ARTICLE OPEN

Variational quantum eigensolver with reduced circuit complexity

Yu Zhang [1✉], Lukasz Cincio[1], Christian F. A. Negre[1], Piotr Czarnik[1], Patrick J. Coles [1], Petr M. Anisimov [2], Susan M. Mniszewski [3], Sergei Tretiak [1,4] and Pavel A. Dub [5]

The variational quantum eigensolver (VQE) is one of the most promising algorithms to find eigenstates of a given Hamiltonian on noisy intermediate-scale quantum devices (NISQ). The practical realization is limited by the complexity of quantum circuits. Here we present an approach to reduce quantum circuit complexity in VQE for electronic structure calculations. Our ClusterVQE algorithm splits the initial qubit space into clusters which are further distributed on individual (shallower) quantum circuits. The clusters are obtained based on mutual information reflecting maximal entanglement between qubits, whereas inter-cluster correlation is taken into account via a new "dressed" Hamiltonian. ClusterVQE therefore allows exact simulation of the problem by using fewer qubits and shallower circuit depth at the cost of additional classical resources, making it a potential leader for quantum chemistry simulations on NISQ devices. Proof-of-principle demonstrations are presented for several molecular systems based on quantum simulators as well as IBM quantum devices.

INTRODUCTION

One of the major goals of computational chemistry is the development of methods and algorithms for the calculation of the molecular electronic ground and excited state energies and corresponding wave functions from first-principles. Such eigenvalues and eigenvectors can be obtained from the solution of the time-independent electronic Schrödinger equation. However, since the invention of classical digital computers in the early 1940s, the exact numerical solution of this central quantum mechanical equation remains infeasible for systems roughly having more than 12 electrons distributed on 184 spin-orbitals[1]. Even though modern quantum many-body methods, such as density matrix renormalization group (DMRG)[2], selected configuration interaction (sCI)[3], and coupled-cluster (CC) theory[4] methods can solve larger systems, their applications are still limited to dozens of electrons and they are subject to truncation or approximations. The reason is that the solution space (i.e., the Fock space) grows factorially with the system size (e.g., the number of electrons and basis functions). One of the most promising and immediate applications of quantum computers is solving classically intractable quantum chemistry problems[5,6].

The direct solution to these problems on currently available quantum computer hardware is intractable. For example, the quantum phase estimation (QPE) algorithm[7] represents the natural translation of the full configuration interaction (FCI) procedure to quantum computers[8]. However, QPE requires millions of qubits and quantum gates even for relatively small systems, making the algorithm unsuitable for practical applications on near-term noisy intermediate-scale quantum (NISQ) devices[9]. Full quantum eigensolver (FQE), in which the complexity of basic gate operators is polylogarithmic in the number of spin-orbitals, represents another fully quantum algorithm to simulate molecular systems[10]. However, FQE might not be suitable for advantageous quantum chemistry demonstrations due to the

limitations of NISQ devices. Indeed, leading digital quantum computers based on superconducting qubit technology have limited coherence times (~100 μs) and gate error rates (~2×10^{-2} for a two-qubit gate), limiting the possible number of operations that can be executed to evolve the quantum state. Under these conditions, leveraging classical resources as much as possible through a hybrid quantum-classical approach seems to be the most promising route toward achieving quantum chemical advantage on NISQ devices[11–13].

The variational quantum eigensolver (VQE) is one of the leading hybrid quantum-classical algorithms developed specifically to simulate molecular systems in their ground[14–17] and ultimately excited states[18–23]. In a typical VQE setup, a variational parameterized ansatz is used to represent the trial wave function prepared with a quantum circuit and the expectation value of the qubit Hamiltonian is measured. Then, the parameters of the ansatz are iteratively optimized on a classical computer using the Rayleigh-Ritz variational principle. Although VQE simulations of small molecules have been performed on various quantum architectures such as photons[14], superconducting qubits[24,25] and trapped ions[26], major efforts are needed to scale up this approach to larger molecular systems of chemical interest. Here, albeit not as large as QPE circuits, VQE would also suffer from the large size of the quantum circuits, which coupled with a classical optimization requiring many variational parameters, can render calculations intractable. Consequently, the construction of an efficacious ansatz and improved classical optimizer is an active area of research[12,13].

There are two main approaches for ansatz design for electronic structure calculations. Hardware-efficient ansatzes[24,25] are constructed from repeated, dense blocks of a limited set of parameterized gates that are easy to implement on a quantum device. While such an ansatz requires fewer resources on the quantum processor, because of its totally chemically agnostic

[1]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [2]Accelerators and Electrodynamics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [3]Computer, Computational and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [4]Center for Integrated Nanotechnologies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [5]Chemistry Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. ✉email: zhy@lanl.gov

nature, it might face difficulties in the parameter optimization due to the barren-plateau problem[27–30]. Consequently, simulation of large molecular systems with hardware-efficient ansatzes might become practically challenging and even impossible. Chemically-inspired ansatzes, such as the Unitary Coupled-Cluster Singles and Doubles (UCCSD)[31–35], are designed by using domain knowledge from classical quantum chemistry. The standard version of UCCSD, however, has unfavorable scaling in the number of gates required for larger molecular systems[36]. Consequently, various approaches are being developed to improve on the UCCSD method to obtain shorter circuits or higher than UCCSD accuracy at comparable circuit depth[17,37–41]. Among them, "iterative VQE" algorithms[42–47], which instead of using a fixed ansatz, construct problem-tailored ansatzes on-the-fly, have gained attention because of their controllable circuit-size.

Two notable iterative VQE algorithms to reduce quantum circuit depth are (fermionic and qubit) ADAPT-VQE[42,43] and iterative Qubit Coupled Cluster (iQCC)[44–46]. Fermionic-ADAPT-VQE employs a predetermined pool of spin-adapted fermionic (generalized) singles and doubles excitation operators from which the ansatz is dynamically constructed[42]. The more measurement- and circuit-efficient (e.g. less CNOT gates in particular) qubit-ADAPT-VQE uses selected Pauli words (obtained from mapping fermionic operators into Pauli operators) to form the "qubit" pool, but the algorithm comes with a larger number of ADAPT iterations because of the larger number of variational parameters (more gradient calculations required)[43]. IQCC uses arbitrarily shallow quantum circuit depth at the expense of an iterative canonical Hamiltonian dressing technique employing the qubit coupled cluster (QCC) ansatz. The exponential growth of the "dressed" Hamiltonian with the size of the QCC transformation in an anti-commuting set could be mitigated using the involuntary linear combinations of Paulis technique[45]. But the algorithm is exponential with respect to the dressing steps.

Although circuit depth is one of the most critical metrics for NISQ devices, the number of physical qubits on the actual quantum chip (circuit width) is the necessary condition that limits the problem that can be solved. Several techniques based on symmetries presented in the Hamiltonian have been used to reduce the number of qubits[48–51]. But these techniques can only reduce a few qubits and have certain symmetry requirements.

Instead, there are recent efforts to solve larger problems with smaller quantum computers by either using tensor network or divide-and-conquer techniques[52–55]. In addition, a dressing technique has been proposed to downfold the correlation effect in a large space into a smaller active space by using the second-order approximation[56]. However, these methods are approximate or only applied to model systems. In this work, we present an algorithm that reduces both quantum circuit depth and width in VQE at the cost of additional classical resources. Our algorithm, called ClusterVQE, uses the mutual information (a measure of the correlation between spin-orbitals[57,58]) to group the qubits into different clusters, which can be distributed to different quantum circuits. The correlation between different clusters is minimized by the mutual information optimization method and is taken into account by building a "dressed" Hamiltonian. Even shallower circuit depths can be achieved compared to the qubit-ADAPT-VQE. Consequently, the ClusterVQE algorithm is more robust to noise, making it particularly attractive for NISQ devices. Because the correlations between different clusters are minimized, the dressed Hamiltonian protocol only introduces minor additional classical resource requirements compared to the iQCC method. Most importantly, the ClusterVQE algorithm removes the entanglement between different clusters via the dressed Hamiltonian. Consequently, ClusterVQE can solve the original problem in a much smaller qubit space (clusters), reducing the number of qubits significantly for simulating large molecules on NISQ devices.

