. For FQA, each layer in the state consists of two unitaries, i.e., $\tilde{U}_1(\beta_k \Delta t)$ and $\tilde{U}(\Delta t)$.

C. Comparison to FQA and the quantum ACSE method

In Fig. 4, we see the NoVa-ADAPT algorithm with γ computed from the 2nd-derivative performs substantially better than an FQA with various values of Δt in terms of the number of iterations. It is worth noting that in the FQA, each layer consists of two unitaries, $\tilde{U}_1(\beta_k \Delta t)$ and $\tilde{U}(\Delta t)$. Each unitary is a product of a series of terms, since both H_1 and H contain multiple single-body or two-body operators. Therefore, even with the same number of iterations, the circuit depth of FQA exceeds that of the NoVa-ADAPT algorithm. For the FQA, the energy reduction rate increases with Δt for $\Delta t = 1$ and below, with $\Delta t = 2$ falling outside the range of energy reduction.

In Fig. 5, we compare the NoVa-ADAPT algorithm and the quantum ACSE method, where all the operators in the pool of the former appear in one iteration of the latter. In the quantum ACSE method, the value of ϵ is optimized at each iteration and the optimized value is fixed for subsequent iterations [49]. We use BFGS for this optimization, although in principle any optimizer can be used. Different choices of ordering can be made for the operators in each Trotter step. We show the results of an ordering following an arbitrary labeling of the operators (green), and of an ordering according to the magnitudes of the energy gradients, from the largest to the smallest (blue). The two ordering lead to similar performances in terms of

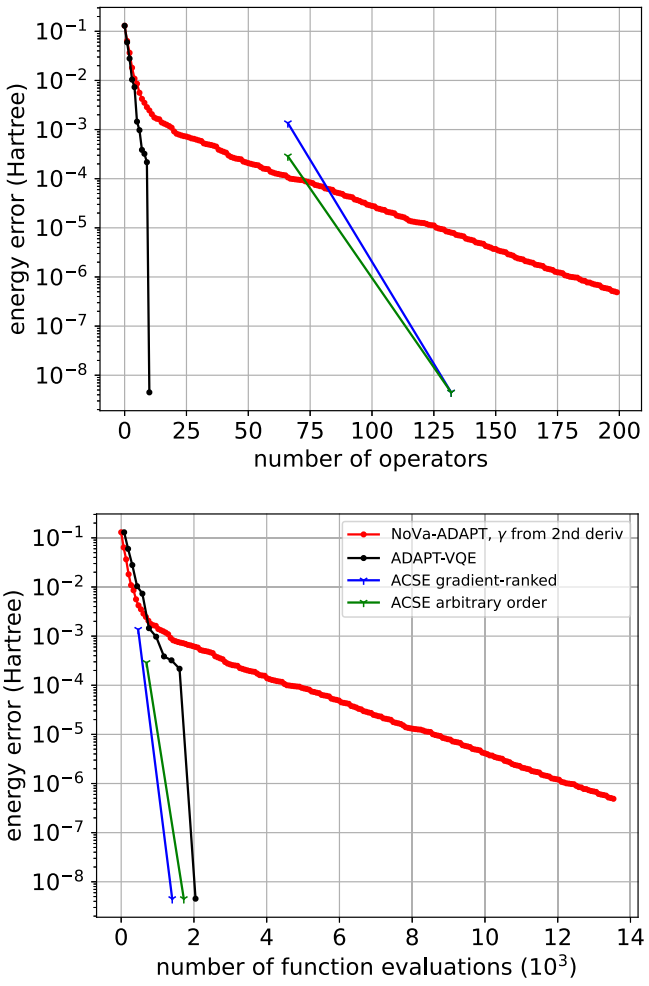

FIG. 5. Results for H_4 molecule of bond length 1.5 Å using the NoVa-ADAPT algorithm and the quantum ACSE method. Energy error is plotted as a function of the number of operators added to the state (upper panel) and as a function of the number of function evaluations needed (lower panel). The red curve shows the result of the NoVa-ADAPT algorithm with γ computed with second derivative using Eq. (10). The green and blue curves show the results of the quantum ACSE method with an arbitrary Trotter ordering and a Trotter ordering according to the energy gradient ranking, respectively.

the energy accuracy and the measurement budget. Since the NoVa-ADAPT algorithm adds only the most relevant operator at each iteration, it reaches chemical accuracy with a smaller number of operators than the quantum ACSE method, leading to more compact quantum circuits. The quantum ACSE method is more economical in the number of measurements required to reach the chemical accuracy, as all the operators in the pool are added in each iteration and therefore fewer iterations are needed. For the small-scale problem studied here, the quantum ACSE method gets below chemical accuracy after two iterations.

From the comparison to FQA and the quantum ACSE method, we show that selecting the most energy-relevant operators improves the circuit depth of the feedback-based state preparation, starting from the Hartree-Fock state.

VI. NOISE ROBUSTNESS

In practice, quantum devices are subject to errors due to imperfect control as well as interactions with the environment. To study the algorithmic performance, we directly model the impact of errors on the algorithms studied in this work, instead of modeling physical errors for a given platform. Possible sources of such effective errors are given following each type of error. In both the NoVa-ADAPT algorithm and ADAPT-VQE, the quantum circuit used for state preparation is determined by the selected operators and the variational or update parameters. Therefore, we model the dominant impact of errors in two ways: errors in the variational or update parameters, which we refer to as rotational errors, and errors in the gradient measurements, which affect operator selection. In the near term, without the protection of error correction, the rotational error stems mainly from the imprecise control of gate parameters. Even in the fault-tolerant regime, the implementation of the algorithm is not completely free of the rotational error. At the logical level, the actual gate parameter can still deviate from the desired value due to fault-tolerant gate synthesis with a finite overhead [53–55]. Since the update parameter does not have to be a precise value for the energy to decrease, we anticipate the NoVa-ADAPT algorithm to be more robust to the rotational error than ADAPT-VQE. On the other hand, both the variational and the nonvariational algorithms in this work select operators based on gradient estimates, which suffer from statistical uncertainty as well as decoherence. The nonvariational approach also relies on these estimates to provide update parameters. Consequently, gradient measurement errors may have a larger impact on its performance. To investigate the robustness to these two error types, we perform simulations with modeled effective errors. The largest number of operators selected by a given algorithm is set to be 200, beyond which the simulation is terminated by hand. In the lower panels of Figs. 6 and 7, different circuit realizations are shown as individual curves for both ADAPT-VQE (blue) and the algorithm with operators randomly selected from the pool (green), for the reasons explained in Secs. V A and V B.

