Accurate harmonic vibrational frequencies for diatomic molecules via quantum computing

Shih-Kai Chou,[1] Jyh-Pin Chou,[2,3] Alice Hu,[4,5] Yuan-Chung Cheng,[6,7,3,*] and Hsi-Sheng Goan[1,7,3,†]

[1]*Department of Physics and Center for Theoretical Physics, National Taiwan University, Taipei 10617, Taiwan*
[2]*Department of Physics, National Changhua University of Education, Changhua 50007, Taiwan*
[3]*Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan*
[4]*Department of Mechanical Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR 999077, China*
[5]*Department of Materials Science and Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR 999077, China*
[6]*Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan*
[7]*Center for Quantum Science and Engineering, National Taiwan University, Taipei 10617, Taiwan*

(Received 4 June 2023; accepted 6 November 2023; published 8 December 2023)

During the noisy intermediate-scale quantum (NISQ) era, quantum computational approaches refined to overcome the challenge of limited quantum resources are highly valuable. A comprehensive benchmark for a quantum computational approach in this spirit could provide insights toward further improvements. On the other hand, the accuracy of the molecular properties predicted by most of the quantum computations nowadays is still far off (not within chemical accuracy) compared to their corresponding experimental data. In this work, we propose a promising qubit-efficient quantum computational approach and present a comprehensive investigation by benchmarking quantum computation of the harmonic vibrational frequencies of a large set of neutral closed-shell diatomic molecules with results in great agreement with their experimental data. To this end, we construct the accurate Hamiltonian using molecular orbitals, derived from density functional theory to account for the electron correlation and expanded in the Daubechies wavelet basis set to allow an accurate representation in real space grid points, where an optimized compact active space is further selected so that only a reduced small number of qubits is sufficient to yield an accurate result. Typically, calculations achieved with 2 to 12 qubits using our approach would need 20 to 60 qubits using a traditional cc-pVDZ basis set with frozen core approximation to achieve similar accuracy. To justify the approach, we benchmark the performance of the Hamiltonians spanned by the selected molecular orbitals by first transforming the molecular Hamiltonians into qubit Hamiltonians and then using the exact diagonalization method to calculate the results, regarded as the best results achievable by quantum computation to compare to the experimental data. Furthermore, using the variational quantum eigensolver algorithm with the constructed qubit Hamiltonians, we show that the variational quantum circuit with the chemistry-inspired UCCSD ansatz can achieve the same accuracy as the exact diagonalization method except for systems whose Mayer bond order indices are larger than 2. For those systems, we then demonstrate that the heuristic hardware-efficient RealAmplitudes ansatz, even with a substantially shorter circuit depth, can provide a significant improvement over the UCCSD ansatz, verifying that the harmonic vibrational frequencies could be calculated accurately by quantum computation in the NISQ era.

DOI: 10.1103/PhysRevResearch.5.043216

I. INTRODUCTION

Recently, quantum computing has emerged as a promising way to potentially solve classically intractable problems, especially in the field of quantum chemistry, whose fundamental goal is to solve the Schrödinger equation for chemical or molecular systems [1–3]. With the quantum nature of wave functions, quantum computing makes use of superposition, entanglement and interference to prepare and manipulate quantum states, offering the potential for exponential speedup over classical computing. Along with this quantum advantage, the unitary coupled cluster with single and double excitations (UCCSD) method [4–6] for quantum computation of the molecular properties can be realized by mapping the exponential operators to qubit operators and using trotterization to approximate the corresponding quantum circuit. This circuit can be efficiently implemented on a quantum computer [7–10], making UCCSD a powerful tool for simulating chemical systems. The number of gates for the UCCSD circuit, using the Jordan-Wigner transformation [11], scales as $\mathcal{O}(N_q^3 N_e^2)$, where N_q is the number of qubits and N_e is the number of electrons [9]. On the other hand, classical implementation of the UCCSD method is impractical due to the nontruncated Baker-Campbell-Hausdorff expansion [12]. In addition, the gold standard method in quantum chemistry by

*yuanchung@ntu.edu.tw
†goan@phys.ntu.edu.tw

classical computation, the coupled cluster with single, double, and perturbative triple excitations (CCSD(T) [13]), scaling as $\mathcal{O}(N^7)$, where N is the number of molecular orbitals (MOs), is applicable only to small systems. Therefore the UCCSD method is considered a promising candidate for the quantum simulation, and its efficient implementation is a major area of research in the field of quantum computational chemistry. As quantum computing technology continues to advance, it is expected that the UCCSD method will play an increasingly important role in the development of applications of quantum chemistry.

In the so-called noisy intermediate-scale quantum (NISQ) era [14], quantum computers have noisy and a limited number of qubits without error correction implemented. As a result, the number of consecutive quantum gates that can be reliably run on the NISQ machines is also restricted. To address this challenge, a hybrid quantum-classical algorithm called the variational quantum eigensolver (VQE) [7] has been proposed and widely used. In this VQE framework, a parameterized quantum circuit as an ansatz to prepare the trial quantum states (wave functions) is optimized to find the ground state energy of the Hamiltonian based on the variational principle. The VQE algorithm leverages the strengths of quantum and classical computation, and distributes the computational workload between quantum and classical computers, with the quantum computer performing the trial quantum state preparation and measurement, and the classical computer performing the parameter optimization. The structure of the circuit ansatz and the ability of parameter optimization determine the accuracy of the result obtained from the VQE algorithm.

Reducing the requirement on the number of qubits and thus the depth of the quantum circuit is one of the major strategies in the NISQ era. For chemical problems, the chemistry-inspired UCCSD ansatz is commonly used. However, due to its complex circuit structure resulting from an exponential operator, the UCCSD ansatz can give rise to a very deep quantum circuit, making it difficult to be really implemented on NISQ devices. In contrast, the heuristic hardware-efficient ansatz is proposed to take advantage of its shorter circuit depth than that of the chemistry-inspired ansatz on NISQ devices [15]. The general construction of a hardware-efficient ansatz consists of alternating layers of parameterized single-qubit rotation gates and two-qubit entangling gates. In general, the true wave function of a quantum system can be expressed as a unitary transformation of the initial state [12]. Despite it is not guaranteed that a hardware-efficient ansatz of a unitary quantum circuit contains the solution of the wave function or it is optimal, and that it preserves the same properties as the true wave function [16,17], a hardware-efficient ansatz is flexible to vary the types and increase the number of layers of parameterized and entangling gates, allowing it to cover more of the solution space where the true wave function may reside. Some reviews about the recent development for the both kinds of ansatzes could be found in [18,19]. Recently, an adaptive variational algorithm, called adaptive derivative-assembled pseudo-Trotter ansatz VQE (ADAPT-VQE) [20], has been proposed to determine adaptively a quasioptimal ansatz with the minimal number of excitation operators (e.g. as in the UCCSD ansatz) for molecular simulations. Although ADAPT-VQE can yield an adaptive ansatz with a considerably reduced number of parameters, the ansatz circuit that becomes shallower than the UCCSD ansatz is still too deep to be implemented successfully on the current NISQ devices.

Besides the variational ansatz encoded in the trial wave function, quantum computation of quantum chemistry also depends on a representation of the molecular Hamiltonian. A common choice as a first demonstration is the minimal basis (MB) set like STO-3G [21], but the calculation using the STO-3G MB set usually does not yield an accurate result. To improve the predicted results to be comparable with their experimental data, a larger basis set such as cc-pVDZ [22–25] is widely adopted in classical computation of quantum chemistry. However, even for the simplest H_2 molecule, the VQE calculation with Hamiltonian in the cc-pVDZ basis set required 20 qubits and the circuit depth estimated by using the UCCSD ansatz would be over 10^4. Such a very deep circuit for H_2 is obviously not realizable on the current NISQ devices, not to mention for larger molecules. In order to represent a high-quality Hamiltonian in a smaller basis set for the NISQ devices without losing much of the accuracy, the choice of the basis set is thus extremely important, which was often overlooked in the quantum computing community. Different from traditional basis sets, the basis set constructed from the real-space numerical grid method, where the MO is expanded in the set of specific real-space basis functions, could fulfill the need here. For example, the multiresolution analysis (MRA) [26–28] method has been used to represent pair-nature orbitals as an efficient basis-set-free approach to simulate molecular systems for VQE by Kottmann *et al.* [29]. Although the qubit requirements can be significantly reduced, the MRA approach is treated as a black box, and for the BeH_2 molecule, their approach would not yield a smooth potential energy curve (PEC). More recently, Hong *et al.* [30] demonstrated that a MB set constructed from Daubechies wavelet [31] MOs calculated from BIGDFT [32–35] can yield accurate results in harmonic vibrational frequencies for H_2, LiH, and BeH_2 on quantum simulator with noisy model implemented from the real devices. The key feature of real-space basis set of, e.g., Daubechies wavelet MOs is that its accuracy could be systematically improvable as the number of basis functions is increased. It was demonstrated in Ref. [30] that VQE quantum computations for vibrational frequencies using the MB set of Daubechies wavelet MOs with accuracy comparable with that of the full configuration interaction (FCI) [36] calculation using the cc-pVDZ basis set, whereas the computational cost the same as that of a STO-3G calculation, have been achieved for this small set of three simple molecules. However, while further inspecting on the application of Daubechies wavelet basis set extended to a larger benchmark molecular dataset considered here (see Sec. II G), we find that the approach of using the MB set of Daubechies wavelet MOs [30] does not provide adequate results of vibrational frequencies as it fails to produce smooth PECs for some molecules, and furthermore by excluding those unavailable vibrational frequency data, it also has a significantly larger root-mean-square deviation (RMSD) value even though the number of active MOs used is considerably higher than that of wour proposed approach here (see Ref. [37]). This intrigues a study of active space since important orbitals may not be inside the MB set.