RESULTS AND DISCUSSION

Simulation details

To demonstrate the validity of the ClusterVQE algorithm, we have simulated several molecules on both a quantum simulator and IBM quantum devices. In order to perform the simulations, an in-house modified Qiskit code[59] has been developed. The Jordan-Wigner operator transformation[60] and the Limited-memory BFGS Bound (L-BFGS-B) optimizer[61] were used in this work. The obtained results are compared with that of exact diagonalization. In order to make the comparison over different methods (VQE, qubit-ADAPT-VQE, iQCC) meaningful, the same ansatz, QUCCSD (details in Supplementary Section II), is used for all simulations
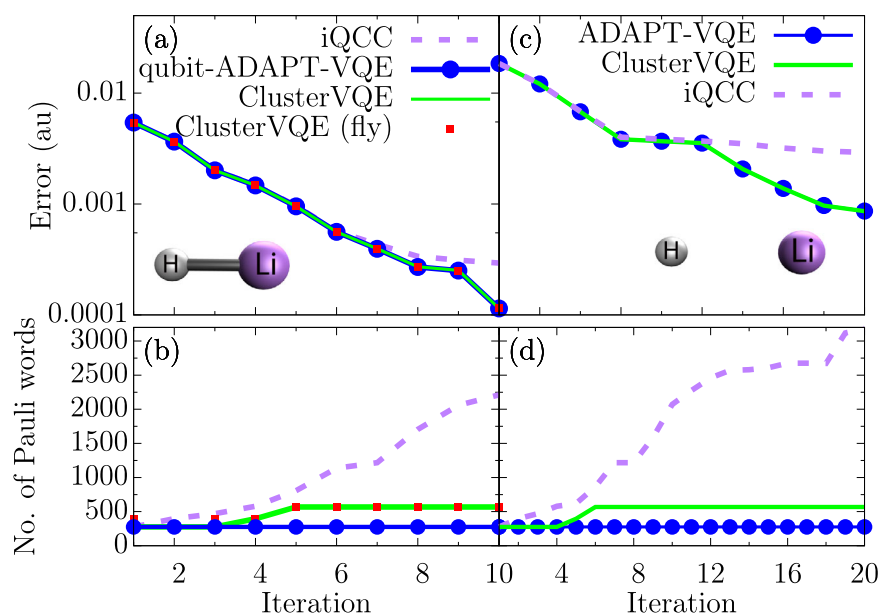

Fig. 1 Simulation of LiH molecule. a, c GS Energy errors and (**b, d**) number of Pauli words per each iteration for the LiH molecule obtained from the qubit-ADAPT-VQE, iQCC, and ClusterVQE algorithms. The Li-H bond lengths are 1.547 Å (**a, b**) and 2.4 Å (**c, d**), respectively.

below unless specified otherwise. The convergence criteria of $\epsilon = 10^{-4}$ and the minimal STO-3G basis set are used for all the molecules. The minimal basis set is chosen in the test because the NISQ devices suffer from noise due to the limitations of NISQ devices. Using a larger basis set directly will significantly increase the number of qubits and circuit depth, further amplifying the noise effect. To address the basis set problem, we recently proposed a trans-correlated Hamiltonian approach which could significantly improve the accuracy by using only the minimal basis set, and consequently, is more NISQ friendly[62].

LiH molecule

We start with simulating the ground state of the LiH molecule at equilibrium (bond distance is 1.547 Å). LiH requires 10 qubits with the minimal STO-3G basis set and frozen core orbitals. The graph representation of the qubit clustering within the ClusterVQE is shown in Supplementary Fig. 2. Qubits [0, 1, 4, 5, 6, 9] and [2, 3, 7, 8] are in the same cluster, respectively. Figure 1(a) and (c) compares the energy convergence of the qubit-ADAPT-VQE, ClusterVQE, and iQCC approaches for different Li-H bond lengths (1.547 Å and 2.4 Å). Because iQCC only implements one entangler on the quantum circuit, the number of parameters to be optimized is also smaller compared to that of qubit-ADAPT-VQE and ClusterVQE. Consequently, a smaller number of optimization cycles is required per iteration. For the iQCC method, once the entanglers in the previous iteration are used to dress the Hamiltonian, these parameters are fixed. In contrast, the parameterized ansatz in the qubit-ADAPT-VQE and ClusterVQE approaches are re-optimized at each iteration, making them more flexible to achieve better convergence. As a result, the convergence of iQCC is slightly worse than that of qubit-ADAPT-VQE for a given number of iterations, as shown in Fig. 1(a) and (c). For ClusterVQE, because only a few operators are used to dress the Hamiltonian, the performance is almost the same as that of qubit-ADAPT-VQE. Because the Hamiltonian of the iQCC and ClusterVQE methods is dressed by the entanglers, the number of Pauli words

in the dressed Hamiltonian is increased. However, compared to iQCC, only two operators are used to dress the Hamiltonian in ClusterVQE as shown in Fig. 1(b). The number of Pauli words of the dressed Hamiltonian does not increase significantly. In contrast, the number of Pauli words in iQCC increases exponentially with the number of iterations before saturation, see Fig. 1(b) and (d). But it should be noted that the performance of ClusterVQE depends on the quality of qubit clustering. If a less optimal qubit clustering is used, more entanglers will be used to dress the Hamiltonian. Consequently, the size of the dressed Hamiltonian increases significantly as shown in Supplementary Fig. 3 even though it is still smaller than that of iQCC.

We then compared the performance of the qubit-ADAPT-VQE, iQCC, and ClusterVQE approaches on the ground-state potential energy surface (PES) of LiH as shown in Fig. 2. Overall, ClusterVQE can achieve the same energy error as that of qubit-ADAPT-VQE but with smaller quantum circuits. And also uses fewer Pauli words compared to iQCC, as shown in Fig. 2.

We also compared the performance of the different clustering methods used (Supplementary Section I). For most geometries of the LiH molecule, these three methods produce the same clustering. However, at the bond length of 2.4 Å, the Metis graph partitioning method results in less optimal clustering. Consequently, more entanglers are needed to dress the Hamiltonian, resulting in 808 Pauli words. This comparison suggests that the QCD and MI-based clustering methods provide better and more stable performance in clustering and help improve the ClusterVQE performance.

Non-covalently-bound (van der Waals) complex $(H_2)_2$

Since ClusterVQE intends to minimize the intercluster correlations, it is intuitively expected that it is ideal for molecular aggregate systems where the intermolecular coupling is usually smaller than the intramolecular correlation. Here, we employ the non-covalently-bond (van der Waals) complex $(H_2)_2$[41] to demonstrate the application of ClusterVQE on molecular aggregates. The spin orbitals of each molecule of the complex are grouped in a cluster. The intramolecular correlation of H_2 is 0.04 au. By tuning the distance between the two H_2 molecules, the intermolecular correlations can be changed. The intermolecular correlations of the $(H_2)_2$ complex with distances 2 Å and 5 Å are 0.59 and 0.35 au, respectively. To demonstrate the complexity of the ClusterVQE method, the qubit-clustering is molecule-wise, i.e., the spin orbitals of the same molecule are grouped together. The ground state energies of the $(H_2)_2$ complex with different intermolecular distances (2 Å and 5 Å, respectively) are calculated, which are shown in Fig. 3. For the shorter distance, the intermolecular correlations are much stronger. Consequently, more inter-cluster entanglers are used to dress the Hamiltonian, making the complexity of ClusterVQE almost the same as iQCC. In contrast, the intermolecular correlations become smaller when the intermolecular distance is 5 Å. In this case, only 2 entanglers are used to dress the Hamiltonian. Moreover, increasing the intermolecular distance further reduces the intermolecular correlations and results in less dressing. This comparison indicates that ClusterVQE is ideal for the molecular aggregates where the intermolecular correlation is weak and the complexity increases with increasing inter-cluster correlations.