The rotational error is simulated by adding a random value to the optimized parameters at the end of each ADAPT iteration,

$$\vec{\theta} \to \vec{\theta} + \vec{\delta\theta}, \tag{23}$$

where $\vec{\delta\theta}$ is a vector of random values sampled from a Gaussian distribution with zero mean and standard deviation $\sigma = 0.01$ or 0.001. It imitates the effect of the over/under-rotations occurring in the circuit preparation. This random deviation is not added for each iteration in the classical optimization part of ADAPT-VQE to keep the simulation fast. This leads to an underestimate of the error impact on ADAPT-VQE. From Fig. 6, we see that the NoVa-ADAPT algorithm suffers moderately from the rotational error, while ADAPT-VQE shows a high sensitivity to the rotational error where the energy starts to increase after a certain number of iterations. This is because the parameters in ADAPT-VQE are first tightly optimized to an optimal point, and the energy is sensitive to any perturbation from the optimal point. Since ADAPT-VQE goes through the VQE subroutine at every adaptive step,

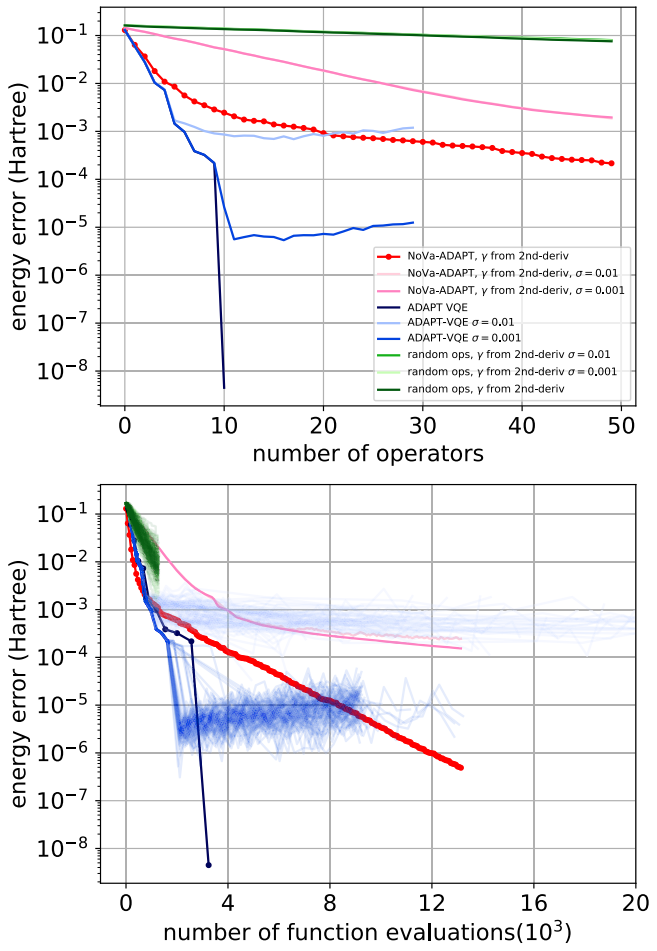

FIG. 6. Results for H_4 molecule of bond length 1.5 Å with rotational errors. Energy error is plotted as a function of the number of operators added to the state (upper panel), and as a function of the number of function evaluations (lower panel). Blue curves show the results of ADAPT-VQE. Red curves show the results of the NoVa-ADAPT algorithm with γ computed from the second derivative using Eq. (10). In the upper panel, the two red curves representing different nonzero error strengths coincide. Green curves show the results of randomly picked operators from the pool with γ computed from the second derivative. In the upper panel, all three green curves coincide. All results shown here are based on 100 circuit realizations, either as individual realizations or as the average.

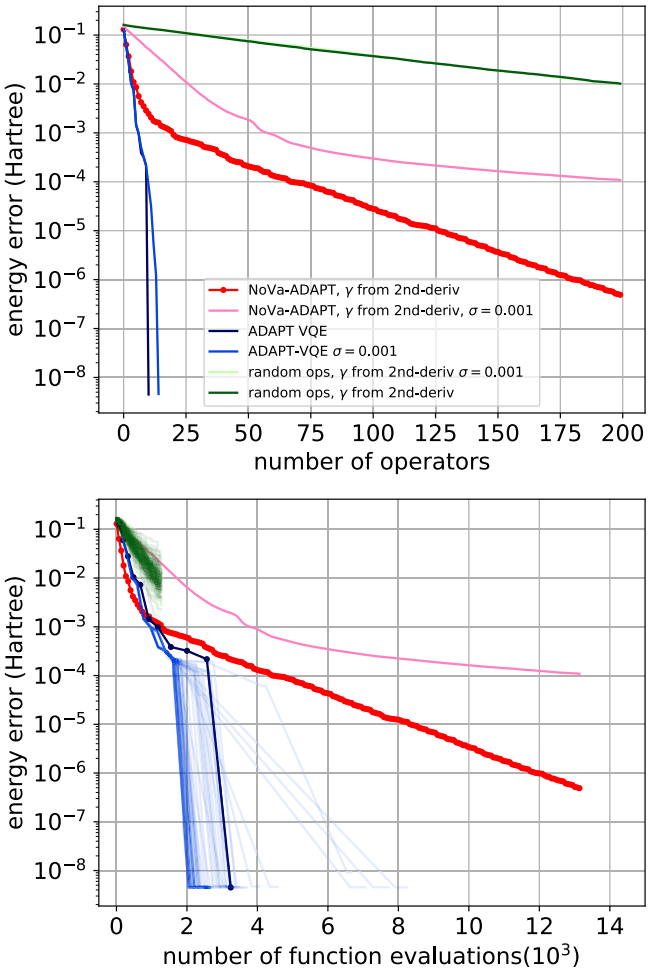

FIG. 7. Results for H_4 molecule of bond length 1.5 Å with gradient measurement errors. Energy error is plotted as a function of the number of operators added to the state (upper panel), and as a function of the number of function evaluations (lower panel). Blue curves show the results of ADAPT-VQE. Red curves show the results of the NoVa-ADAPT algorithm with γ computed from the second derivative using Eq. (10). Green curves show the results of randomly picked operators from the pool with γ computed from the second derivative. In the upper panel, both green curves coincide. All results shown here are based on 100 circuit realizations, either as individual realizations or as the average.