Selecting active space that includes most important MOs participating in physical or chemical reactions is an effective approach to form compact representations of molecular orbital spaces of a given basis set. Properly constructed active space reduces the number of MOs used, and is critical for performing quantum computations for large systems in the NISQ era; however, a comprehensive method that can select an appropriate active space for correlated molecular calculations, regardless of the basis sets, is still desirable. For example, a conventional active space selection scheme based on the occupation numbers of nature orbitals often identifies important orbitals successfully, but they are not foolproof and might falter in specific scenarios [38]. Therefore it remains worth inspecting suitable ways to choose an active space for the system under consideration.

In addition, several studies have demonstrated that using the MOs derived from the Kohn-Sham (KS) density functional theory (DFT) in correlated wave function theory calculations could provide improved results. [39–46] For example, Bertels *et al.* [46] have recently benchmarked classical CCSD(T) calculations with a variety of KS MOs on many diatomic molecules. In their paper, MOs derived from a hierarchy of different exchange-correlation (XC) functionals were compared and an improvement on predicting vibrational frequencies was observed. Their results indicate that the electron correlation effect could be incorporated into the MOs via the XC functional, which suggests that a possible further improvement on the accuracy of molecular vibrational frequency calculations could be realized when using KS MOs without increasing the number of MOs used.

In this work, we propose a quantum computational approach that adopts the molecular orbitals derived from KS DFT with XC functionals and expanded in the Daubechies wavelet basis set, where a reduced active space based on an energy criterion of a first-order pair energy in the theory of independent electron pair approximation (IEPA) [47,48], denoted by IEPA1, is selected to further reduce the required number of qubits. We perform a VQE quantum computing benchmark investigation using our proposed approach on the harmonic vibrational frequencies of 43 neutral, closed-shell diatomic molecules with results in great agreement with their corresponding experimental data. We attribute its excellent performance to three factors: (i) a better description of the Hamiltonian by the Daubechies wavelets MOs, (ii) better reference for the electron correlation effect in the MOs via the XC functional, (iii) an improved selection of active space by IEPA1 energy criteria. Remarkably, our proposed approach significantly reduces the number of qubits for the 43 diatomic molecules from 20 to 60 using the traditional cc-pVDZ basis set with frozen core approximation to only 2 to 12 but with similar accuracy for the obtained results. VQE calculations by our approach with a significantly reduced qubit number imply that a considerably reduced ansatz circuit depth can reach the same level of accuracy as those by using the traditional cc-pVDZ basis set. For example, the VQE calculation of a H_2 molecule using the cc-pVDZ basis set requires 20 qubits and the circuit depth estimated by using the UCCSD ansatz would be, as mentioned above, over 10^4, and even for the case of using the heuristic ansatz of a RealAmplitudes circuit, to have a result reaching the same level of accuracy the estimated circuit depth would be still about a few hundreds. In contrast, the required circuit depth of our proposed approach to obtain a result of harmonic vibrational frequency of H_2 in great agreement with the experimental data is only 4. For all the 43 molecules we investigate, the required circuit depths to reach accurate vibrational frequency results are all less than 100. Thus, VQE quantum computations of these molecules using our proposed approach are promisingly realizable on the NISQ devices given the recent advance on quantum utility demonstration [49] and the projected achievement of the so-called 100×100 Challenge in 2024 [50].

The paper is organized as follows. In Sec. II, we introduce the methods we use to calculate the molecular vibrational frequencies in our proposed approach. The active space selection criterion by IEPA1 will be described in Sec. II C. In Sec. III A, the performance of the proposed approach is benchmarked against the classical WFT methods and traditional basis sets on a large dataset of 43 neutral closed-shell diatomic molecules. Moreover, we use VQE to evaluate the performance of the UCCSD ansatz with the Hamiltonian represented by the selected Daubechies wavelet basis set in Sec. III B. The results show that the UCCSD ansatz can yield accurate results except for systems whose Mayer bond order indices [51] are larger than 2. For those systems, we then demonstrate that the heuristic hardware-efficient ansatz even with a substantially shorter circuit depth can provide significant improvement over the UCCSD ansatz in Sec. III C. To the best of our knowledge, our investigation is the first systematical benchmark study to demonstrate that a heuristic hardware-efficient ansatz could outperform a chemistry-inspired UCCSD ansatz in predicting accurate molecular properties by quantum computation. Such a comprehensive benchmark study enables us to establish an accurate approach for the vibrational frequencies of diatomic molecules, which could be realized on near-term NISQ computers. A conclusion and the outlook of this work are presented in Sec. IV.

II. METHODS

A. Daubechies wavelet

In this work, MOs expanded in the Daubechies wavelet basis set [31] are generated from the BIGDFT package [32–35]. The KS MO is expanded in the Daubechies wavelets of order 16 with one scaling function ϕ^0 and seven augmented wavelets $\psi^1, \ldots, \psi^7$ [30,32]:

$$\Psi^{\mathrm{KS}}(\mathbf{r}) = \sum_{\mathbf{i}} s_{\mathbf{i}}\, \phi^0_{\mathbf{i}}(\mathbf{r}) + \sum_{\mathbf{j}} \sum_{\lambda=1}^{7} w^{\lambda}_{\mathbf{j}}\, \psi^{\lambda}_{\mathbf{j}}(\mathbf{r}), \qquad (1)$$

where $s_{\mathbf{i}}, w_{\mathbf{j}}$ are expansion coefficients, and indices $\mathbf{i} = \{i_1, i_2, i_3\}$ and $\mathbf{j} = \{j_1, j_2, j_3\}$ are summed over the low (coarse) and high (fine) resolution regions, respectively, in three-dimensional real space grid points $\mathbf{r} = \{x, y, z\}$ with grid spacing h. The three-dimensional basis functions $\phi^0_{i_1,i_2,i_3}(\mathbf{r})$ and $\phi^{\lambda}_{j_1,j_2,j_3}(\mathbf{r})$ are a tensor product of one-dimensional scaling

function ϕ and wavelet ψ, which read as

$$\begin{aligned}
\phi^0_{i_1,i_2,i_3}(\mathbf{r}) &= \phi(x/h - i_1)\phi(y/h - i_2)\phi(z/h - i_3),\\
\psi^1_{j_1,j_2,j_3}(\mathbf{r}) &= \psi(x/h - j_1)\phi(y/h - j_2)\phi(z/h - j_3),\\
\psi^2_{j_1,j_2,j_3}(\mathbf{r}) &= \phi(x/h - j_1)\psi(y/h - j_2)\phi(z/h - j_3),\\
\psi^3_{j_1,j_2,j_3}(\mathbf{r}) &= \psi(x/h - j_1)\psi(y/h - j_2)\phi(z/h - j_3),\\
\psi^4_{j_1,j_2,j_3}(\mathbf{r}) &= \phi(x/h - j_1)\phi(y/h - j_2)\psi(z/h - j_3),\\
\psi^5_{j_1,j_2,j_3}(\mathbf{r}) &= \psi(x/h - j_1)\phi(y/h - j_2)\psi(z/h - j_3),\\
\psi^6_{j_1,j_2,j_3}(\mathbf{r}) &= \phi(x/h - j_1)\psi(y/h - j_2)\psi(z/h - j_3),\\
\psi^7_{j_1,j_2,j_3}(\mathbf{r}) &= \psi(x/h - j_1)\psi(y/h - j_2)\psi(z/h - j_3). \qquad (2)
\end{aligned}$$

The multiresolution of the Daubechies wavelets of order 16 is featured by the refinement equations

$$\begin{aligned}
\phi(x) &= \sqrt{2}\sum_{l=-7}^{8} h_l \phi(2x - l),\\
\psi(x) &= \sqrt{2}\sum_{l=-7}^{8} g_l \phi(2x - l),
\end{aligned} \qquad (3)$$

which establishes a relation between the scaling functions on a twice coarser grid and a finer grid. The coefficients h_l and $g_l = (-1)^l h_{-l}$ are filters that characterizes the scaling function and wavelet.