N$_2$ molecule

We next demonstrated the validity of ClusterVQE in the strongly correlated regime. The N_2 molecule is a prototype system for testing multi-reference methods in quantum chemistry due to its strong electronic correlations, particularly when the bond is stretched. Here N_2 at equilibrium (1.09 Å) and stretched configurations (1.6 Å) are used as examples. The 12-qubit
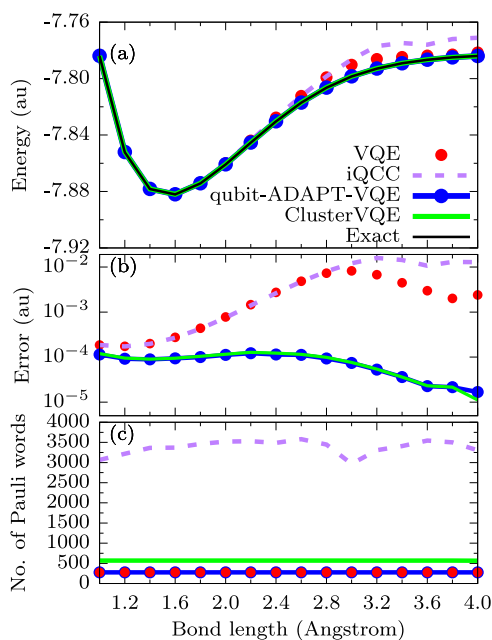

Fig. 2 Potential energy surface of LiH molecule. a GS energy, (**b**) energy error, and (**c**) number of Pauli words in the qubit Hamiltonian of LiH for different bond lengths obtained from different methods. For VQE, one layer of the QUCCSD ansatz is used. One layer of the QUCCSD ansatz is also used as the operator pool for the other methods.
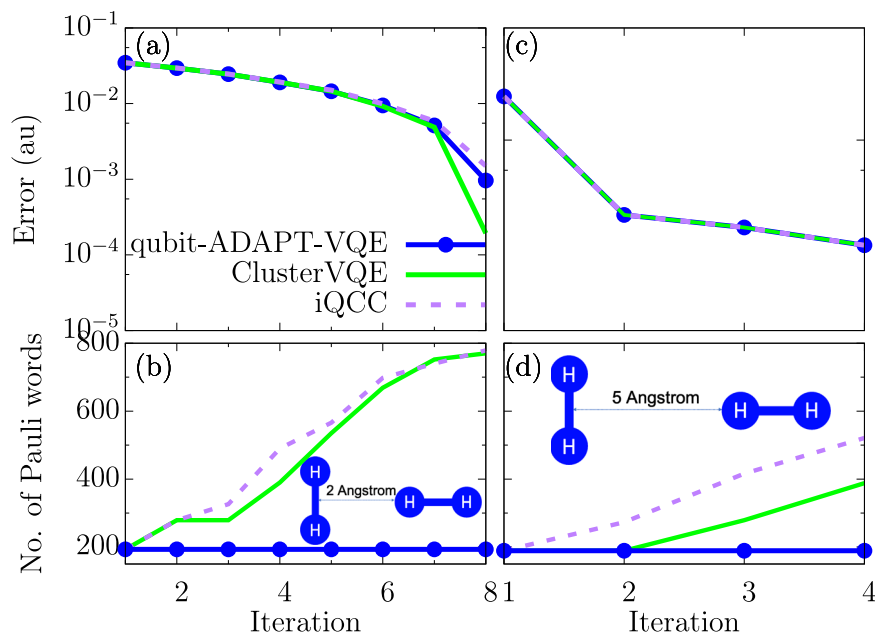
Fig. 3 Simulation of the (H₂)₂ complex molecule. a, c GS Energy errors and (**b, d**) number of Pauli words per each iteration for the $(H_2)_2$ complex molecule obtained from the qubit-ADAPT-VQE (blue), iQCC (purple) and ClusterVQE (green) methods. The intermolecular distances are 2 Å (**a, b**) and 5 Å (**c, d**), respectively.
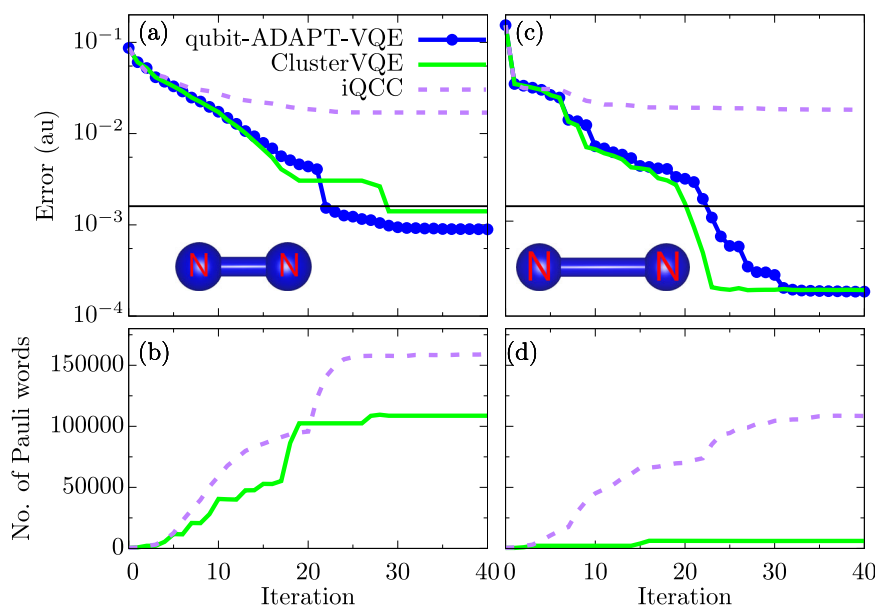

Fig. 4 N₂ simulation. a, c GS Energy errors and (**b, d**) number of Pauli words per each iteration for the N_2 molecule obtained from the qubit-ADAPT-VQE (blue), iQCC (purple) and ClusterVQE (green) methods. The N-N bond lengths are 1.09 Å (**a, b**) and 1.6 Å (**c, d**), respectively. The numbers of Pauli words in the qubit-ADAPT-VQE are 539 and 383 for 1.09 Å and 1.6 Å, respectively, which are not shown in (**b**) and (**d**).

N_2/STO-3G system is simulated after freezing the lowest two spin-orbitals.

Since the N_2 molecule has a stronger correlation, the QUCCSD ansatz (i.e., choosing only one Pauli word from each UCCSD excitation) could result in slow convergence[43]. Instead, adding one more adjoint Pauli word for each excitation (known as the QUCCSD2 ansatz with more details found in the SI) could significantly improve the accuracy. The results obtained from the QUCCSD and the QUCCSD2 ansatz are shown in Supplementary Fig. 4 and Fig. 4, respectively. Comparison of these figures reveals that QUCCSD2 significantly improves the convergence but at the same time increases the circuit depth and the Hamiltonian

size in the ClusterVQE and iQCC methods. The development of a more efficient ansatz that balances accuracy and efficiency for ClusterVQE could be a subject of separate investigations. As shown in Fig. 4(a) and (c), the energy convergence of ClusterVQE is almost the same as that of qubit-ADAPT-VQE and is better than that of iQCC. In contrast, the VQE optimization may lead to a small improvement after certain iterations due to the limited correlation that could be further introduced by one entangler. The convergence can be further improved by using more Pauli words for each UCCSD excitation[43]. Nevertheless, for a given ansatz, the size of the dressed Hamiltonian in ClusterVQE is much smaller than that of iQCC. It should be noted that the stretched N_2
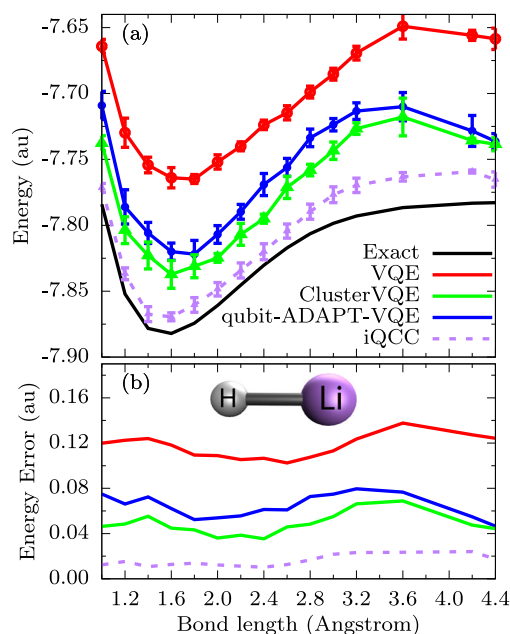
Fig. 5 PES of LiH molecule in the presence of noise. a Dissociation curves of LiH and (b) corresponding errors with exact diagonalization (black), VQE (red), qubit-ADAPT-VQE (blue), iQCC (purple), and ClusterVQE (green) methods on a noisy simulator. The custom noise in Qiskit[59] was used, where the error rates of single-qubit and two-qubit gates are 0.001 and 0.005, respectively. The means and error bars are calculated from 5 independent simulations.

Table 1. Circuit depth of each method with the QUCCSD ansatz. The LiH and N_2 molecules used in this table are in equilibrium configurations. The intermolecular distance of $(H_2)_2$ is 2 Å. The circuit depths of VQE with 1 layer QUCCSD ansatz are 352, 1535, and 331 for LiH, N_2, and $(H_2)_2$, respectively.