each from the optimized parameters from the last step, the rotational error leads to a worse initial point for the next round of classical optimization. In contrast, perturbations in the update parameters have less impact in the nonvariational approach since a range of parameter values can lead to energy reduction. The number of function evaluations required by noisy simulations of ADAPT-VQE to reach the same energy accuracy is thus larger than the NoVa-ADAPT algorithm. Finally, the randomized adaptive method is almost unchanged under the same level of rotation error, because the energy error it can attain is already larger than the noise level in the simulation.

The gradient measurement error is inserted into the simulation by adding a random value to the calculated first derivative of the pool operators during the operator selection procedure, as follows:

$$\left.\frac{\partial E_{A_i}}{\partial\theta}\right|_{\theta=0} \rightarrow \left.\frac{\partial E_{A_i}}{\partial\theta}\right|_{\theta=0} + \delta_i, \tag{24}$$

where δ_i is a random value sampled from Gaussian distribution with zero mean and standard deviation σ. This form of error captures the effect of statistical uncertainty associated with repeated measurements. It will affect the choice of operators for ADAPT-VQE and the NoVa-ADAPT algorithm. Since the NoVa-ADAPT algorithm and the randomized algorithm also use the measured gradient to estimate the update parameter, they are affected by the gradient error in two ways. From Fig. 7, we see that the NoVa-ADAPT algorithm is affected moderately by the gradient error, while ADAPT-VQE is robust against the gradient error, since a less-than-optimal

choice of operator can still produce a decent variational ansatz. Strategies of operator selection based on perturbation theory have been designed to reduce or avoid the impact of measurement noise on variational ansatz construction [56,57]. We remark that gradient errors are not introduced in the gradient-based classical optimization part of ADAPT-VQE, so the energy error in the noisy ADAPT-VQE simulation is underestimated. Meanwhile, the randomized adaptive method is also unchanged under the same level of gradient error.

VII. ADAPT-VQE—NOVA-ADAPT HYBRID APPROACH

While the NoVa-ADAPT algorithm can potentially lower the measurement cost, it adds more operators to the state and leads to a deeper quantum circuit. To compensate for this drawback on the circuit depth and leverage the desired features of each algorithm, we explore the strategy of combining ADAPT-VQE and the NoVa-ADAPT algorithm. In Fig. 1(b), we observe that the NoVa-ADAPT algorithm performs better at the early stage of the algorithm, offering an energy reduction comparable to that of ADAPT-VQE. Based on this observation, we start with the NoVa-ADAPT algorithm and then switch to ADAPT-VQE after a certain number of iterations. In Fig. 8(b), we see that for the H_4 Hamiltonian with a bond distance of 1.5 Å, switching at the 5th iteration (blue) can slightly reduce the number of function evaluations required to reach chemical accuracy. This simulation demonstrates that the hybrid approach can perform similarly to both algorithms. In terms of potential advantages, a hybrid approach starting from the nonvariational approach can be seen as a pre-optimization technique that can lead to different initial states for ADAPT-VQE. Applied to problems where ADAPT-VQE converges to a local minimum prematurely, it may produce similar effects as the basin-hopping approach, from which the lowest can be selected. Further work is required to study whether this hybrid approach can save quantum resources for larger-scale problems.

VIII. CONCLUSIONS

We introduce the NoVa-ADAPT algorithm for preparing ground states, which both selects the operators and estimates the coefficients for updating the state based on the energy gradient information, without performing optimization. Through numerical simulation, we show that the NoVa-ADAPT algorithm performs comparably to ADAPT-VQE for a small molecular Hamiltonian in terms of measurement cost at the early stage of the algorithm, but has more operators in the circuit. By getting rid of the classical optimization, the NoVa-ADAPT algorithm may provide more robustness against errors in circuit parameters due to imperfect control or gate synthesis. It appears more prone to gradient measurement errors than ADAPT-VQE, although the impact of gradient measurement errors in the latter is underestimated in our simulation. Further work is required to determine whether noisy measurement favors the variational approach, and whether the level of noise plays a role. By choosing the most energy-relevant operators, the NoVa-ADAPT algorithm saves quantum and classical resources compared to previous feedback-based quantum algorithms that either include all