In a simulation domain, the chemical bonds are described in a high resolution region (fine region) which is composed of one scaling function and seven wavelets, and the exponentially decaying tails of the wave functions are described in a low resolution region (coarse region) which is only composed of scaling functions.

The number of virtual orbitals is a parameter and it is chosen to be equal to the total number of atomic input orbitals of the system as implemented in the BIGDFT package, while degenerate orbitals will be considered together. This forms the initial truncated MO space. The spin treatment does not involve spin polarization. The XC functionals considered in this work are Hartree-Fock (HF), Perdew-Burke-Ernzerhof (PBE) [52], and PBE0 [53]. Other popular XC functionals which have no suitable pseudopotential in the package, like B3LYP [54], are excluded. All the three XC functionals use the same norm-conserving Hartwigsen-Goedeker-Hutter Krack (HGH-K) [55–57] pseudopotential generated with the PBE functional.

One caveat needed to be mentioned in this work is about the grid parameters in BIGDFT. The grid parameters, hgrids which controls the grid spacing and rmult which controls the size of simulation space, are determined from the analysis of the grid convergence for each molecule. However, such determination is analyzed in the framework of DFT where virtual orbitals are not used, and thus might not be optimal for WFT since these grid parameters significantly affect the properties such as the shapes and energies of the virtual orbitals. Generally, setting better grid parameters (smaller hgrids and larger rmult) helps the calculation converge to a lower energy but with increasing computing cost, so the optimal choice should consider both accuracy and efficiency. The setting with better and better grid parameters could generate continuumlike orbitals, but WFT with continuumlike virtual orbitals would suffer from the vanishing electron correlation [58]. Therefore to find a better way to determine the grid parameters for WFT is crucial and will be investigated further in the future.

B. Second-quantized Hamiltonian

Given a set of KS MOs, the second-quantized Hamiltonian is constructed as

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2}\sum_{p,q,r,s} h_{\mathrm{pqrs}} a_p^\dagger a_q^\dagger a_r a_s, \qquad (4)$$

where the one-electron integral

$$h_{pq} = \int d\mathbf{x}\ \Psi_p^{\mathrm{KS}^*}(\mathbf{x})\left(-\frac{\nabla^2}{2} - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|}\right)\Psi_q^{\mathrm{KS}}(\mathbf{x}), \qquad (5)$$

the two-electron integral

$$h_{\mathrm{pqrs}} = \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{\Psi_p^{\mathrm{KS}^*}(\mathbf{x}_1)\Psi_q^{\mathrm{KS}^*}(\mathbf{x}_2)\Psi_r^{\mathrm{KS}}(\mathbf{x}_2)\Psi_s^{\mathrm{KS}}(\mathbf{x}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \qquad (6)$$

and $a_p^\dagger$ and a_q are creation and annihilation operators acting on the pth and qth component of the occupation number vector in Fock space, respectively. The $\mathbf{x} = (\mathbf{r}, \sigma)$ denotes the position and the spin of the electron, and $\mathbf{R}_A$ and Z_A denote the position and the atomic number of the Ath nucleus, respectively. The values of the integrals h_{pq} and h_{pars} are calculated via BIGDFT subroutines.

The correlation energy E_{corr} defined in WFT with KS MOs is [43]

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - \langle\Psi^{\mathrm{KS}}|H|\Psi^{\mathrm{KS}}\rangle, \qquad (7)$$

where E_{exact} is the exact energy, $|\Psi^{\mathrm{KS}}\rangle$ is the single Slater determinant formed with a set of KS spin orbitals and $\langle\Psi^{\mathrm{KS}}|H|\Psi^{\mathrm{KS}}\rangle$ is the reference value mimicking the HF energy.

C. Active space

We discuss here how the [MB] active space and [IEPA1] active space are selected. The [MB] active space selected from the initial truncated MO space is chosen to imitate the complete active space constructed from the minimal basis set. Let us take LiF as an example. The minimal basis set for this system is constructed from the $1s$, $2s$, and $2p$ orbitals of Li, and the $1s$, $2s$, and $2p$ orbitals of F, giving a total of 10 MOs from a linear combination of atomic orbitals (LCAO). Since this system has 12 electrons resulting in 6 occupied MOs, one would take the 6 occupied MOs and 4 lowest-energy unoccupied MOs in the [MB] method to select a total of 10 MOs in the active space. However, the usual frozen core approximation is used here for F's $1s$ orbital, but we do not apply this approximation to Li's $1s$ orbital since the pseudopotential used in this work includes that orbital (the same treatment applied to the alkaline earth metal element Be, while for Na atom the 1s orbital is frozen). By taking these considerations into account, the active space of LiF consists of 10 active electrons and 9 active MOs, denoted as [10,9], in

our [MB] approach. Note that for group 13 - group 17 atoms in the periodic table of the elements, the inner-shell frozen core approximation is used and the total number of minimal-basis valence orbitals for these atoms is 4.

The active orbitals in the [IEPA1] active space are selected from the initial truncated MO space by calculating the first-order pair energy in the theory of the independent electron pair approximation. In the following, we describe briefly the theory of IEPA and then present an energy criterion derived from IEPA1 to select active orbitals. The theory of IEPA considers the correlation energy associated with a pair of electrons independently of other pairs in a configuration interaction way. The correlated wave function for the pair ij, denoted as the pair function $|\Psi_{ij}\rangle$, is

$$|\Psi_{ij}\rangle = |\Psi_0\rangle + \sum_{a<b} c_{ij}^{ab} \left|\Psi_{ij}^{ab}\right\rangle, \tag{8}$$

where i and j denote the occupied spin orbital indices, a and b denote the virtual spin orbital indices, and c_{ij}^{ab} is the wave function coefficient. Here, $|\Psi_0\rangle$ and $|\Psi_{ij}^{ab}\rangle$ are ground and doubly excited Slater determinants formed with a set of HF spin orbitals. The energy of this correlated wave function, denoted by E_{ij}^{IEPA}, is

$$E_{ij}^{\text{IEPA}} = \langle\Psi_{ij}|H|\Psi_{ij}\rangle = E_0 + e_{ij}^{\text{IEPA}}, \tag{9}$$

where E_0 is the HF reference energy and e_{ij}^{IEPA} is the pair (correlation) energy. Under the first-order approximation to IEPA, which neglects coupling between excited determinants, the first-order pair energy [59], denoted by e_{ij}^{IEPA1}, reads

$$e_{ij}^{\text{IEPA1}} = \sum_{a<b}^{\text{vir}} \frac{\left|\langle\Psi_0|H\left|\Psi_{ij}^{ab}\right\rangle\right|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} = \sum_{a<b}^{\text{vir}} \frac{(h_{ijab} - h_{ijba})^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \tag{10}$$

where h_{ijab} is the two-electron integral defined in Eq. (6) and the energy difference in the denominator, $\langle\Psi_{ij}^{ab}|H - E_0|\Psi_{ij}^{ab}\rangle$, has been approximated by the difference of orbital energies, $\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$. The total first-order pair energy is then

$$E_{\text{IEPA1}} = \sum_{i<j}^{\text{occ}} e_{ij}^{\text{IEPA1}} = \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} \frac{(h_{ijab} - h_{ijba})^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \tag{11}$$

$$= E_{\text{MP2}}, \tag{12}$$

which is identical to the energy correction by the second-order Møller-Plesset perturbation theory (MP2), denoted by E_{MP2} [60].

The pair energy would suggest a selection criterion for the active space. The active orbitals of the [IEPA1] active space are selected from the MOs in initial truncated MO space by using the first-order pair energy. In order to determine which MO is important, we calculate E_{IEPA1} of the MO by considering the sum of all pair energies involved that MO, that is, the terms in the summation involving only the spin indices of that MO. For example, the first occupied MO is denoted as OccMO[0], and its spin-up and spin-down orbitals are labeled with indices $i = 0$ and $i = 1$, respectively; in this case, E_{IEPA1} of OccMO[0] is

$$E_{\text{IEPA1}}(\text{OccMO}[0]) = \left(\sum_{i=0<j}^{\text{occ}} + \sum_{i=1<j}^{\text{occ}}\right) e_{ij}^{\text{IEPA1}}$$

$$= \left(\sum_{i=0<j}^{\text{occ}} + \sum_{i=1<j}^{\text{occ}}\right) \sum_{a<b}^{\text{vir}} \frac{(h_{ijab} - h_{ijba})^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \tag{13}$$

Take each E_{IEPA1} of the MO divided by the total E_{IEPA1} as a percentage, and then with a target of selecting a small number of MOs, the MOs with relatively large percentages are chosen into the IEPA1 active space. A different flavor to directly determine MRA-represented pair-natural orbitals on the level of MP2 can be found in Refs. [28,29]; the number of active pair-natural orbitals (approximate nature orbitals), is truncated based on occupation numbers. Compared to their approach [28,29], our approach directly analyzes on canonical orbitals without additional transformation to nature orbitals.