Molecule	Iteration	Circuit depth		
		ADAPT-VQE	iQCC	ClusterVQE
LiH	2	41	18	41
	4	73	16	58
	6	103	12	69
	8	117	8	83
	10	152	18	118
N_2	10	163	16	47
	20	349	22	134
	30	525	20	214
	40	701	16	246
	50	826	24	296
$(H_2)_2$	5	72	14	30
	10	144	14	45

molecule has stronger correlations (0.274 au) compared to the equilibrium one (0.113 au). However, accompanied by a better separation between clusters. Consequently, less dressing is required for the stretched N_2 to obtain an even larger amount of correlation and achieve better accuracy. Hence, ClusterVQE only slightly increases the size of the dressed Hamiltonian. In contrast, the size of the Hamiltonian exceeds 100,000 Pauli words in iQCC after 25 iterations. These results suggest that, even for strongly correlated systems, our ClusterVQE reduces circuit complexity without a significant increase in Hamiltonian size.

Noisy simulation and clustering on-the-fly

We next examine the dissociation curve of the LiH molecule in the presence of noise to demonstrate ClusterVQE's resilience to this common feature of all NISQ architectures. Figure 5 shows the dissociation curve of the LiH molecule calculated with different algorithms. The energy values are averaged over 5 runs. As shown by the curves, in the presence of noise, the dissociation curves obtained from the VQE, qubit-ADAPT-VQE, iQCC, and ClusterVQE methods are different from the exact solution. Since the VQE algorithm requires more parameters and the deepest circuit, the noise influence is the largest across the set, showing the strongest deviation from the reference. For example, one layer of the QUCCSD ansatz for the LiH (N_2) molecule requires 15 (92) parameters/entanglers. The qubit-ADAPT-VQE method, in contrast, is able to grow a compact ansatz with fewer parameters (the number of parameters in qubit-ADAPT-VQE, ClusterVQE, and iQCC is equal to the iteration number) and lower circuit depth as shown in Table 1. Consequently, the influence of noise on qubit-ADAPT-VQE is suppressed compared to VQE. On the other hand, the iQCC algorithm uses a much shallower circuit depth. As a result, iQCC is most immune to noise, and the corresponding dissociation curve is the closest to the reference. Next is ClusterVQE which partially dresses the Hamiltonian and grows the ansatz. The circuit depth of ClusterVQE lies between that of iQCC and qubit-ADAPT-VQE as shown by the comparison in Table 1. Consequently, the energy error in the presence of noise is also in between that of the iQCC and qubit-ADAPT-VQE methods.

We also compare the performance of the four algorithms with different error rates. Supplementary Fig. 1(a) and (b) compare the energy errors with respect to different single- and two-qubit error rates, respectively. The single-qubit gates usually have much lower error rates on real quantum devices. For example, the single- and two-qubit gate error rates of the IBM-Yorktown[63] device are around $\sim10^{-3}$ and $\sim10^{-2}$, respectively. Consequently, the error of single-qubit gates has a much smaller influence on the results. In contrast, the results are significantly affected by the error of two-qubit gates. In general, the error of ClusterVQE lies between that of the qubit-ADAPT-VQE and iQCC methods on the noisy simulator, as expected.

In the previous examples, the MI is calculated based on the exact wavefunction (WF) for proof-of-principle. But, the exact WF is not known before the VQE calculation. However, an approximate calculation of MI provides a sufficient estimate of the correlation between different qubits, as shown in the previous DMRG-type selection of active space[64] and our recent development of the PermVQE algorithm[41]. The approximate MI can be estimated on classical computers by using either MP2 or DMRG (with small bond-dimension) methods. Here, we also demonstrate that ClusterVQE even works with MI calculated on-the-fly. At each iteration, the MI is updated and used for re-clustering the qubits. As shown by the red dots in Fig. 1(a) and (c), the performance of ClusterVQE on-the-fly is almost the same as that using the exact MI. The clustering is suboptimal only for the first three iterations and the dressed Hamiltonian has a slightly larger size as a result. The composition of each cluster continues to change until a stable clustering is found. These results illustrate that ClusterVQE also works when clustering on-the-fly.

Simulation on a quantum computer and error mitigation

Finally, we implement ClusterVQE on an actual quantum backend. The IBMQ-Bogota device with 5 qubits was chosen to simulate the LiH molecule described by 10 qubits (orbital with the lowest energy is assumed to be frozen). In order to run on IBMQ-Bogota, we set the size of each cluster to be 5. Figure 6 shows the energies
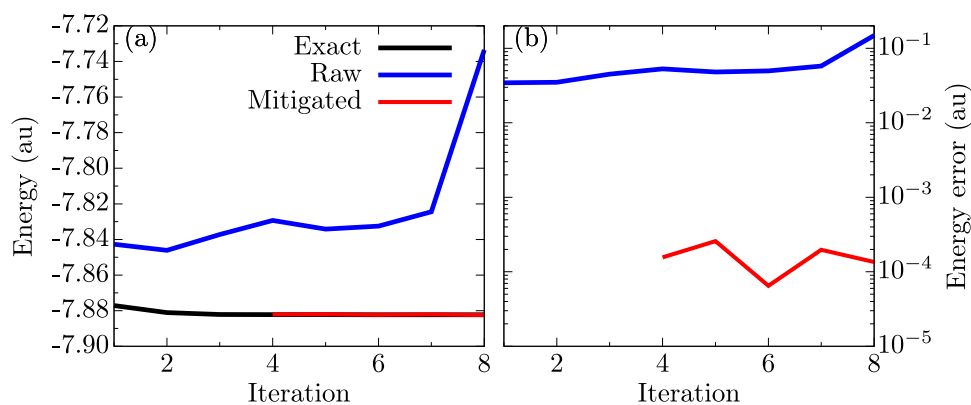
Fig. 6 Simulation of the LiH molecule (10 qubits) on the IBMQ-Bogota quantum device (5 qubits) with (red) and without (blue) error mitigation and their comparison to the exact solution (black). We do not show mitigated results for the first three iterations as Clifford data regression (CDR) gives perfect correction for these circuits since they contain at most one non-Clifford gate. The error rate of the CNOT gate is between $1.06 \times 10^{-2} \sim 2.02 \times 10^{-2}$ and 1.62×10^{-2} on average at runtime.

as a function of ClusterVQE iteration without error mitigation. Due to the large CNOT error rate, the energies go up in the first few ClusterVQE iterations without error mitigation. The number of entanglers encoded on the quantum circuits increases with the number of iterations and introduces more errors.

The ClusterVQE performance on real quantum devices can be further improved with recently developed error mitigation techniques[65,66]. For proof-of-principle, we first perform ClusterVQE on the noiseless simulator to generate 8 circuits of the first 8 iterations and run the circuits on the IBM quantum device (IBMQ-Bogota). We perform the mitigation using Clifford data regression (CDR)[67] on the IBMQ-Bogota device as well. The CDR training set is constructed using near-Clifford circuits with at most 1 non-Clifford gate in each of 2 clusters. The training set contains 3–4 near-Clifford circuits depending on the circuit of interest. The training circuits are chosen by replacing non-Clifford gates with the closest Clifford gates. The results are shown in Fig. 6 in which the blue and red lines plot the raw results obtained from the IBM-Bogota and mitigated results, respectively. We do not show mitigated results for the first three iterations as CDR gives a perfect correction for these circuits since they contain at most one non-Clifford gate. Figure 6 shows that error mitigation reduces the error by several orders of magnitude. Such simulations demonstrate that the ClusterVQE method, combined with error mitigation, can readily achieve "chemical accuracy" on real quantum devices. It should be noted that the term "chemical accuracy" used throughout this manuscript refers to the relative accuracy compared to the exact (FCI) solution of the given basis set, which is not the true chemical accuracy. The rigorous chemical accuracy can only be achieved by improving the theoretical complexity (HF, CCSD, CCSDT, …, FCI, etc) and basis set size. However, increasing the basis set size in quantum algorithms is not NISQ friendly, significantly increasing the needed quantum resources. Alternatively, by using a trans-correlated Hamiltonian, one can achieve the accuracy at the cc-pVTZ basis even with a minimal basis set[62].