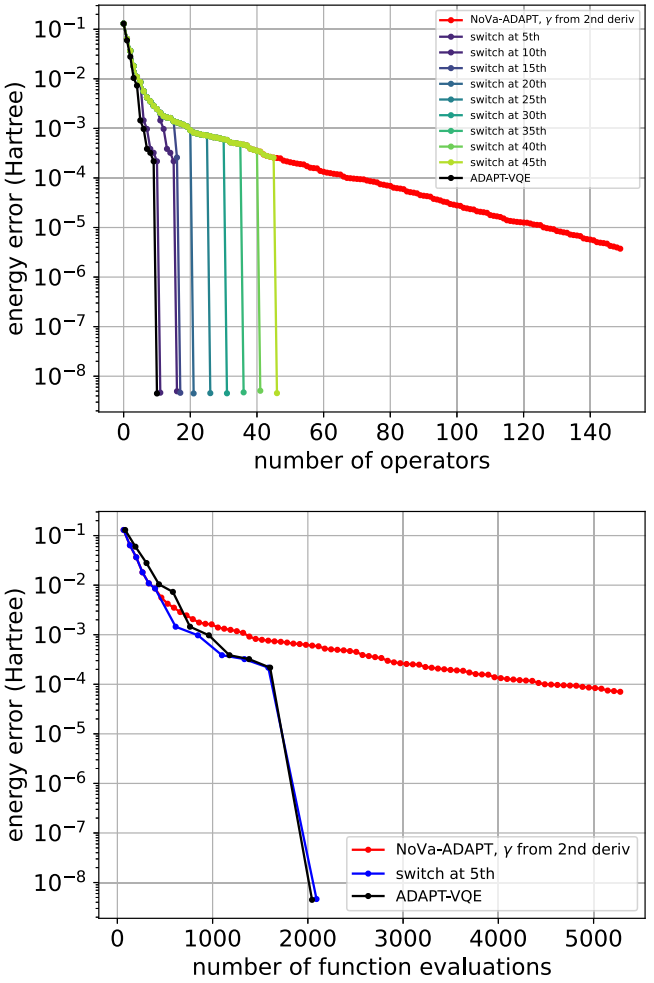

FIG. 8. Results for H_4 molecule of bond length 1.5 Å using the NoVa-ADAPT algorithm (red), ADAPT-VQE (black), and procedures starting with the nonvariational update and switching to ADAPT-VQE at the 5, 10, 15, 20, 25, 30, 35, 40, 45-th iteration (see legend). Energy error is plotted as a function of the number of operators in the state (upper panel), and as a function of the number of function evaluations (lower panel).

operators in the Hamiltonian or use random operators. In the case of a sub-optimal initial state, we find that both the nonvariational ADAPT approach and the approach of adding randomized operators perform better than ADAPT-VQE in terms of measurement cost. We also explore the strategy of combining ADAPT-VQE and the nonvariational approach, which may provide a realistic way to construct the ground state using limited resources.

Due to limitation in classical computational power, our simulation is carried out only for small-scale problems. We would like to stress that simulation for larger-scale problems (e.g., with techniques in Ref. [58]) is needed to infer whether the NoVa-ADAPT algorithm can outperform ADAPT-VQE either in the ideal situation or with errors for problems beyond classical simulability. More generally, one can adjust the extent of optimization in the ground-state-finding algorithm and explore the resource costs required for a given energy accuracy, a point alluded to in Sec. VII. Other than the hybrid approach in Sec. VII, the extent of optimization

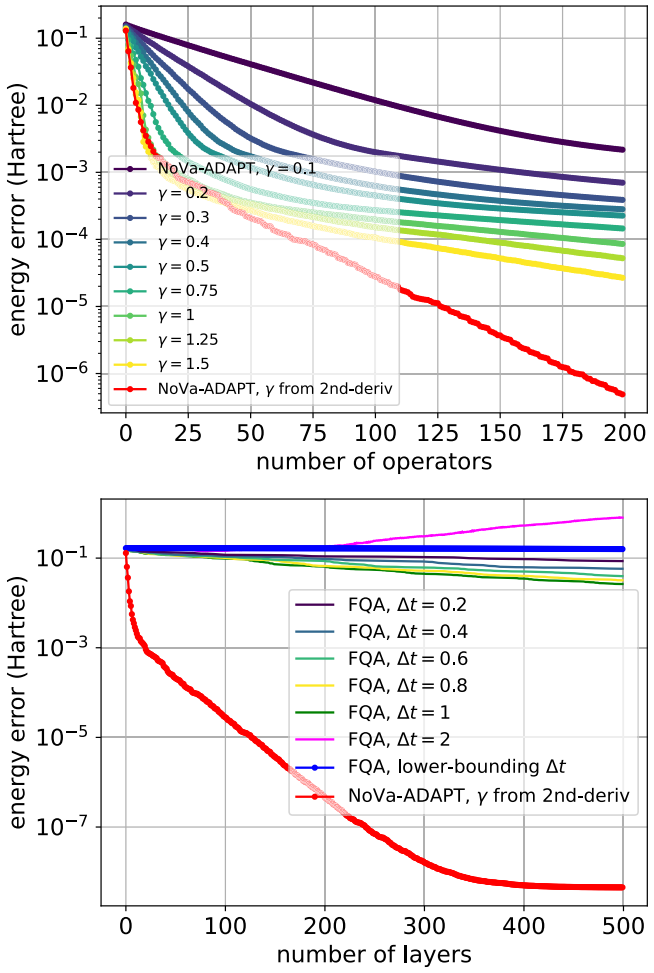

FIG. 9. Results for H_4 molecule of bond length 1.5 Å produced by the NoVa-ADAPT algorithm with the spin-adapted fermionic operator pool (upper panel) and FQA (lower panel), where energy error is plotted as a function of the number of operators added to the state. Different curves show the results with different constant γ values. The red curve is the result from the NoVa-ADAPT algorithm with γ calculated from the second derivative using Eq. (10).

can also be engineered within each adaptive step, by including optimization of only the new parameter (like in Ref. [49]) or the last few added parameters. The accuracy of this optimization procedure can be further relaxed to save resource cost by limiting the number of iterations in the optimization, i.e., terminating without convergence. Furthermore, measurement cost associated with the VQE subroutine in ADAPT-VQE can be reduced by leveraging classical simulation techniques for pre-optimization [58,59]. We expect this pre-optimization strategy to also benefit the NoVa-ADAPT algorithm, as it can be applied to the gradient measurement procedure in the NoVa-ADAPT algorithm. It would be interesting to compare the improvements provided by such approaches in the NoVa-ADAPT algorithm and in ADAPT-VQE.

Our results suggest that algorithms for finding the ground state can benefit from combining the adaptive operator selection strategy and techniques in classical optimization algorithms. The update rule in the NoVa-ADAPT algorithm is inspired by first-order methods. In a similar way, one can design an update rule according to techniques based on higher-order derivatives. We have observed that partial information about second-order derivatives, as in Eq. (8), can speed up energy reduction in the NoVa-ADAPT algorithm. More complete information about the Hessian may bring further improvement, as shown in Ref. [60], where quasi-second-order techniques can accelerate the quantum contracted-eigenvalue-equation solver.