While there is no distinguishable difference between the energies E_{IEPA1} and E_{MP2} for the standard HF orbitals, we follow by the idea of IEPA to evaluate the first-order pair energy when the KS orbitals are used. However, in this case the Brillouin's theorem, which states that the matrix element contributed from singly excited Slater determinants formed with a set of HF spin orbitals is zero, does not hold due to the fact that the KS MO is not the eigenfunction of the Fock operator.

D. Quantum computing

Quantum computing in VQE starts from mapping the second-quantized Hamiltonian to the qubit Hamiltonian, and in this work the quantum computing package, QISKIT [61], is used. Common encoding methods to encode the fermionic operators to qubit operators transform a fermionic system of m active MOs ($2m$ spin orbitals) into an $2m$-qubit system. Here, the parity encoding scheme [62,63] is chosen to further reduce the number of qubits by two due to $\mathbb{Z}_2$ symmetry reduction [64], so the number of qubits used is $2m - 2$.

The ground state energy of the qubit Hamiltonian is calculated via different methods. In Sec. III A, the exact diagonalization of the qubit Hamiltonian is used and we denote this approach as the exact diagonalization method of quantum computing (EDQC). Hence, the results of the EDQC method could be regarded as the best results achievable by quantum computation. Consequently, the EDQC method is also used as the standard to investigate the performance of the VQE method in Secs. III B and III C, where we demonstrate the result of the vibrational frequencies obtained by our proposed approach for VQE quantum computation is as accurate as those by the EDQC method. At the same time, the state fidelity between the EDQC ground state and the VQE ground state calculated via different circuit ansatzes is evaluated as another verification indicator (see Table III).

The quantum circuit ansatzes used here for VQE are the UCCSD ansatz and the heuristic hardware-efficient ansatz of RealAmplitudes from the QISKIT circuit library, where these VQE approaches are denoted as VQE(UCCSD) and VQE(RealAmplitudes), respectively. The Hartree-Fock state

as the initial state is prepended to both the quantum circuits. The entanglement type for the RealAmplitudes ansatz is chosen to be the linear entanglement. The optimizers used are SLSQP and L-BFGS-B of SCIPY [65]. All the VQE calculations are performed in the noiseless situation with the state-vector simulation method.

E. Classical computing

We use the PYSCF package [66] for the classical computations. The methods used include CCSD(T) and complete active space configuration interaction (CASCI) [67], where the former is chosen since it is the golden standard method in quantum chemistry and the latter is chosen since it is the FCI on the active space which corresponds to the exact diagonalization method in this work. The KS MOs with XC functional of HF or PBE0 are used, and the nature orbital [68] is additionally considered in the CASCI method. The traditional basis set used is the Dunning correlation-consistent basis set, cc-pVDZ. The spin treatment is spin-restricted Hartree-Fock method. The number of frozen cores is chosen to be the same as that of the pseudopotential considered in this work. In CASCI, for a rapid convergence, the MP2 natural orbitals, denoted by MP2NO, transformed from the standard HF MOs are used and then the active space is determined by the natural orbital occupation number, denoted by NOON.

In the calculation of Mayer bond order indices, we use PYSCF functions to evaluate the formula. This index between atoms A and B of a closed-shell molecule is defined as

$$M_{AB} = \sum_{\mu\in A}\sum_{\nu\in B}(\mathbf{DS})_{\mu\nu}(\mathbf{DS})_{\nu\mu}, \tag{14}$$

where μ and ν are indices for the basis functions belonged to the assigned atom, and **DS** denotes the product the spinless density matrix **D** and the overlap matrix **S**.

F. Harmonic vibrational frequency

The quantity to be calculated in the benchmark is the harmonic vibrational frequency. For diatomic molecules, there is only one vibrational mode, the stretching mode. The corresponding diatomic harmonic vibrational frequency is calculated by quadratic polynomial curve fitting using five points around the minimal energy point of the equilibrium bond length with step size 0.01 Å on the PEC. In order to put different comparative methods on equal footing, this calculation procedure applies to all the methods considered here. The equilibrium bond lengths calculated by different methods are presented in Ref. [37].

G. Dataset

For simplicity, the dataset considered consists of diatomic molecules that are neutral, closed-shell and formed by atoms (elements) in row 1 to row 4 of the periodic table, excluding the transition metal elements, but the diatomic molecules whose experimental data are not available on the Computational Chemistry Comparison and Benchmark Database [69] are also not considered. For comparison, since there is no K atom in the cc-pVDZ basis set, the molecules involved K atom are excluded. In BIGDFT, we can not generate smooth PECs for NaLi, Na_2 and NaK, and therefore these three molecules are excluded. C_2 with multireference character and F_2 owing to severe static correlation [70] are taken to be the overall outliers and then are excluded. In the end, the benchmark dataset contains 43 neutral closed-shell diatomic molecules.

H. Notations

We use the following notations to denote different approaches used in this work: "Method[active space selection method]-XC/Basis Set," where XC can be HF, PBE, or PBE0 to keep track of the type of MOs. For example, EDQC[IEPA1]-PBE0/Wavelet, denotes using the EDQC method with PBE0 exchange functional for the Daubechies Wavelet MO basis set, and the method of the active space selection is IEPA1, where Wavelet is shorthand for Daubechies Wavelet MO basis set.

The notation [MB] or [IEPA1] indicates how the active space are selected as we discuss in Sec. II C. The active space indices $[n, m]$ in Table I denote n active electrons and m active MOs ($2m$ spin orbitals).

III. RESULTS

A. Performance of the Daubechies wavelet basis set

We evaluate the performance of KS MOs generated from HF, PBE, and PBE0 XC functionals in the Daubechies wavelet basis set. To further reduce the qubit number requirement, we use a reduced active space based on a IEPA1 energy criterion in our proposed approach. In order to fully reveal the performance of the Daubechies wavelet basis set, we adopt the result of the EDQC method of the qubit Hamiltonian in the given active space to compare to the experimental value so that the error is fully attributed to the inadequacy of the basis set.

1. H_2*, LiH, and HF*

To begin with, we choose three simple molecules, H_2, LiH, and HF, from Table I, to benchmark the accuracies of the different methods and their respective errors in predicting the harmonic vibrational frequencies when compared to the experimental values (Fig. 1).

For H_2, the approach EDQC[2,2]-XC/Wavelet (exact diagonalization method; the active space [2,2] determined by IEPA1) predicts vibrational frequencies of 4462.40 (61.19) cm^{-1}, 4365.33 (−35.88) cm^{-1}, and 4391.45 (−9.76) cm^{-1} for XC = HF, PBE, and PBE0 functional, respectively, where the value inside the parenthesis denotes the error (difference) with respect to the experimental value and the same notation will be used in the following text. The improvement in the predicted vibrational frequency with the increasing level of the XC functionals is obvious. Notably, CASCI[2,2]-PBE0/cc-pVDZ (the active space [2,2] imitating the minimal basis set case), CASCI[2,2]-MP2NO/cc-pVDZ (using MP2 nature orbital; the active space [2,2] determined by NOON), and CASCI[2,3]-PBE0/cc-pVDZ (the active space [2,3] determined by IEPA1), yield vibrational frequencies of 4291.21 (−110.0) cm^{-1}, 4224.68 (−176.53) cm^{-1}, and 4194.85 (−206.35) cm^{-1}, respectively, and all of which perform quite poorly. The results indicate that classical methods with a traditional atom-centered basis set do not perform

TABLE I. Harmonic vibrational frequencies (in cm^{-1}) of neutral closed-shell diatomic molecules obtained by different methods.