Complexity and scalability

Even though ClusterVQE reduces the circuit complexity (circuit depth and number of qubits), the computational complexity is increased due to the dressing. As shown in the previous section, ClusterVQE has almost the same performance as qubit-ADAPT-VQE in terms of energy errors. But, more Pauli words are measured due to the dressing. The computational scaling of ClusterVQE is dependent on the number of dressing operators. That being said, the scaling is determined by the amount of inter-cluster correlation. Even though the scaling of computing the dressed Hamiltonian is exponential against the number of dressing

operators[68], the scaling of ClusterVQE against iterations is not exponential in general since not every entangler is used to dress the Hamiltonian. For the best cases, only a limited number of entanglers is used to dress the Hamiltonian (LiH and $(H_2)_2$ complex with a large intermolecular distance in section "The ClusterVQE algorithm" for example). Such a situation can be intuitively expected in molecular aggregate systems, where the intermolecular coupling is usually smaller than the intramolecular correlation. For the worse cases (usually in strongly correlated systems in which qubit clustering is not optimal), the scaling of ClusterVQE may be exponential if a larger number of entanglers are used to dress the Hamiltonian ($(H_2)_2$ complex with large intermolecular distance 2 Å for example). However, even in strongly correlated molecules, the scaling of different geometries can be different. For example, ClusterVQE simulation of N_2 at equilibrium with 12-qubits ends up with a large number of Pauli words (Fig. 4(b)). However, the simulation of N_2 with a stretched bond has much better scaling (Fig. 4(d)) even though the stretched configuration has a stronger correlation. That being said, the bottleneck of our approach relies on whether the inter-cluster MI can be efficiently minimized. In summary, ClusterVQE does not introduce exponential scaling for general cases and the complexity of ClusterVQE increases with increasing inter-cluster correlations.

Discussion

We have developed a ClusterVQE algorithm to reduce quantum circuit complexity, including both the number of qubits and circuit depth. Our algorithm groups qubits into clusters based on the maximal correlation between them and further takes into account the residual entanglement between clusters in a "dressed" Hamiltonian. ClusterVQE naturally implements the benefits of the previously developed PermVQE algorithm, which significantly reduces the number of controlled-NOT gates to connect strongly correlated qubits[41]. The entanglers that couple different clusters are further used to dress the Hamiltonian and remove the entanglement between different clusters. Consequently, the Hamiltonian can be decomposed as a tensor product of each cluster. Each cluster can then be measured separately with a smaller number of qubits and shallower circuit depth. Therefore, ClusterVQE is a particularly promising algorithmic development for quantum chemistry applications, where it is able to reduce the quantum hardware requirements significantly. Moreover, our algorithm has the flexibility of reducing qubits to any number at the expense of additional computational costs on classical computers. Since fewer entanglers are encoded on quantum circuits, ClusterVQE is more robust to noise than qubit-ADAPT-VQE,

being one of the most circuit-efficient VQE algorithms, which is verified by our numerical experiments on both the noisy simulator and real quantum platforms. Since the correlation between different clusters is minimized, only a few operators are used to dress the Hamiltonian. Consequently, the Hamiltonian size does not increase significantly compared to that of the iQCC method.

Finally, it should be noted that cluster VQE uses a newly proposed gradient measurement method for VQE without using an ancillary qubit. The analytical gradient can significantly improve the performance of ClusterVQE in noisy environments. Simulation of the LiH molecule on a noisy quantum simulator showed that the energy error is smaller than that of the qubit-ADAPT-VQE method due to the shallower circuit depth. A groundbreaking example of the exact error-corrected LiH molecule simulation described by 10 qubits on the 5-qubits IBMQ-Bogota quantum computer with ClusterVQE is demonstrated. Results further demonstrate the validity and advantage of the ClusterVQE algorithm. We believe the algorithm developed in this work can pave the way toward the adaptation of molecular quantum chemistry simulation on NISQ devices. It should be mentioned that since the size of the qubit Hamiltonian increases with the number of operators used to dress the Hamiltonian, it increases both computational costs on classical computers and measurements on quantum computers. Recently developed measurement reduction techniques[69–75] are likely a path forward to further reduce the simulation cost on quantum devices.

METHODS

Mutual information based clustering of qubits

The first essential step of ClusterVQE is to split the qubits into different clusters and minimize the correlations between different clusters. With this aim, an entanglement map based on the mutual information (MI) that reflects the electronic correlations among all pairs of qubits is calculated[41]. Previous results obtained for the orbital ordering problem in the Density Matrix Renormalization Group (DMRG) method in classical quantum chemistry calculations[57] showed that the MI is a reliable parameter to quantify the correlation between two quantum particles. Our previous work also demonstrated that an approximate MI is sufficient for the qubit re-arrangement[41] for the reduction of circuit depth. The MI has also been used to construct a compact entangler pool for ADAPT-VQE[76]. According to quantum information theory, the MI between qubits i and j is defined as follows[57]:

$$I_{ij} = \frac{1}{2}(S_i + S_j - S_{ij})(1 - \delta_{ij}), \tag{1}$$

where S_i and S_{ij} are the single- and two-qubit Von Neumann entropies, respectively, and δ_{ij} is the Kronecker delta. S_i and S_{ij} are obtained from the corresponding one- and two-body density matrix[41].

Once the MI is obtained, it can be used to cluster the qubits into different groups by minimizing the MI between different clusters which is defined as the summation of MIs between inter-cluster qubits. The qubit clustering can be achieved by using a classical graph partitioning method, such as the graph partition function of the Metis library[77] (Supplementary Section I(A)). Alternatively, this can be mapped into a quadratic unconstrained binary optimization (QUBO) problem and solved on a quantum annealer by either the MI-based clustering method (details are given in Supplementary Section I(B)) or the quantum community detection method (QCD)[78] (details are given in Supplementary Section I(C)). Problems with up to 64 variables or nodes can be solved directly on the D-Wave 2000Q quantum annealer. Larger sizes require a quantum-classical approach using *qbsolv*[79] or alternative solvers[80] + D-Wave. We have chosen to use quantum graph partitioning/clustering methods since they produce comparable results to classical methods[81,82]. Classical partitioning/clustering methods could also be used as an alternative.

VQE, qubit-ADAPT-VQE, and iQCC algorithms

Within the VQE algorithm, a parameterized ansatz $|\Phi\rangle = \hat{U}(\theta)|\Phi_r\rangle$ is suggested and encoded on the quantum circuit, where $|\Phi_r\rangle$ is the reference state. In this work, the mean-field Hartree-Fock (HF) state is used as a reference. The unitary operator $\hat{U}(\theta)$ introduces the correlations between qubits. Since current quantum gates can only operate on a few qubits at a time, $\hat{U}(\theta)$ is decomposed as a tensor product of many unitary operators, denoted as entanglers, with each entangler acting on a few qubits, i.e., $\hat{U}(\theta) = \prod_j^{\mathcal{D}} \hat{U}_j$ where $\hat{U}_j = e^{i\theta_j \hat{A}_j}$ are the entanglers and $\mathcal{D}$ is the number of entanglers in the ansatz. In the widely used UCCSD ansatz, the $\{\hat{A}_j\}$ can be readily obtained by mapping the single/double Fermionic excitation operators into qubit-space as shown in Supplementary Section II. With the ansatz, the ground-state (GS) energy of a given Hamiltonian can be obtained by imposing the variational principle within the VQE algorithm,

$$E_{GS} = \min_\theta \langle \Phi_r | \hat{U}^\dagger(\theta)\hat{H}\hat{U}(\theta)|\Phi_r\rangle \equiv \min_\theta f(\theta), \tag{2}$$

where $\hat{H} = \sum_k a_k \hat{P}_k$ qubit form of the electronic Hamiltonian, which is obtained by one of the fermion-to-qubit transformations. a_k are the coefficients and $\hat{P}_k$ are the Pauli words. Hence, the expectation value (or cost function) in Eq. (2) is further decomposed as

$$f(\theta) = \sum_k a_k \langle \Phi_r | \hat{U}^\dagger(\theta)\hat{P}_K \hat{U}(\theta)|\Phi_r\rangle \equiv \sum_k a_k f_k(\theta) \tag{3}$$

and each term ($f_k(\theta)$) is measured on a quantum circuit.

However, the contribution of each entangler to the GS energy is different. Thus, the UCCSD ansatz can be made even more compact by the adaptive construction[42,43] in the qubit-ADAPT-VQE, at the cost of more VQE calculations. At the n-th iteration of qubit-ADAPT-VQE, the ADAPT ansatz is

$$|\Phi_n^{ADAPT}(\theta)\rangle = \prod_{j=1}^n e^{i\theta_j \hat{P}_j}|\Phi_r\rangle. \tag{4}$$

Where the operators ($\hat{P}_j$) are selected from an operator pool $\{\hat{A}_j\}$. Within the qubit-ADAPT-VQE algorithm, the sequence of operators is selected based on their contributions to the energy weighted by the gradients with respect to the corresponding parameters[42],

$$\frac{\partial E}{\partial \theta_j} = \langle \Phi^{ADAPT}(\theta)|[\hat{H}, \hat{P}_j]|\Phi^{ADAPT}(\theta)\rangle. \tag{5}$$

The gradients can be readily measured on quantum computers. Once the gradients of all operators are measured, the operator with the largest gradient ($\frac{\partial E}{\partial \theta_j}$) is selected into the ansatz and the ansatz keeps growing with each iteration until convergence.