ACKNOWLEDGMENTS

This work was supported by NSF Grant No. 2231328. A.B.M. acknowledges support from Sandia National Laboratories' Laboratory Directed Research and Development Program under the Truman Fellowship. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

APPENDIX A: CONSTANT γ VALUES

If γ is set to be a constant in the nonvariational state update approach, it is then a hyperparameter in the algorithm. A suitable value can lead to fast convergence to a local minimum without oscillatory or divergent behavior. Figure 9 shows the performance of the NoVa-ADAPT algorithm and FQA with different values of a constant γ. In the lower panel of Fig. 9, we see that with different constant γ's, the FQA

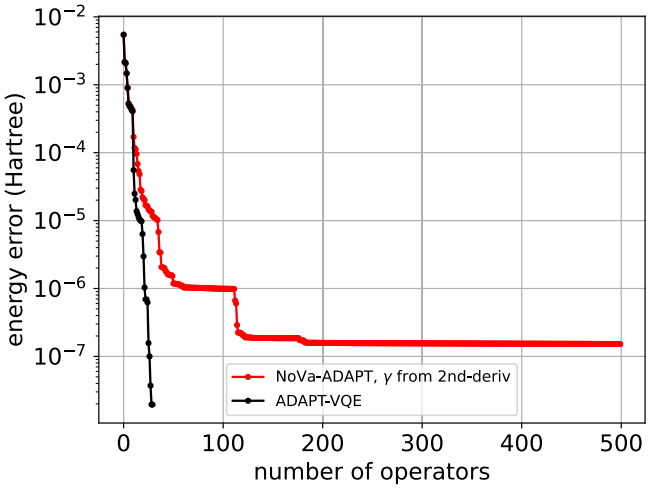

FIG. 10. Results for LiH molecule of bond length 1.5 Å with the spin-adapted fermionic operator pool. Energy error is plotted as a function of the number of operators added to the state. Red curve shows the result of the NoVa-ADAPT algorithm with γ calculated from the second derivative using Eq. (10). Black curve shows the result of ADAPT-VQE, where classical optimization is implemented.

lowers the energy at different rates. Note that the energy will increase with the number of layers when $\gamma = 2.0$. We display the case with $\gamma = 1$ for these two algorithms in the main text, as an empirical choice leading to reasonably fast energy reduction.

APPENDIX B: SIMULATED HAMILTONIAN IN QUBIT OPERATIONS

We use Jordan-Wigner transformation to map the Hamiltonian for the H_4 molecule in the STO-3G basis with a bond distance of 1.5 Å to the following:

$$
\begin{aligned}
H = & -0.92094 - 0.03974X_0X_1Y_2Y_3 - 0.00906X_0X_1Y_2Z_3Z_4Z_5Z_6Y_7 - 0.00906X_0X_1X_3Z_4Z_5X_6 - 0.02877X_0X_1Y_4Y_5 \\
& - 0.02749X_0X_1Y_6Y_7 + 0.03974X_0Y_1Y_2X_3 + 0.00906X_0Y_1Y_2Z_3Z_4Z_5Z_6X_7 - 0.00906X_0Y_1Y_3Z_4Z_5X_6 \\
& + 0.02877X_0Y_1Y_4X_5 + 0.02749X_0Y_1Y_6X_7 + 0.02080X_0Z_1X_2X_3Z_4X_5 + 0.02080X_0Z_1X_2Y_3Z_4Y_5 \\
& - 0.01998X_0Z_1X_2X_4Z_5X_6 - 0.01161X_0Z_1X_2Y_4Z_5Y_6 - 0.04000X_0Z_1X_2X_5Z_6X_7 - 0.04000X_0Z_1X_2Y_5Z_6Y_7 \\
& - 0.00837X_0Z_1Y_2Y_4Z_5X_6 + 0.02839X_0Z_1Z_2X_3Y_4Z_5Z_6Y_7 + 0.02001X_0Z_1Z_2X_3X_5X_6 - 0.02839X_0Z_1Z_2Y_3Y_4Z_5Z_6X_7 \\
& + 0.02001X_0Z_1Z_2Y_3Y_5X_6 + 0.00650X_0Z_1Z_2Z_3X_4 - 0.00297X_0Z_1Z_2Z_3X_4Z_5 + 0.00860X_0Z_1Z_2Z_3X_4Z_6 \\
& + 0.01748X_0Z_1Z_2Z_3X_4Z_7 + 0.00888X_0Z_1Z_2Z_3Z_4X_5Y_6Y_7 - 0.00888X_0Z_1Z_2Z_3Z_4Y_5Y_6X_7 - 0.00402X_0Z_1Z_2X_4 \\
& + 0.01678X_0Z_1Z_3X_4 + 0.01684X_0Z_2Z_3X_4 + 0.03974Y_0X_1X_2Y_3 + 0.00906Y_0X_1X_2Z_3Z_4Z_5Z_6Y_7 \\
& - 0.00906Y_0X_1X_3Z_4Z_5Y_6 + 0.02877Y_0X_1X_4Y_5 + 0.02749Y_0X_1X_6Y_7 - 0.03974Y_0Y_1X_2X_3 \\
& - 0.00906Y_0Y_1X_2Z_3Z_4Z_5Z_6X_7 - 0.00906Y_0Y_1Y_3Z_4Z_5Y_6 - 0.02877Y_0Y_1X_4X_5 - 0.02749Y_0Y_1X_6X_7 \\
& - 0.00837Y_0Z_1X_2X_4Z_5Y_6 + 0.02080Y_0Z_1Y_2X_3Z_4X_5 + 0.02080Y_0Z_1Y_2Y_3Z_4Y_5 - 0.01161Y_0Z_1Y_2X_4Z_5X_6 \\
& - 0.01998Y_0Z_1Y_2Y_4Z_5Y_6 - 0.04000Y_0Z_1Y_2X_5Z_6X_7 - 0.04000Y_0Z_1Y_2Y_5Z_6Y_7 - 0.02839Y_0Z_1Z_2X_3X_4Z_5Z_6Y_7 \\
& + 0.02001Y_0Z_1Z_2X_3X_5Y_6 + 0.02839Y_0Z_1Z_2Y_3X_4Z_5Z_6X_7 + 0.02001Y_0Z_1Z_2Y_3Y_5Y_6 + 0.00650Y_0Z_1Z_2Z_3Y_4 \\
& - 0.00297Y_0Z_1Z_2Z_3Y_4Z_5 + 0.00860Y_0Z_1Z_2Z_3Y_4Z_6 + 0.01748Y_0Z_1Z_2Z_3Y_4Z_7 - 0.00888Y_0Z_1Z_2Z_3Z_4X_5X_6Y_7 \\
& + 0.00888Y_0Z_1Z_2Z_3Z_4Y_5X_6X_7 - 0.00402Y_0Z_1Z_2Y_4 + 0.01678Y_0Z_1Z_3Y_4 + 0.01684Y_0Z_2Z_3Y_4 \\
& + 0.11933Z_0 + 0.01684Z_0X_1Z_2Z_3Z_4X_5 + 0.01684Z_0Y_1Z_2Z_3Z_4Y_5 + 0.10125Z_0Z_1 \\
& - 0.00839Z_0X_2Z_3Z_4Z_5X_6 - 0.00839Z_0Y_2Z_3Z_4Z_5Y_6 + 0.05022Z_0Z_2 - 0.01746Z_0X_3Z_4Z_5Z_6X_7 \\
& - 0.01746Z_0Y_3Z_4Z_5Z_6Y_7 + 0.08996Z_0Z_3 + 0.06236Z_0Z_4 + 0.09114Z_0Z_5 \\
& + 0.07784Z_0Z_6 + 0.10533Z_0Z_7 - 0.02080X_1X_2Y_3Y_4 + 0.02001X_1X_2X_4Z_5Z_6X_7 \\
& + 0.02839X_1X_2Y_5Y_6 + 0.02080X_1Y_2Y_3X_4 + 0.02001X_1Y_2Y_4Z_5Z_6X_7 - 0.02839X_1Y_2Y_5X_6 \\
& - 0.04000X_1Z_2X_3X_4Z_5X_6 - 0.04000X_1Z_2X_3Y_4Z_5Y_6 - 0.01998X_1Z_2X_3X_5Z_6X_7 - 0.01161X_1Z_2X_3Y_5Z_6Y_7 \\
& - 0.00837X_1Z_2Y_3Y_5Z_6X_7 + 0.00888X_1Z_2Z_3X_4X_6X_7 + 0.00888X_1Z_2Z_3Y_4Y_6X_7 + 0.00650X_1Z_2Z_3Z_4X_5 \\
& + 0.01748X_1Z_2Z_3Z_4X_5Z_6 + 0.00860X_1Z_2Z_3Z_4X_5Z_7 - 0.00297X_1Z_2Z_3X_5 + 0.01678X_1Z_2Z_4X_5 \\
& - 0.00402X_1Z_3Z_4X_5 + 0.02080Y_1X_2X_3Y_4 + 0.02001Y_1X_2X_4Z_5Z_6Y_7 - 0.02839Y_1X_2X_5Y_6 \\
& - 0.02080Y_1Y_2X_3X_4 + 0.02001Y_1Y_2Y_4Z_5Z_6Y_7 + 0.02839Y_1Y_2X_5X_6 - 0.00837Y_1Z_2X_3X_5Z_6Y_7 \\
& - 0.04000Y_1Z_2Y_3X_4Z_5X_6 - 0.04000Y_1Z_2Y_3Y_4Z_5Y_6 - 0.01161Y_1Z_2Y_3X_5Z_6X_7 - 0.01998Y_1Z_2Y_3Y_5Z_6Y_7 \\
& + 0.00888Y_1Z_2Z_3X_4X_6Y_7 + 0.00888Y_1Z_2Z_3Y_4Y_6Y_7 + 0.00650Y_1Z_2Z_3Z_4Y_5 + 0.01748Y_1Z_2Z_3Z_4Y_5Z_6 \\
& + 0.00860Y_1Z_2Z_3Z_4Y_5Z_7 - 0.00297Y_1Z_2Z_3Y_5 + 0.01678Y_1Z_2Z_4Y_5 - 0.00402Y_1Z_3Z_4Y_5 \\
& + 0.11933Z_1 - 0.01746Z_1X_2Z_3Z_4Z_5X_6 - 0.01746Z_1Y_2Z_3Z_4Z_5Y_6 + 0.08996Z_1Z_2 \\
& - 0.00839Z_1X_3Z_4Z_5Z_6X_7 - 0.00839Z_1Y_3Z_4Z_5Z_6Y_7 + 0.05022Z_1Z_3 + 0.09114Z_1Z_4 \\
& + 0.06236Z_1Z_5 + 0.10533Z_1Z_6 + 0.07784Z_1Z_7 - 0.03517X_2X_3Y_4Y_5 \\
& - 0.02944X_2X_3Y_6Y_7 + 0.03517X_2Y_3Y_4X_5 + 0.02944X_2Y_3Y_6X_7 - 0.02174X_2Z_3X_4X_5Z_6X_7 \\
& - 0.02174X_2Z_3X_4Y_5Z_6Y_7 + 0.01055X_2Z_3Z_4Z_5X_6 - 0.01865X_2Z_3Z_4Z_5X_6Z_7 + 0.00330X_2Z_3Z_4X_6 \\
& - 0.01844X_2Z_3Z_5X_6 + 0.00261X_2Z_4Z_5X_6 + 0.03517Y_2X_3X_4Y_5 + 0.02944Y_2X_3X_6Y_7 \\
& - 0.03517Y_2Y_3X_4X_5 - 0.02944Y_2Y_3X_6X_7 - 0.02174Y_2Z_3Y_4X_5Z_6X_7 - 0.02174Y_2Z_3Y_4Y_5Z_6Y_7 \\
& + 0.01055Y_2Z_3Z_4Z_5Y_6 - 0.01865Y_2Z_3Z_4Z_5Y_6Z_7 + 0.00330Y_2Z_3Z_4Y_6 - 0.01844Y_2Z_3Z_5Y_6
\end{aligned}
$$