Mol.	EDQC [IEPA1]-HF /Wavelet	EDQC [IEPA1]-PBE /Wavelet	EDQC [IEPA1]-PBE0 /Wavelet	CASCI [IEPA1]-PBE0 /cc-pVDZ	CASCI [MB]-PBE0 /cc-pVDZ	CASCI [NOON]-MP2NO /cc-pVDZ	CCSD(T) -PBE0 /cc-pVDZ	CCSD(T) -HF /cc-pVDZ	Expt.
H_2	[2,2]4462.40	[2,2]4365.33	[2,2]4391.45	[2,3]4194.85	[2,2]4291.21	[2,2]4224.68	[2,10]4397.89	[2,10]4397.70	4401.21
LiH	[2,3]1413.65	[2,3]1393.42	[2,3]1405.36	[2,5]1309.67	[4,6]1372.62	[2,5]1393.25	[4,19]1350.18	[4,19]1350.18	1405.50
NaH	[2,3]1178.10	[2,3]1149.01	[2,3]1143.76	[2,5]1153.83	[10,9]1107.69	[2,5]1128.44	[10,22]1107.59	[10,22]1107.47	1171.97
BH	[4,5]2476.84	[4,7]2418.72	[4,7]2386.09	[4,9]2353.90	[4,5]2166.48	[4,9]2390.04	[4,18]2318.76	[4,18]2319.91	2366.73
AlH	[4,6]1707.87	[4,7]1756.43	[4,7]1709.33	[4,9]1649.35	[4,5]1554.54	[4,9]1666.97	[4,18]1672.14	[4,18]1672.42	1682.38
GaH	[4,6]1600.18	[4,7]1575.06	[4,7]1593.80	[4,8]1600.89	[4,5]1509.07	[4,9]1612.96	[4,18]1602.76	[4,18]1597.07	1604.52
HF	[6,6]4432.96	[2,3]4152.30	[2,3]4148.44	[8,10]4631.00	[8,5]4141.97	[6,6]4171.69	[8,18]4151.90	[8,18]4151.79	4138.39
HCl	[2,3]3008.83	[2,3]2883.96	[2,3]2874.66	[8,9]2900.94	[8,5]2858.48	[8,9]3162.91	[8,18]3018.33	[8,18]3017.70	2990.93
HBr	[2,4]2623.92	[2,3]2515.62	[2,3]2511.84	[8,9]2615.01	[8,5]2541.63	[8,9]2710.44	[8,18]2654.50	[8,18]2654.02	2649.00
LiF	[4,6] 932.22	[4,6] 939.94	[4,6] 938.00	[8,7] 959.51	[10,9] 999.36	[8,7] 980.57	[10,27] 975.01	[10,27] 975.36	910.57
LiCl	[4,6] 648.66	[4,4] 663.29	[4,4] 649.26	[8,7] 656.13	[10,9] 632.17	[6,6] 635.55	[10,27] 625.11	[10,27] 624.96	642.95
LiBr	[6,5] 558.04	[6,5] 571.19	[6,5] 559.84	[8,7] 540.32	[10,9] 555.71	[6,8] 582.25	[10,27] 554.50	[10,27] 554.87	563.00
NaF	[6,5] 541.11	[6,4] 552.30	[6,4] 548.70	[8,7] 589.84	[16,12] 594.09	[6,6] 584.83	[16,30] 582.15	[16,30] 582.36	535.66
NaCl	[6,4] 362.28	[6,4] 363.57	[6,4] 362.88	[8,7] 373.37	[16,12] 354.34	[6,6] 343.89	[16,30] 349.63	[16,30] 349.83	364.68
NaBr	[6,4] 292.53	[6,4] 293.19	[6,4] 292.99	[8,7] 354.87	[16,12] 289.84	[6,8] 311.71	[16,30] 288.17	[16,30] 288.17	302.00
BeO	[6,9]1506.27	[6,6]1322.15	[6,6]1436.71	[8,6]1635.71	[10,9]1262.73	[6,6]1500.43	[10,27]1365.29	[10,27]1358.46	1457.09
BeS	[6,8] 866.53	[6,6]1002.45	[6,6] 949.56	[8,6]1085.78	[10,9] 911.52	[6,6] 984.06	[10,27] 960.90	[10,27] 959.81	997.94
BF	[6,7]1353.21	[8,7]1495.16	[8,7]1462.33	[10,8]1504.25	[10,8]1277.04	[8,7]1349.93	[10,26]1304.94	[10,26]1306.53	1402.16
BCl	[8,6] 859.42	[8,6] 873.82	[8,6] 852.18	[10,8] 772.05	[10,8] 772.15	[8,7] 867.14	[10,26] 841.81	[10,26] 840.77	840.29
BBr	[8,6] 676.96	[8,6] 686.53	[8,6] 686.03	[10,8] 644.16	[10,8] 644.07	[8,7] 700.09	[10,26] 695.24	[10,26] 694.65	684.31
AlF	[8,7] 815.15	[8,7] 839.01	[8,7] 814.55	[10,8] 791.47	[10,8] 787.41	[8,7] 777.61	[10,26] 788.84	[10,26] 789.40	802.32
AlCl	[2,3] 480.17	[4,4] 486.10	[4,4] 474.55	[10,7] 455.72	[10,8] 438.92	[8,7] 462.90	[10,26] 474.84	[10,26] 474.51	481.77
AlBr	[2,3] 374.09	[4,4] 375.63	[4,4] 371.20	[10,7] 363.31	[10,8] 344.91	[8,7] 374.39	[10,26] 375.61	[10,26] 375.12	378.00
GaF	[8,6] 591.41	[8,6] 623.13	[8,6] 609.27	[10,8] 641.78	[10,8] 646.70	[8,7] 654.89	[10,26] 659.63	[10,26] 659.25	622.20
GaCl	[4,4] 333.49	[4,4] 353.81	[4,4] 340.81	[10,7] 356.66	[10,8] 340.84	[8,7] 365.68	[10,26] 389.50	[10,26] 372.43	365.30
CO	[8,8]2126.40	[8,6]2297.41	[8,6]2248.86	[10,8]2415.09	[10,8]2118.43	[8,7]2189.97	[10,26]2101.17	[10,26]2180.17	2169.76
CS	[8,6]1334.09	[8,7]1338.66	[8,7]1298.67	[10,8]1185.46	[10,8]1185.46	[8,7]1282.47	[10,26]1248.77	[10,26]1250.42	1285.16
CSe	[10,7]1086.94	[8,8] 903.59	[87]1010.42	[10,8] 964.48	[10,8] 963.57	[8,7]1036.65	[10,26]1006.64	[10,26]1007.86	1035.36
SiO	[8,8]1311.41	[8,7]1238.18	[8,7]1247.22	[10,7]1239.35	[10,8]1100.60	[8,7]1200.23	[10,26]1139.71	[10,26]1140.03	1241.54
SiS	[8,8] 784.94	[8,8] 732.18	[8,8] 716.80	[10,8] 682.38	[10,8] 682.32	[8,7] 732.23	[10,26] 727.85	[10,26] 727.80	749.65
SiSe	[8,6] 554.84	[8,6] 591.19	[8,6] 582.23	[10,8] 538.16	[10,8] 538.59	[8,7] 571.14	[10,26] 569.19	[10,26] 569.55	580.00
GeO	[8,8]1074.76	[8,6] 995.56	[8,6] 999.61	[10,8] 885.21	[10,8] 885.82	[8,7] 989.61	[10,26] 939.51	[10,26] 957.82	985.50
N_2	[4,4]2345.03	[4,4]2400.36	[4,4]2399.42	[10,8]2636.51	[10,8]2146.61	[8,7]2332.91	[10,26]2329.44	[10,26]2331.16	2358.57
PN	[4,4]1365.63	[4,4]1375.26	[4,4]1375.98	[10,8]1220.52	[10,8]1220.52	[8,7]1301.49	[10,26]1254.95	[10,26]1293.57	1336.95
P_2	[4,4] 800.32	[4,4] 796.93	[4,4] 778.53	[10,8] 722.56	[10,8] 722.56	[8,7] 747.38	[10,26] 749.68	[10,26] 749.67	780.77
AsN	[4,4]1052.25	[4,4]1085.11	[4,4]1088.54	[10,8] 934.37	[10,8] 934.65	[8,7]1055.11	[10,26]1023.62	[10,26]1023.89	1068.54
As_2	[4,4] 414.06	[4,4] 413.52	[4,4] 414.28	[10,8] 388.45	[10,8] 388.42	[8,7] 410.69	[10,26] 414.24	[10,26] 414.24	430.00
Li_2	[2,5] 388.38	[2,5] 321.77	[2,5] 315.30	[2,5] 315.00	[6,10] 308.12	[2,5] 343.25	[6,28] 337.73	[6,28] 337.73	351.41
ClF	[10,8] 752.08	[10,7] 808.43	[10,7] 689.23	[14,10] 569.73	[14,8] 578.98	[12,9] 745.00	[14,26] 717.57	[14,26] 696.26	783.45
Cl_2	[10,7] 539.06	[10,7] 498.40	[10,7] 504.52	[14,10] 448.17	[14,8] 450.35	[12,9] 522.17	[14,26] 509.86	[14,26] 509.71	559.75
BrF	[10,8] 649.39	[10,7] 598.63	[10,7] 625.71	[14,10] 526.99	[14,8] 533.89	[12,9] 662.93	[14,26] 619.66	[14,26] 619.35	669.68
BrCl	[6,5] 414.75	[10,7] 410.42	[10,7] 400.00	[14,10] 424.17	[14,8] 366.81	[12,9] 422.58	[14,26] 402.30	[14,26] 402.07	444.32
Br_2	[6,5] 306.19	[8,6] 285.86	[8,6] 288.04	[12,9] 311.25	[14,8] 267.78	[12,9] 314.74	[14,26] 294.13	[14,26] 294.17	325.00
RMSD	60.55	54.07	41.44	123.68	98.11	46.83	43.74	41.73	
MSD	11.18	−4.04	−9.75	−8.70	−72.73	−2.76	−23.84	−21.77	
MAD	36.16	37.99	28.53	82.27	80.88	30.25	34.60	32.10	