Because $\langle \Phi_r | \hat{U}^\dagger(\theta)\hat{H}\hat{U}(\theta)|\Phi_r\rangle = \langle \Phi_r | \hat{H}_d |\Phi_r\rangle$ where $\hat{H}_d = \hat{U}^\dagger(\theta)\hat{H}\hat{U}(\theta)$, the operator can be folded into a dressed Hamiltonian. Thus, instead of growing the ansatz, the iQCC algorithm grows the Hamiltonian[68]. After each iteration, the operators and corresponding parameters are used to dress the Hamiltonian,

$$\hat{H}_j = e^{-i\theta_j \hat{P}_j}\hat{H}_{j-1}e^{i\theta_j \hat{P}_j}. \tag{6}$$

where $\hat{H}_j$ is the dressed Hamiltonian of the j-th iteration.

The ClusterVQE algorithm

According to the qubit clustering introduced in Sec. 3.1, the entanglers used in the ansatz can be grouped into different sets of intracluster and intercluster entanglers,

$$\hat{U}(\theta) = \{\hat{U}_{c_1}, \cdots, \hat{U}_{c_M}, \hat{U}_{c_{12}}, \cdots, \hat{U}_{c_{M-1,M}}\} \tag{7}$$

where M is the number of clusters. In this way, the ordering of entanglers in $\hat{U}(\theta)$ is changed and the corresponding parameters can be different as well. The entangler set $\{\hat{U}_{c_i}\}$ only acts on the corresponding cluster i, and $\{\hat{U}_{c_{ij}}\}$ represents the entanglers that couple clusters i and j. Consequently, the cost function in Eq. (2) can be rewritten as

$$f(\theta) = \langle \Phi_r | \prod_i \hat{U}_{c_i}^\dagger \left[\prod_{i \neq j} \hat{U}_{c_{ij}}^\dagger \hat{H}\hat{U}_{c_{ij}}\right] \hat{U}_{c_i}|\Phi_r\rangle$$
$$= \langle \Phi_r | \prod_i \hat{U}_{c_i}^\dagger \hat{H}_d \hat{U}_{c_i}|\Phi_r\rangle \tag{8}$$

where $\hat{H}_d \equiv \prod_{i \neq j} \hat{U}_{c_{ij}}^\dagger \hat{H}\hat{U}_{c_{ij}}$ is the dressed Hamiltonian. Because the $\hat{H}_d$ can be represented in terms of Pauli words where each Pauli word can be trivially rewritten as the tensor product of different clusters, i.e., $\hat{H}_d = \sum_k a_k \hat{P}_k = \sum_k a_k \prod_i \hat{P}_k(c_i)$ and the initial state can be decomposed as
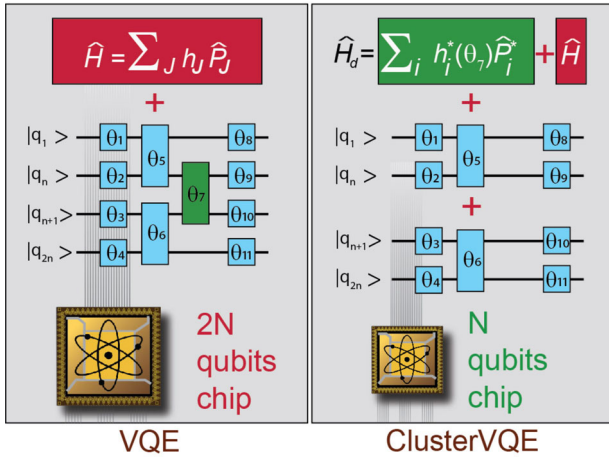
Fig. 7 Finding the lowest eigenvalue of a given qubit Hamiltonian
$H = \sum_J P_J$ **with VQE and ClusterVQE, where 2N is the number of qubits.** VQE uses a quantum circuit defined on 2N qubits. In contrast, ClusterVQE splits the original 2N-qubit circuit into 2 N-qubit circuits by removing the entanglement between them. Here, shorter circuits are achieved at the expense of a dressed Hamiltonian $\hat{H}_d$.
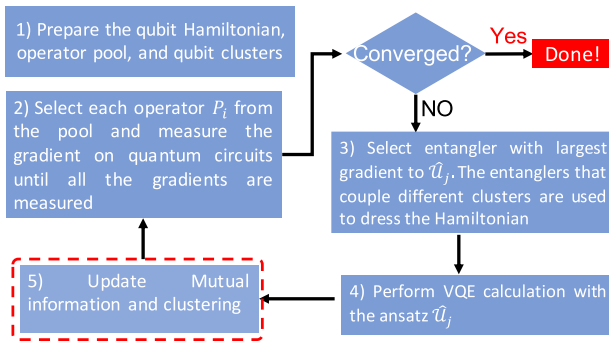

Fig. 8 Flowchart of the ClusterVQE algorithm. The fifth step in the red-dashed box is skipped if MI is provided.

$\prod_i \Phi_r(c_i)$. The energy in Eq. (8) can be further rewritten as

$$E_{GS} = \min_\theta \sum_k a_k \left[\prod_i \langle \Phi_r(c_i) | \hat{U}_{c_i}^\dagger \hat{P}_k(c_i) \hat{U}_{c_i} | \Phi_r(c_i) \rangle \right]. \qquad (9)$$

Here each expectation value $\langle \Phi_r(c_i) | \hat{U}_{c_i}^\dagger \hat{P}_k(c_i) \hat{U}_{c_i} | \Phi_r(c_i) \rangle$ only involves the qubits in the cluster i. Hence, the energy of a larger problem can be measured on M smaller quantum circuits after Hamiltonian dressing, as shown by the schematic diagram in Fig. 7. In practice, like the qubit-ADAPT-VQE algorithm, ClusterVQE adaptively grows the ansatz in Eq. (8). The ansatz used in the jth iteration is denoted as $\hat{\mathcal{U}}_j$. However, in contrast to the growing ansatz in qubit-ADAPT-VQE[42] and growing Hamiltonian in iQCC[44], the ansatz in ClusterVQE is decomposed onto smaller quantum circuits, and only the entanglers between different clusters are used to dress the Hamiltonian. Consequently, the iterative construction breaks the original VQE algorithm with deeper and wider quantum circuits into a series of quantum circuits with a shallower depth and a smaller number of qubits, as illustrated by Fig. 7.

The entangler $\hat{U}_k$ chosen in the ansatz $\hat{\mathcal{U}}_j$ for the jth iteration is determined by its contribution to the energy weighted by the gradient with respect to the parameter, i.e., the $\hat{U}_k$ terms with largest gradients $\frac{\partial E}{\partial \theta_k}$ are selected[42], where the gradient is defined as

$$
\begin{aligned}
\frac{\partial E}{\partial \theta_k} &= \frac{\partial}{\partial \theta_k} \langle \Phi_r | \hat{U}_k^\dagger(\theta_k) \hat{H}_{d,j-1} \hat{U}_k(\theta_k) | \Phi_r \rangle |_{\theta_k=0} \\
&= i \langle \Phi_r | \hat{U}_k^\dagger(\theta_k) [\hat{H}_{d,j-1}, \hat{P}_k] \hat{U}_k(\theta_k) | \Phi_r \rangle.
\end{aligned}
\qquad (10)
$$

Where $\hat{H}_{d,j-1}$ is the dressed Hamiltonian of the $j-1$ iteration. It should be noted that $\hat{H}_{d,j}$ may be the same as the $\hat{H}_{d,j-1}$ if the selected entangler of

the jth iteration is not used to dress the Hamiltonian. Since the measurement of each gradient term is independent, all the gradients can be measured in parallel with several uncoupled quantum computers[42].

The flowchart of the ClusterVQE scheme is presented in Fig. 8. In the first step, the qubit Hamiltonian and operator pool are prepared: (a) Provided classical mean-field calculation of the molecule, map the second quantized Hamiltonian into its qubit counterpart by using either Jordan-Wigner (JW)[60], Bravyi-Kitaev (BK)[83], or parity[48,84] encoding. (b) Generate the UCCSD excitation operators and map them into Pauli words. Here only Pauli words with the different flip indices (Supplementary Section II) are chosen and added to the excitation pool $\{\hat{A}_i\}$. In addition, clustering of qubits is performed if MI is provided. Otherwise, the qubits in the first iteration are simply grouped according to their spin index. If the largest gradient in the second step is smaller than the convergence criteria ϵ and the energy minimization criteria are satisfied, the algorithm exits from the iterative loop. The fifth step is only triggered if clustering on-the-fly is desired.