$$
\begin{aligned}
&+ 0.00261 Y_2 Z_4 Z_5 Y_6 + 0.07128 Z_2 + 0.00261 Z_2 X_3 Z_4 Z_5 Z_6 X_7 + 0.00261 Z_2 Y_3 Z_4 Z_5 Z_6 Y_7 \\
&+ 0.09406 Z_2 Z_3 + 0.05893 Z_2 Z_4 + 0.09410 Z_2 Z_5 + 0.06483 Z_2 Z_6 \\
&+ 0.09428 Z_2 Z_7 + 0.02174 X_3 X_4 Y_5 Y_6 - 0.02174 X_3 Y_4 Y_5 X_6 + 0.01055 X_3 Z_4 Z_5 Z_6 X_7 \\
&- 0.01865 X_3 Z_4 Z_5 X_7 - 0.01844 X_3 Z_4 Z_6 X_7 + 0.00330 X_3 Z_5 Z_6 X_7 - 0.02174 Y_3 X_4 X_5 Y_6 \\
&+ 0.02174 Y_3 Y_4 X_5 X_6 + 0.01055 Y_3 Z_4 Z_5 Z_6 Y_7 - 0.01865 Y_3 Z_4 Z_5 Y_7 - 0.01844 Y_3 Z_4 Z_6 Y_7 \\
&+ 0.00330 Y_3 Z_5 Z_6 Y_7 + 0.07128 Z_3 + 0.09410 Z_3 Z_4 + 0.05893 Z_3 Z_5 \\
&+ 0.09428 Z_3 Z_6 + 0.06483 Z_3 Z_7 - 0.04234 X_4 X_5 Y_6 Y_7 + 0.04234 X_4 Y_5 Y_6 X_7 \\
&+ 0.04234 Y_4 X_5 X_6 Y_7 - 0.04234 Y_4 Y_5 X_6 X_7 - 0.00689 Z_4 + 0.09690 Z_4 Z_5 \\
&+ 0.05391 Z_4 Z_6 + 0.09626 Z_4 Z_7 - 0.00689 Z_5 + 0.09626 Z_5 Z_6 \\
&+ 0.05391 Z_5 Z_7 - 0.10062 Z_6 + 0.11281 Z_6 Z_7 - 0.10062 Z_7
\end{aligned}
\tag{B1}
$$

APPENDIX C: SIMULATION FOR LiH MOLECULE

In Fig. 10, we show the performances of the NoVa-ADAPT algorithm and ADAPT-VQE in the simulation for LiH molecule of bond length 1.5 Å with 12 qubits. Both algorithms require similar numbers of operators to reach chemical accuracy. Beyond that, the NoVa-ADAPT algorithm requires more operators to reach the same energy accuracy, similar to the case of H_4 in the main text.

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[2] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, arXiv:quant-ph/9511026.

[3] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Simulation of electronic structure hamiltonians using quantum computers, Mol. Phys. **109**, 735 (2011).

[4] D. S. Abrams and S. Lloyd, Simulation of many-body Fermi systems on a universal quantum computer, Phys. Rev. Lett. **79**, 2586 (1997).

[5] D. S. Abrams and S. Lloyd, Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors, Phys. Rev. Lett. **83**, 5162 (1999).

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

[7] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[8] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, Hybrid quantum-classical approach to correlated materials, Phys. Rev. X **6**, 031045 (2016).

[9] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum **3**, 191 (2019).

[10] M. Motta, C. Sun, A. T. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. Brandao, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

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

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

[13] 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).

[14] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature (London) **549**, 242 (2017).

[15] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure, Phys. Rev. A **95**, 020501(R) (2017).

[16] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Quantum chemistry calculations on a trapped-ion quantum simulator, Phys. Rev. X **8**, 031022 (2018).

[17] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).

[18] H. L. Tang, V. O. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor, PRX Quantum **2**, 020310 (2021).

[19] P. G. Anastasiou, Y. Chen, N. J. Mayhall, E. Barnes, and S. E. Economou, TETRIS-ADAPT-VQE: An adaptive algorithm that yields shallower, denser circuit ansätze, Phys. Rev. Res. **6**, 013254 (2024).

[20] C.-F. Chen, H.-Y. Huang, J. Preskill, and L. Zhou, Local minima in quantum systems, Nat. Phys. **21**, 654 (2025).

[21] L. Bittel and M. Kliesch, Training variational quantum algorithms is NP-hard, Phys. Rev. Lett. **127**, 120502 (2021).

[22] D. A. Mazziotti, Anti-Hermitian contracted Schrödinger equation: Direct determination of the two-electron reduced density

matrices of many-electron molecules, Phys. Rev. Lett. **97**, 143002 (2006).

[23] A. B. Magann, K. M. Rudinger, M. D. Grace, and M. Sarovar, Feedback-based quantum optimization, Phys. Rev. Lett. **129**, 250502 (2022).

[24] A. B. Magann, K. M. Rudinger, M. D. Grace, and M. Sarovar, Lyapunov-control-inspired strategies for quantum combinatorial optimization, Phys. Rev. A **106**, 062414 (2022).

[25] S. A. Rahman, H. G. Clausen, O. Karabacak, and R. Wisniewski, Adaptive sampling noise mitigation technique for feedback-based quantum algorithms, in *Computational Science ICCS 2024: 24th International Conference, Malaga, Spain, July 2–4, 2024, Proceedings, Part VI* (Springer-Verlag, Berlin, Heidelberg, 2024), pp. 321–329.

[26] G. E. L. Pexe, L. A. M. Rattighieri, A. L. Malvezzi, and F. F. Fanchini, Using a feedback-based quantum algorithm to analyze the critical properties of the Annni model without classical optimization, Phys. Rev. B **110**, 224422 (2024).

[27] S. A. Rahman, Ö. Karabacak, and R. Wisniewski, Feedback-based quantum algorithm for constrained optimization problems, in *Parallel Processing and Applied Mathematics*, edited by R. Wyrzykowski, J. Dongarra, E. Deelman, and K. Karczewski (Springer Nature Switzerland, Cham, 2025), pp. 277–289.

[28] S. A. Rahman, Ö. Karabacak, and R. Wisniewski, Feedback-based quantum algorithm for excited states calculation, arXiv:2404.04620.

[29] R. K. Malla, H. Sukeno, H. Yu, T.-C. Wei, A. Weichselbaum, and R. M. Konik, Feedback-based quantum algorithm inspired by counterdiabatic driving, Phys. Rev. Res. **6**, 043068 (2024).