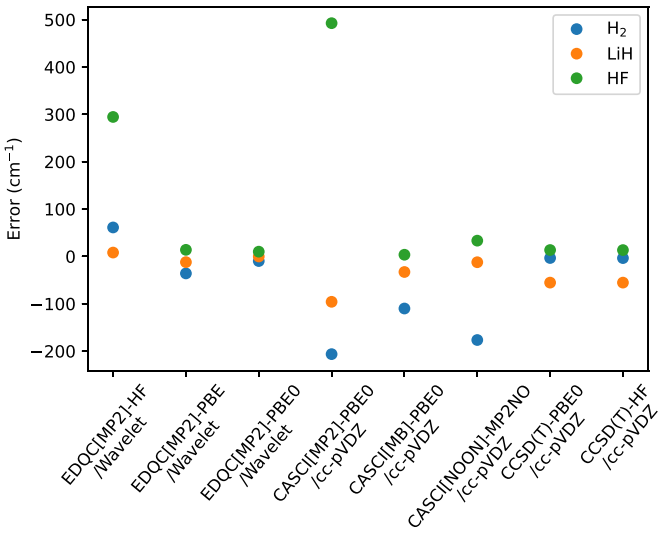

FIG. 1. Errors in the harmonic vibrational frequencies of H_2, LiH, and HF molecules (in cm^{-1}) calculated by a variety of methods. The errors (differences) are obtained by comparing the results to the corresponding experimental values.

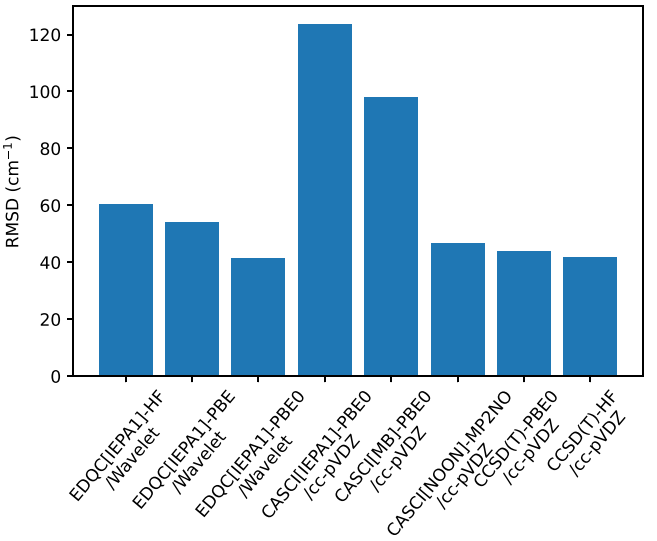

FIG. 2. RMSD of the harmonic vibrational frequencies (in cm^{-1}) obtained by comparing the results to their corresponding experimental values for a variety of methods.

well with a truncated small active space. The best performance in classical methods as expected is from CCSD(T)-HF/cc-pVDZ, which uses 10 MOs to yield a result of 4397.70 (-3.51) cm^{-1}, while the result obtained by CCSD(T)-PBE0/cc-pVDZ is about the same.

For LiH, EDQC[2,3]-PBE0/Wavelet which gives 1405.36 (-0.14) cm^{-1} performs best, and furthermore the size of its active space is the smallest. Similarly, for the HF (hydrogen fluoride) molecule, EDQC[2,3]-PBE/Wavelet which gives 4152.30 (13.91) cm^{-1} and EDQC[2,3]-PBE0/Wavelet which gives 4148.44 (10.04) cm^{-1} both using only 3 MOs determined by IEPA1 outperform the other methods.

To sum up, for these three cases, the EDQC[IEPA1]-PBE0/Wavelet approach dominates the performance in both accuracy and efficiency at the same time. Clearly, the adaptation of a Daubechies wavelet basis set and a small number of selected KS MOs can yield the vibrational frequencies as accurate as those obtained by the high-level WFT methods with a much larger number of MOs in the cc-pVDZ basis set. Therefore the KS Daubechies wavelet MOs in the active space selected by IEPA1 can provide significant improvement in describing molecular Hamiltonians in both size and quality.

2. *Performance on the overall dataset*

To evaluate the overall performance of all the methods, we present statistical measures of errors of the harmonic vibrational frequencies including RMSD, mean signed deviation (MSD), and mean absolute deviation (MAD) in Table I. In addition, the RMSD for each method is presented in Fig. 2. Moreover, Fig. 3 presents the scatter plots to display the correlation between theoretical predictions and experimental data in different approaches. As shown at the bottom of Table I, the EDQC[IEPA1]-XC/Wavelet approach gives the RMSD values of 60.55, 54.07, and 41.44 cm^{-1} for XC = HF, PBE, and PBE0 functional, respectively. Overall, the EDQC[IEPA1]-XC/Wavelet approach improves with the increasing level of the XC functionals. This trend is consistent with that of using the classical CCSD(T) method with MOs in similar XC functionals for the pruned closed-shell dataset in Ref. [46]. As clearly seen from Table I and Fig. 2, the approach with the best performance among all the methods is EDQC[IEPA1]-PBE0/Wavelet. So EDQC[IEPA1]-PBE0/Wavelet is the approach proposed in this work, and it gives results in excellent agreement with the experimental data. We attribute its great performance to three factors: (i) a better description of the Hamiltonian by introducing the Daubechies wavelets MOs, (ii) incorporating the electron correlation effect into the MOs via the XC functional, (iii) a suitable selection of active space by IEPA1. In the following, we will discuss and emphasize these points through the comparison with relevant classical methods.

The Daubechies wavelet MOs clearly outperform conventional cc-pVDZ basis set in our benchmark (see Fig. 2). For example, EDQC[IEPA1]-HF/Wavelet has a decent performance, and if the HF molecule [the orange point in Fig. 3(a)] is excluded, its RMSD value can be significantly improved to 41.06 cm^{-1}. One of the errors for the outliers of EDQC[IEPA1]-HF/Wavelet might be due to the choice on the size of the initial truncated MO space, as more virtual orbitals generated by the HF XC functional outside the initial truncated MO space should be considered to reduce the error. On the other hand, if the electron correlation can be incorporated into the MOs, as in the EDQC[IEPA1]-XC/Wavelet method, using the same size of the initial truncated MO space could produce more accurate results than EDQC[IEPA1]-HF/Wavelet.

For the KS MOs with XC functional beyond HF, the electron correlation effect is incorporated into the basis set via the XC functional. As one can see from Fig. 2, the RMSD values of EDQC[IEPA1]-XC/Wavelet can be reduced with the increasing level of the XC functionals from XC = HF to PBE and then to PBE0. The two outliers, HCl and HBr [the middle green points in Fig. 3(c)], with errors larger than