Since some of the entanglers in $\hat{\mathcal{U}}_j$ are used to dress the Hamiltonian, the circuit depth is smaller than that of qubit-ADAPT-VQE as shown by Table 1. As a result, ClusterVQE is more robust to noise compared to qubit-ADAPT-VQE, as shown and discussed later, making the ClusterVQE algorithm more suitable for NISQ devices than qubit-ADAPT-VQE. Besides, because the clustering minimizes the entanglement between different clusters, the number of dressing entanglers is much smaller than that of iQCC. Consequently, ClusterVQE can balance the computational costs between the Quantum Processing Unit (QPU) and CPU. The exponential scaling in the growing of the Hamiltonian in iQCC is avoided. Most importantly, the circuit depth/width is further reduced to qubit-ADAPT-VQE by splitting the measurement into cluster-wise measurements (Eq. (8)). Thus, compared to previous approaches, ClusterVQE reduces the circuit complexity (depth and width) and can solve larger problems on smaller quantum circuits.

It should be noted that even though the ClusterVQE method can reduce the number of qubits and ansatz complexity (depth and width) by distributing the measurements on smaller quantum circuits, the Hamiltonian size increases exponentially with the number of dressing entanglers. Consequently, the number of measurements is increased according to Eq. (3). Future work will consider measurement reduction techniques[69–75] to reduce the number of measurements required. Among the many measurement reduction techniques, the classical shadow approach is a potential candidate for reducing the number of measurements for ClusterVQE[72,85].

Analytical gradients for VQE optimization

In order to improve the algorithm performance in the presence of noise, an analytical gradient is essential for classical optimization. It should be noted that the gradient for VQE optimization is different from the gradients in Eq. (10) which are used to choose entanglers for each iteration. Ref. [86] suggests calculation of the analytical gradients by performing a measurement on an ancillary qubit[35,87]. An analytical approach has been developed in this work for the analytical gradient measurements but without introducing an ancillary qubit. The gradient of energy with respect to the variational parameter reads as follows.

$$\frac{\partial E}{\partial \theta_i} = 2\mathrm{Im}\left\{ \langle \Phi_r | \hat{U}^\dagger \hat{H} \hat{P}_i^{\mathrm{eff}} \hat{U} | \Phi_r \rangle \right\} \qquad (11)$$

where $\hat{P}_i^{\mathrm{eff}}$ is the dressed Pauli word as shown in Supplementary Section III. Therefore, the energy gradient can be measured in the same way as the energy measurement by replacing $\hat{H}$ with $\hat{H}\hat{P}_j^{\mathrm{eff}}$ and no ancillary qubit is required. Moreover, a related approach without using an ancillary qubit was recently proposed as well within the unitary block optimization scheme[88]. On a noiseless simulator, VQE calculations with numerical gradients or analytical gradients give the same results. However, analytical gradients are much more stable on noisy simulators and quantum computers.

DATA AVAILABILITY

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

CODE AVAILABILITY

The code that supports the findings of this study are available from the authors upon request.

REFERENCES

1. Gan, Z. & Harrison, R. J. Calibrating quantum chemistry: A multi-teraflop, parallel-vector, full-configuration interaction program for the cray-x1, in *SC '05: Proceedings of the 2005 ACM/IEEE Conference on Supercomputing* (2005) pp. 22–22.
2. Chan, G. K.-L. & Sharma, S. The density matrix renormalization group in quantum chemistry. *Ann. Rev. Phys. Chem.* **62**, 465–481 (2011).
3. Tubman, N. M. et al. Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling ci method. *J. Chem. Theory Comput.* **16**, 2139–2159 (2020).
4. Lyakh, D. I., Musiał, M., Lotrich, V. F. & Bartlett, R. J. Multireference nature of chemistry: The coupled-cluster view. *Chem. Rev.* **112**, 182–243 (2012).
5. Cao, Y. et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).
6. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. *Rev. Mod. Phys.* **92**, 015003 (2020).
7. Kitaev, A. Y. Quantum measurements and the abelian stabilizer problem. Preprint at arXiv https://arxiv.org/abs/quant-ph/9511026 (1995).
8. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
9. Preskill, J. Quantum computing in the nisq era and beyond. *Quantum* **2**, 79 (2018).
10. Wei, S., Li, H. & Long, GuiLu A full quantum eigensolver for quantum chemistry simulations. *Research* **2020**, 1486935 (2020).
11. Moll, N. et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Sci. Technol.* **3**, 030503 (2018).
12. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021).
13. Bharti, K. et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **94**, 015004 (2022).
14. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
15. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 023023 (2016).
16. Li, Y., Hu, J., Zhang, X.-M., Song, Z. & Yung, M.-H. Variational quantum simulation for quantum chemistry. *Adv. Theory Simul.* **2**, 1800182 (2019).
17. Fedorov, D. A., Peng, B., Govind, N. & Alexeev, Y. VQE method: A short survey and recent developments. *Mat. Theory* **6**, 2 (2022).
18. Shen, Y. et al. Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. *Phys. Rev. A* **95**, 020501 (2017).
19. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308 (2017).
20. Nakanishi, K. M., Mitarai, K. & Fujii, K. Subspace-search variational quantum eigensolver for excited states. *Phys. Rev. Res.* **1**, 033062 (2019).
21. Higgott, O., Wang, D. & Brierley, S. Variational quantum computation of excited states. *Quantum* **3**, 156 (2019).
22. Kawai, H. & Nakagawa, Y. O. Predicting excited states from ground state wavefunction by supervised quantum machine learning. *Mach. Learn.* **1**, 045027 (2020).
23. Greene-Diniz, G. & Muñoz Ramo, D. Generalized unitary coupled cluster excitations for multireference molecular states optimized by the variational quantum eigensolver. *Int. J. Quantum Chem.* **121**, e26352 (2021).
24. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).
25. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
26. Nam, Y. et al. Ground-state energy estimation of the water molecule on a trapped-ion quantum computer. *npj Quantum Info.* **6**, 33 (2020).
27. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 4812 (2018).
28. Wang, S. et al. Noise-induced barren plateaus in variational quantum algorithms. *Nat. Commun.* **12**, 6961 (2021).
29. Cerezo, M., Sone, A., Volkoff, T., Cincio, L. & Coles, P. J. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat. Commun.* **12**, 1791 (2021).
30. Abbas, A. et al. The power of quantum neural networks. *Nat. Comput. Sci.* **1**, 403–409 (2021).
31. Bartlett, R. J., Kucharski, S. A. & Noga, J. Alternative coupled-cluster ansätze ii. the unitary coupled-cluster method. *Chem. Phys. Lett.* **155**, 133–140 (1989).
32. Bartlett, R. J. & Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **79**, 291–352 (2007).
33. Seeley, J. T., Richard, M. J. & Love, P. J. The bravyi-kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.* **137**, 224109 (2012).
34. Barkoutsos, P. et al. Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wave-function expansions. *Phys. Rev. A* **98**, 022322 (2018).
35. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **4**, 14008 (2018).
36. Kühn, M., Zanker, S., Deglmann, P., Marthaler, M. & Weiß, H. Accuracy and resource estimations for quantum chemistry on a near-term quantum computer. *J. Chem. Theory Comput.* **15**, 4764–4780 (2019).
37. Ryabinkin, I. G., Yen, Tzu-Ching, Genin, S. N. & Izmaylov, A. F. Qubit coupled cluster method: A systematic approach to quantum chemistry on a quantum computer. *J. Chem. Theory Comput.* **14**, 6317–6326 (2018).
38. Lee, J., Huggins, W. J., Head-Gordon, M. & Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **15**, 311–324 (2019).
39. Matsuzawa, Y. & Kurashige, Y. Jastrow-type decomposition in quantum chemistry for low-depth quantum circuits. *J. Chem. Theory Comput.* **16**, 944–952 (2020).
40. Xia, R. & Kais, S. Qubit coupled cluster singles and doubles variational quantum eigensolver ansatz for electronic structure calculations. *Quantum Sci. Tech.* **6**, 015001 (2020).
41. Tkachenko, N. V. et al. Correlation-informed permutation of qubits for reducing ansatz depth in the variational quantum eigensolver. *PRX Quantum* **2**, 020337 (2021).
42. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* **10**, 3007 (2019).
43. Tang, H. L. et al. Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. *PRX Quantum* **2**, 020310 (2021).
44. Ryabinkin, I. G., Lang, R. A., Genin, S. N. & Izmaylov, A. F. Iterative qubit coupled cluster approach with efficient screening of generators. *J. Chem. Theory Comput.* **16**, 1055–1063 (2020).
45. Lang, R. A., Ryabinkin, I. G. & Izmaylov, A. F. Unitary transformation of the electronic hamiltonian with an exact quadratic truncation of the baker-campbell-hausdorff expansion. *J. Chem. Theory Comput.* **17**, 66–78 (2021).
46. Ryabinkin, I. G., Izmaylov, A. F. & Genin, S. N. A posteriori corrections to the iterative qubit coupled cluster method to minimize the use of quantum resources in large-scale calculations. *Quantum Sci. Tech.* **6**, 024012 (2021).
47. Yordanov, Y. S., Armaos, V., Barnes, CrispinH. W. & Arvidsson-Shukur, DavidR. M. Qubit-excitation-based adaptive variational quantum eigensolver. *Commun. Phys.* **4**, 228 (2021).
48. Bravyi, S., Gambetta, J. M. Mezzacapo, A. & Temme, K. Tapering off qubits to simulate fermionic hamiltonians. Preprint at arXiv https://arxiv.org/abs/1701.08213 (2017).
49. Setia, K. et al. Reducing qubit requirements for quantum simulations using molecular point group symmetries. *J. Chem. Theory Comput.* **16**, 6091–6097 (2020).
50. Yen, Tzu-Ching, Lang, R. A. & Izmaylov, A. F. Exact and approximate symmetry projectors for the electronic structure problem on a quantum computer. *J. Chem. Phys.* **151**, 164111 (2019).
51. Zhang, F. et al. Shallow-circuit variational quantum eigensolver based on symmetry-inspired hilbert space partitioning for quantum chemical calculations. *Phys. Rev. Res.* **3**, 013039 (2021).
52. Barratt, F. et al. Parallel quantum simulation of large systems on small nisq computers. *npj Quantum Info.* **7**, 79 (2021).
53. Fujii, K., Mizuta, K., Ueda, K., Mitarai, K., Mizukami, W. & Nakagawa, Y. O. Deep variational quantum eigensolver: a divide-and-conquer method for solving a larger problem with smaller size quantum computers. *PRX Quantum* **3**, 010346 (2022).
54. Yuan, X., Sun, J., Liu, J., Zhao, Q. & Zhou, Y. Quantum simulation with hybrid tensor networks. *Phys. Rev. Lett.* **127**, 040501 (2022).
55. Eddins, A. et al. Doubling the size of quantum simulators by entanglement forging. *PRX Quantum* **3**, 010309 (2022).
56. Metcalf, M., Bauman, N. P., Kowalski, K. & de Jong, W. A. Resource efficient chemistry on quantum computers with the variational quantum eigensolver and the double unitary coupled-cluster approach. *J. Chem. Theory Comput.* **16**, 6165–6175 (2022).
57. Rissler, J., Noack, R. M. & White, S. R. Measuring orbital interaction using quantum information theory. *Chem. Phys* **323**, 519–531 (2006).