[30] D. Arai, K. N. Okada, Y. Nakano, K. Mitarai, and K. Fujii, Scalable circuit depth reduction in feedback-based quantum optimization with a quadratic approximation, Phys. Rev. Res. **7**, 013035 (2025).

[31] L. T. Brady and S. Hadfield, FOCQS: Feedback optimally controlled quantum states, arXiv:2409.15426.

[32] P. Chandarana, K. Paul, K. Ranjan Swain, X. Chen, and A. del Campo, Lyapunov controlled counterdiabatic quantum optimization, arXiv:2409.12525.

[33] J. B. Larsen, M. D. Grace, A. D. Baczewski, and A. B. Magann, Feedback-based quantum algorithms for ground state preparation, Phys. Rev. Res. **6**, 033336 (2024).

[34] T. Schulte-herbrüggen, S. J. Glaser, G. Dirr, and U. Helmke, Gradient flows for optimization in quantum information and quantum dynamics: Foundations and applications, Rev. Math. Phys. **22**, 597 (2010).

[35] R. Wiersema and N. Killoran, Optimizing quantum circuits with riemannian gradient flow, Phys. Rev. A **107**, 062421 (2023).

[36] A. B. Magann, S. E. Economou, and C. Arenz, Randomized adaptive quantum state preparation, Phys. Rev. Res. **5**, 033227 (2023).

[37] E. Malvetti, C. Arenz, G. Dirr, and T. Schulte-Herbrüggen, Randomized gradient descents on riemannian manifolds: Almost sure convergence to global minima in and beyond quantum optimization, arXiv:2405.12039.

[38] V. O. Shkolnikov, N. J. Mayhall, S. E. Economou, and E. Barnes, Avoiding symmetry roadblocks and minimizing the measurement overhead of adaptive variational quantum eigensolvers, Quantum **7**, 1040 (2023).

[39] L. W. Bertels, H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, Symmetry breaking slows convergence of the adapt variational quantum eigensolver, J. Chem. Theory Comput. **18**, 6656 (2022).

[40] A. Warren, L. Zhu, N. J. Mayhall, E. Barnes, and S. E. Economou, Adaptive variational algorithms for quantum Gibbs state preparation, arXiv:2203.12757.

[41] A. M. Romero, J. Engel, H. L. Tang, and S. E. Economou, Solving nuclear structure problems with the adaptive variational quantum algorithm, Phys. Rev. C **105**, 064317 (2022).

[42] H. Grimsley, G. Barron, E. Barnes, S. Economou, and N. Mayhall, Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus, npj Quantum Inf. **9**, 19 (2023).

[43] C. K. Long, K. Dalton, C. H. W. Barnes, D. R. M. Arvidsson-Shukur, and N. Mertig, Layering and subpool exploration for adaptive variational quantum eigensolvers: Reducing circuit depth, runtime, and susceptibility to noise, Phys. Rev. A **109**, 042413 (2024).

[44] J. S. Van Dyke, K. Shirali, G. S. Barron, N. J. Mayhall, E. Barnes, and S. E. Economou, Scaling adaptive quantum simulation algorithms via operator pool tiling, Phys. Rev. Res. **6**, L012030 (2024).

[45] M. Ramôa, L. Paulo Santos, N. J. Mayhall, E. Barnes, and S. E. Economou, Reducing measurement costs by recycling the hessian in adaptive variational quantum algorithms, Quantum Sci. Technol. **10**, 015031 (2025).

[46] M. Ramôa, P. G. Anastasiou, L. P. Santos, N. J. Mayhall, E. Barnes, and S. E. Economou, Reducing the resources required by ADAPT-VQE using coupled exchange operators and improved subroutines, npj Quantum Info. **11**, 86 (2025).

[47] R. C. Farrell, M. Illa, A. N. Ciavarella, and M. J. Savage, Scalable circuits for preparing ground states on digital quantum computers: The schwinger model vacuum on 100 qubits, PRX Quantum **5**, 020315 (2024).

[48] A. Beck, *First-Order Methods in Optimization* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017), Chap. 5, https://epubs.siam.org/doi/pdf/10.1137/1.9781611974997.

[49] S. E. Smart and D. A. Mazziotti, Quantum solver of contracted eigenvalue equations for scalable molecular simulations on quantum computing devices, Phys. Rev. Lett. **126**, 070504 (2021).

[50] Y. Nakata, C. Hirche, C. Morgan, and A. Winter, Unitary 2-designs from random x- and z-diagonal unitaries, J. Math. Phys. **58**, 052203 (2017).

[51] J.-N. Boyn, A. O. Lykhin, S. E. Smart, L. Gagliardi, and D. A. Mazziotti, Quantum-classical hybrid algorithm for the simulation of all-electron correlation, J. Chem. Phys. **155**, 244106 (2021).

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

[53] A. Y. Kitaev, Quantum computations: Algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[54] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algorithm, Quantum Inf. Comput. 6, 81 (2006), arXiv:quant-ph/0505030.

[55] N. J. Ross and P. Selinger, Optimal ancilla-free clifford+t approximation of z-rotations, Quantum Inf. Comput. **16**, 901 (2016).

[56] D. Mondal, D. Halder, S. Halder, and R. Maitra, Development of a compact ansatz via operator commutativity screening: Digital quantum simulation of molecular systems, J. Chem. Phys. **159**, 014105 (2023).

[57] D. Halder, D. Mondal, and R. Maitra, Noise-independent route toward the genesis of a compact ansatz for molecular energetics: A dynamic approach, J. Chem. Phys. **160**, 124104 (2024).

[58] J. W. Mullinax and N. M. Tubman, Large-scale sparse wave function circuit simulator for applications with the variational quantum eigensolver, Chem. Phys. **162**, 074114 (2025).

[59] J. W. Mullinax, P. G. Anastasiou, J. Larson, S. E. Economou, and N. M. Tubman, Classical preoptimization approach for ADAPT-VQE: Maximizing the potential of high-performance computing resources to improve quantum simulation of chemical applications, J. Chem. Theory Comput. **21**, 4006 (2025).

[60] S. E. Smart and D. A. Mazziotti, Accelerated convergence of contracted quantum eigensolvers through a quasi-second-order, locally parameterized optimization, J. Chem. Theory Comput. **18**, 5286 (2022).