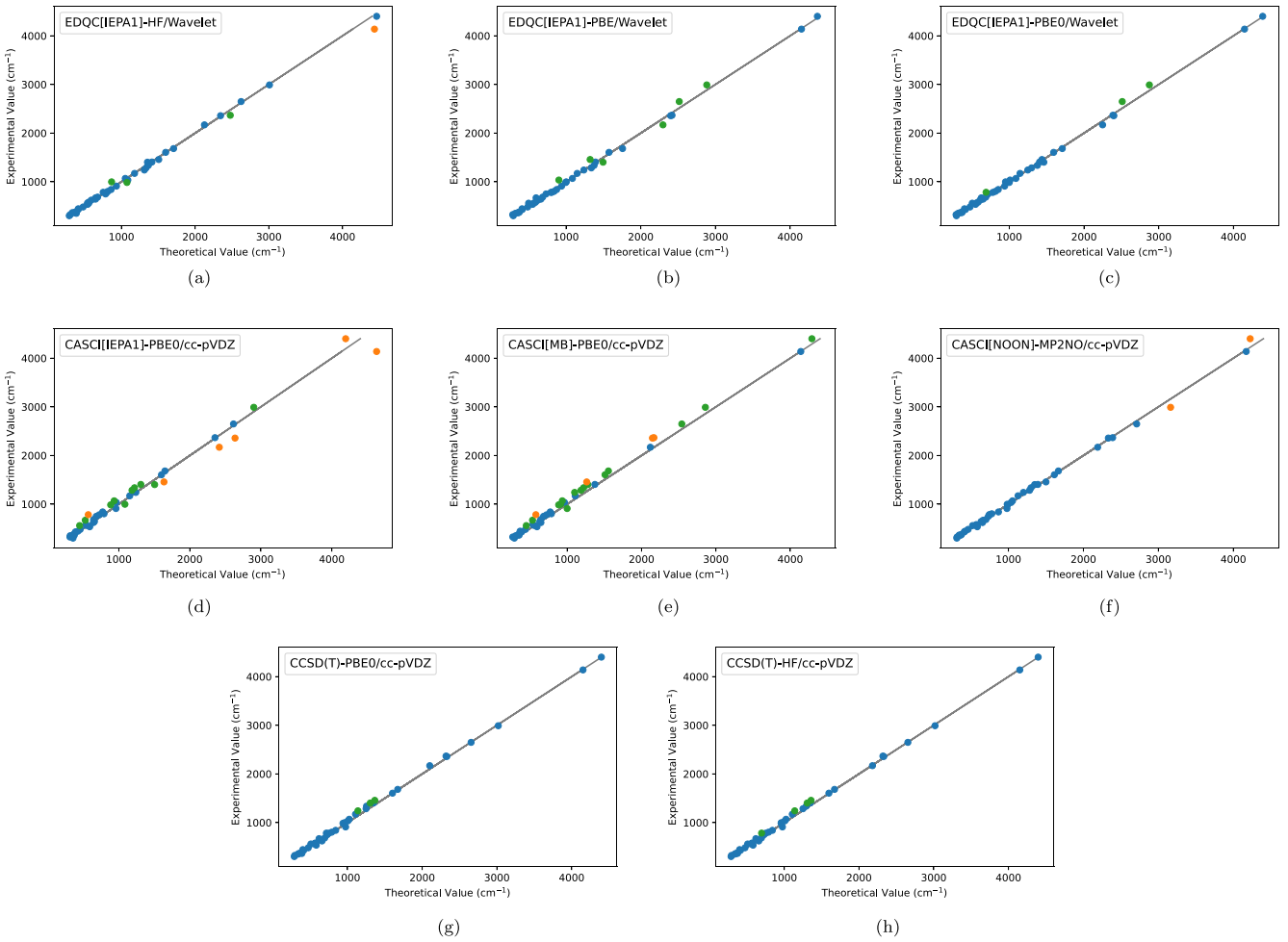

FIG. 3. Scatter plots of the harmonic vibrational frequencies of the diatomic molecules versus their corresponding experimental values for each benchmark method. The blue, green, and orange dots denote outliers with absolute error deviation (in cm^{-1}) less than 85, between 85 and 150, and larger than 150, respectively.

100 cm^{-1} for EDQC[IEPA1]-PBE0/Wavelet can be attributed to the over-corrections resulting in the error noncancellation because DFT-PBE0/Wavelet (presented in Ref. [37]) and EDQC[IEPA1]-HF/Wavelet perform well for these two molecules.

Finally, it is evident that selecting the active MOs correctly is essential in order to yield accurate results with minimal computational cost. It is necessary to reduced the number of qubits required in order to make a calculation successful in the NISQ era, and we propose to utilize a IEPA1-based approach for active MO selection. The success of the EDQC[IEPA1]-XC/Wavelet approaches demonstrates the effectiveness of the IEPA1 MO selection method. In contrast, the EDQC[MB]-XC/Wavelet approaches whose active orbitals are selected in the order of ascending orbital energies do not provide adequate results (see the data presented in Ref. [37]). Compared with the EDQC[IEPA1]-XC/Wavelet approaches, all the EDQC[MB]-XC/Wavelet approaches have significantly larger RMSD values, even though the number of active MOs used in these [MB] approaches are significantly higher.

To reveal the advantage of using the Daubechies wavelet basis set and IEPA1 active space selection, we compare the accuracy of the best achievable quantum computing method (EDQC[IEPA1]-PBE0/Wavelet) to that of the gold standard classical method (CCSD(T)-HF/cc-pVDZ, RMSD = 41.73 cm^{-1}). The accuracies of these two approaches are comparable; however, the number of orbitals used in EDQC[IEPA1]-PBE0/Wavelet is much less than those in CCSD(T)-HF/cc-pVDZ, showing a polynomial advantage in the number of MOs. In the whole dataset, one can see that only 2 to 7 MOs are used in EDQC[IEPA1]-PBE0/Wavelet for quantum computing, whereas the CCSD(T)-HF/cc-pVDZ methods require 10 to 30 MOs in order to achieve similar accuracies. This difference in the number of orbitals leads to a significant change in computational complexity, as measured by the number of terms in the Hamiltonian that need to be evaluated. As a result, even in our scenario with a small number of MOs, the number of evaluations needed for the quantum computation of the EDQC[IEPA1]-PBE0/Wavelet approach scales as at most $(2 \times 7)^4 = 14^4 \approx 3.84 \times 10^4$ while the scaling of CCSD(T) is at most $30^7 \approx 2.19 \times 10^{10}$. This clearly demonstrates the potential of the quantum computing method in achieving a quantum advantage.

To display our proposed approach has its uniqueness, we compare EDQC[IEPA1]-PBE0/Wavelet to the classical

corresponding method, CASCI[IEPA1]-PBE0/cc-pVDZ, where the dramatic difference in RMSD (the latter is 123.68 cm^{-1}) comes from the use of different basis sets, indicating that the same approach (IEPA1 active space and XC=PBE0) with the traditional basis sets is quite inaccurate. Moreover, we observe from Fig. 3(e) that the results of the same classical approach with active space imitating the minimal basis set, CASCI[MB]-PBE0/cc-pVDZ, with RMSD being 98.11 cm^{-1} are apparently red-shifted (see also the corresponding MSDs in Table I). On the other hand, both appreciably red-shifted and blue-shifted behaviors can be observed in Fig. 3(d) for quite a few molecules for the CASCI[EPA1]-PBE0/cc-pVDZ approach, showing that the active space determined by IEPA1 is not even useful for the traditional basis set.

The appropriate approach for classical CASCI method is to introduce the nature orbital. The accuracy of CASCI[NOON]-MP2NO/cc-pVDZ whose RMSD is 46.83 cm^{-1} [the two obvious outliers are H_2 and HCl, marked by the orange points in Fig. 3(f)] is significantly better than that of CASCI[MB]-PBE0/cc-pVDZ or CASCI[IEPA1]-PBE0/cc-pVDZ and is comparable to that of EDQC[IEPA1]-PBE0/Wavelet. Note that the size of the active space selected by NOON is slightly larger than that by IEPA1.

The successful application of the KS MOs in the proposed EDQC[IEPA1]-PBE0/Wavelet approach might not be duplicated in the traditional basis set. The comparison between the two classical methods, CCSD(T)-PBE0/cc-pVDZ (RMSD = 43.74 cm^{-1}) and CCSD(T)-HF/cc-pVDZ (RMSD = 41.73 cm^{-1}) shows that the use of KS MOs with traditional basis set does not give much benefit in this closed-shell dataset (see also the result of Δ(CCSD(T):RHF) in [46]). Nevertheless, we remark that previous studies showed some kind of improvement for radicals [43] and open-shell systems [46].

B. VQE(UCCSD) benchmark

For large molecular systems, the exact diagonalization method would not be feasible anymore. Instead, VQE is an algorithm that will be used practically on the near-term quantum computers. In Table II, the harmonic vibrational frequencies obtained by VQE with the chemistry-inspired UCCSD ansatz and the SLSQP optimizer for the [IEPA1]-PBE0/Wavelet approach are presented (the equilibrium bond lengths are presented in Ref. [37]). The results show that the VQE(UCCSD) approach can be as accurate as the exact diagonalization method except for the BeO family, the CO family, and some of the N_2 family.

Previous study [71] showed that for systems with strongly correlated electrons, UCCSD would not give results achieving chemical accuracy even in the region near the equilibrium (bond-length) point. In strongly correlated systems, the states resulting from the action of the UCCSD exponential operators that include only single and double coupled-cluster excitations might not encompass all those important configurations where the strongly correlated electrons would also be present. This motivates us to investigate the Mayer bond order [51], a good electron correlation descriptor applicable to multiconfigurational (strongly correlated) systems to quantify the degrees of bonding suitable for our analysis. That is, higher Mayer bond order indices correspond to more strongly correlated electrons. We calculate and show the Mayer bond order for the neutral closed-shell diatomic molecules in Ref. [37], and

TABLE II. Harmonic vibrational frequencies (in cm^{-1}) of neutral closed-shell diatomic molecules for VQE(UCCSD)[IEPA1]-PBE0/Wavelet. The approach EDQC[IEPA1]-PBE0/Wavelet is excerpted from Table I for the comparison. The value in the parenthesis denotes the difference in the results between VQE(UCCSD) and EDQC.