58. Huang, Z. & Kais, S. Entanglement as measure of electron-electron correlation in quantum chemistry calculations. *Chem. Phys. Lett.* **413**, 1–5 (2005).

59. Héctor, A. et al. Qiskit: An open-source framework for quantum computing, https://github.com/Qiskit/qiskit/blob/master/Qiskit.bib (2019).

60. Jordan, P. & Wigner, E. Über das paulische Äquivalenzverbot. *Zeitschrift für Physik* **47**, 631–651 (1928).

61. Morales, L. J. & Nocedal, J. English (US) Remark on "algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization", *ACM Trans. Math. Softw.* **38**, 1–4 (2011).

62. Kumar, A. et al. Accurate quantum simulation of molecular ground and excited states with a transcorrelated hamiltonian. Preprint at https://arxiv.org/abs/2201.09852 (2022).

63. Ibm quantum experience webpage, https://quantum-computing.ibm.com.

64. Stein, C. J. & Reiher, M. Automated selection of active orbital spaces. *J. Chem. Theory Comput.* **12**, 1760–1771 (2016).

65. Cincio, L., Rudinger, K., Sarovar, M. & Coles, P. J. Machine learning of noise-resilient quantum circuits. *PRX Quantum* **2**, 010324 (2021).

66. Endo, S., Cai, Z., Benjamin, S. C. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Jpn.* **90**, 032001 (2021).

67. Czarnik, P., Arrasmith, A., Coles, P. J. & Cincio, L. Error mitigation with Clifford quantum-circuit data. *Quantum* **5**, 592 (2021).

68. Ryabinkin, I. G., Izmaylov, A. F. & Genin, S. N. A posteriori corrections to the iterative qubit coupled cluster method to minimize the use of quantum resources in large-scale calculations. *Quantum Sci. Tech.* **6**, 024012 (2021).

69. Bonet-Monroig, X., Babbush, R. & O'Brien, T. E. Nearly optimal measurement scheduling for partial tomography of quantum states. *Phys. Rev. X* **10**, 031064 (2020).

70. Yen, T.-C. & Izmaylov, A. F. Cartan sub-algebra approach to efficient measurements of quantum observables. *PRX Quantum* **2**, 040320 (2021).

71. Zhao, A. et al. Measurement reduction in variational quantum algorithms. *Phys. Rev. A* **101**, 062322 (2020).

72. Huang, H.-Y., Kueng R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.*, 1050–1057. https://www.nature.com/articles/s41567-020-0932-7 (2020).

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

74. Izmaylov, A. F., Yen, T.-C., Lang, R. A. & Verteletskyi, V. Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method. *J. Chem. Theory Comput.* https://doi.org/10.1021/acs.jctc.9b00791 (2019).

75. Huggins, W. J. et al. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. *npj Quantum Info.* **7**, 23 (2021).

76. Zhang, Z.-J., Kyaw, T. H., Kottmann, J., Degroote, M. & Aspuru-Guzik, A. Mutual information-assisted adaptive variational quantum eigensolver. *Quantum Sci.-Tech.* https://doi.org/10.1088/2058-9565/abdca4 (2021).

77. https://metis.readthedocs.io/en/latest/ Metis for python (2021).

78. Mniszewski, S. M. et al. Reduction of the molecular hamiltonian matrix using quantum community detection. *Sci. Rep.* **11**, 4099 (2021).

79. Booth, M., Reinhardt, S. P. & Roy, A. *Partitioning Optimization Problems for Hybrid Classical/Quantum Execution*, Tech. Rep. (2017).

80. Teplukhin, A., Kendrick, B. K., Mniszewski, S. M., Tretiak, S. & Dub, P. A. Sampling electronic structure quadratic unconstrained binary optimization problems (qubos) with ocean and mukai solvers. *PLOS ONE* **17**, 1–8 (2022).

81. Ushijima-Mwesigwa, H., Negre, C. F. A. & Mniszewski, S. M. Graph partitioning using quantum annealing on the D-Wave system, in *Proceedings of the Second International Workshop on Post Moores Era Supercomputing (PMES'17)* (ACM New York, NY, USA, 2017) pp. 22–29.

82. Negre, C. F. A., Ushijima-Mwesigwa, H. & Mniszewski, S. M. Detecting multiple communities using quantum annealing on the D-Wave system. *PLOS ONE* **15**, e0227538 (2020).

83. Bravyi, S. B. & Kitaev, A. Y. Fermionic quantum computation. *Ann. Phys.* **298**, 210–226 (2002).

84. Seeley, J. T., Richard, M. J. & Love, P. J. The bravyi-kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.* **137**, 224109 (2012).

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

86. Li, Y., Hu, J., Zhang, Xiao-Ming, Song, Z. & Yung, M.-H. Variational quantum simulation for quantum chemistry. *Adv. Theory and Simul.* **2**, 1800182 (2019).

87. Dallaire-Demers, P.-L., Romero, J., Veis, L., Sim, S. & Aspuru-Guzik, A. Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer. *Quantum Sci. Technol.* **4**, 045005 (2019).

88. Slattery, L., Villalonga, B. & Clark, B. K. Unitary block optimization for variational quantum algorithms. *Phys. Rev. Res.* **4**, 023072 (2022).

AUTHOR CONTRIBUTIONS

Y.Z., L.C., and P.A.D. conceived the project. Y.Z. wrote the code of ClusterVQE and performed the calculations. C.F.A.N. and S.M.M. wrote the code for qubit clustering. L.C., P.C., and P.J.C. conducted the error mitigation. All authors discussed the algorithm development and analyzed the results. All authors wrote the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-022-00599-z.

Correspondence and requests for materials should be addressed to Yu Zhang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.