Mol.	EDQC[IEPA1] -PBE0/Wavelet	VQE(UCCSD)[IEPA1] -PBE0/Wavelet
H_2	[2,2]4391.45	[2,2]4391.44 (−0.01)
LiH	[2,3]1405.36	[2,3]1405.34 (−0.02)
NaH	[2,3]1143.76	[2,3]1143.89 (0.13)
BH	[4,7]2386.09	[4,7]2385.96 (−0.13)
AlH	[4,7]1709.33	[4,7]1712.38 (3.05)
GaH	[4,7]1593.80	[4,7]1594.50 (0.70)
HF	[2,3]4148.44	[2,3]4146.67 (−1.77)
HCl	[2,3]2874.66	[2,3]2874.69 (0.03)
HBr	[2,3]2511.84	[2,3]2511.84 (0.00)
LiF	[4,6] 938.00	[4,6] 937.94 (−0.06)
LiCl	[4,4] 649.26	[4,4] 649.26 (0.00)
LiBr	[6,5] 559.84	[6,5] 559.81 (−0.03)
NaF	[6,4] 548.70	[6,4] 548.69 (−0.01)
NaCl	[6,4] 362.88	[6,4] 362.89 (0.01)
NaBr	[6,4] 292.99	[6,4] 292.99 (0.00)
BeO	[6,6]1436.71	[6,6]1508.87 (72.16)
BeS	[6,6] 949.56	[6,6] 983.82 (34.26)
BF	[8,7]1462.33	[8,7]1460.42 (−1.91)
BCl	[8,6] 852.18	[8,6] 849.61 (−2.57)
BBr	[8,6] 686.03	[8,6] 683.69 (−2.34)
AlF	[8,7] 814.55	[8,7] 814.80 (0.25)
AlCl	[4,4] 474.55	[4,4] 474.55 (0.00)
AlBr	[4,4] 371.20	[4,4] 371.20 (0.00)
GaF	[8,6] 609.27	[8,6] 609.01 (−0.26)
GaCl	[4,4] 340.81	[4,4] 340.81 (0.00)
CO	[8,6]2248.86	[8,6]2336.76 (87.90)
CS	[8,7]1298.67	[8,7]1370.75 (72.08)
CSe	[8,7]1010.42	[8,7]1056.19 (45.77)
SiO	[8,7]1247.22	[8,7]1316.40 (69.18)
SiS	[8,8] 716.80	[8,8] 747.32 (30.52)
SiSe	[8,6] 582.23	[8,6] 613.98 (31.75)
GeO	[8,6] 999.61	[8,6]1066.79 (67.18)
N_2	[4,4]2399.42	[4,4]2402.52 (3.10)
PN	[4,4]1375.98	[4,4]1372.77 (−3.21)
P_2	[4,4] 778.53	[4,4] 798.40 (19.87)
AsN	[4,4]1088.54	[4,4]1055.19 (−33.35)
As_2	[4,4] 414.28	[4,4] 414.10 (−0.18)
Li_2	[2,5] 315.30	[2,5] 315.30 (0.00)
ClF	[10,7] 689.23	[10,7] 689.11 (−0.12)
Cl_2	[10,7] 504.52	[10,7] 504.56 (0.04)
BrF	[10,7] 625.71	[10,7] 625.79 (0.08)
BrCl	[10,7] 400.00	[10,7] 400.03 (0.03)
Br_2	[8,6] 288.04	[8,6] 288.16 (0.12)
RMSD	41.44	51.62
MSD	−9.75	1.70
MAD	28.53	35.37

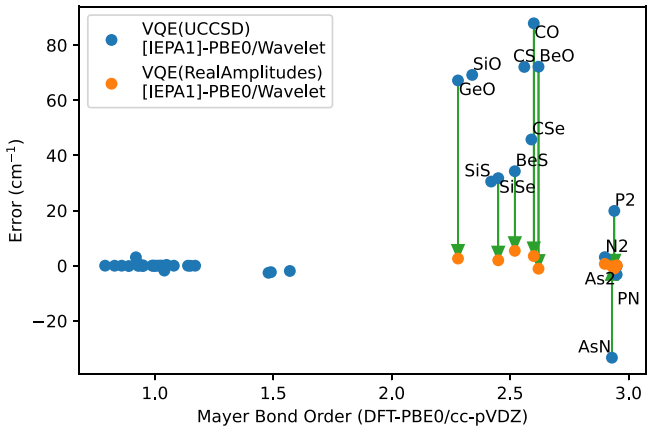

FIG. 4. Mayer bond order indices calculated by DFT-PBE0/cc-pVDZ versus the error (difference) in the harmonic vibrational frequencies (in cm^{-1}) calculated by VQE(UCCSD)[IEPA1]-PBE0/Wavelet with respect to those by EDQC[IEPA1]-PBE0/Wavelet for the diatomic molecule in the benchmark dataset. The orange dots denote the vibrational frequency results for the specified molecules calculated by VQE(RealAmplitudes)[IEPA1]-PBE0/Wavelet. The relation between blue and orange dots is from Table III, and the green arrows point toward the directions of improvement from the UCCSD ansatz to the RealAmplitudes ansatz.

the trend of the Mayer bond order is similar for different XC functionals considered here. We choose the Mayer bond order indices calculated by DFT-PBE0/cc-pVDZ (to be consistent with the XC functional and the traditional basis set used in this work) to present the relation with the harmonic vibrational frequencies calculated by VQE(UCCSD) in Fig. 4. As clearly indicated in Fig. 4, systems for which UCCSD does not yield accurate harmonic vibrational frequencies correspond to those whose Mayer bond order indices are larger than 2, which are the CO family, the N_2 family, and the BeO family (the blue points in the region where the Mayer bond order indices >2 in Fig. 4), notably for the BeO family as they are traditionally thought of as the single-bond molecules. We conclude that the index of the Mayer bond order larger than 2 is a good descriptor to indicate that UCC truncated to SD might be too restricted to describe the harmonic vibrational frequency accurately.

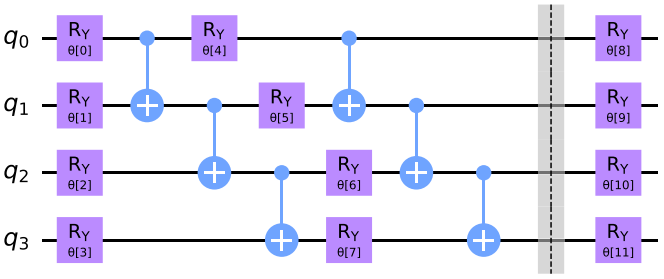

FIG. 5. Typical quantum circuit of the hardware-efficient RealAmplitudes ansatz. For simplicity, a four-qubit two-local quantum circuit with two repeated unit pattern circuits is shown. The unit pattern circuit consists of a layer of parameterized R_Y rotational gates applied on all qubits and a layer of CNOT gates in linear entanglement. For clarity, the barrier is added to separate the repeated unit pattern circuits from the final rotation layer and the Hartree-Fock initial state is skipped to draw. The rotational angles θs in the R_Y gates denote the tunable circuit parameters.

C. VQE(UCCSD) versus VQE(RealAmplitudes)

For those systems whose Mayer bond order indices are larger than 2, we then consider a heuristic hardware-efficient ansatz, the RealAmplitudes ansatz (see Fig. 5) implemented in QISKIT [61], since it can go beyond the restriction of the accessible Hilbert space of the chemistry-inspired UCCSD ansatz. That is, the hardware-efficient ansatz can increase its expressibility [72] by increasing the number of repeated unit pattern circuit consisting of a layer of parameterized R_Y rotational gates and a layer of entanglement circuit with

TABLE III. Comparisons of the harmonic vibrational frequencies and the relevant circuit information between the VQE(UCCSD) and VQE(RealAmplitudes) calculations using the Hamiltonian in the [IEPA1]-PBE0/Wavelet approach for the systems whose Mayer bond order indices are larger than 2 with the required number of qubits up to 10. The value inside the parenthesis in the harmonic vibrational frequency denotes the difference between VQE and EDQC.

Mol.	VQE(UCCSD) [IEPA1]-PBE0/Wavelet	Dep[a]	Nθ[b]	State Fidelity[c]	VQE(RealAmplitudes) [IEPA1]-PBE0/Wavelet	Dep[a] (Rep[d])	Nθ[b]	State Fidelity[c]
BeO	[6,6]1508.87 (72.15)	10914	117	0.98634	[6,6]1435.63 (−1.09)	99(30)	310	0.9990823
BeS	[6,6] 983.82 (34.27)	10914	117	0.99086	[6,6] 955.98 (5.39)	99(30)	310	0.9981542
CO	[8,6]2336.76 (87.90)	8460	92	0.99729	[8,6]2249.18 (3.49)	84(25)	260	0.9999526
SiSe	[8,6] 613.98 (31.75)	8460	92	0.99515	[8,6] 584.19 (1.96)	99(30)	310	0.9996153
GeO	[8,6]1066.79 (67.18)	8460	92	0.99068	[8,6]1012.25 (2.59)	84(25)	260	0.9988251
N_2	[4,4]2402.52 (3.10)	1480	26	0.99991	[4,4]2400.07 (0.65)	35(10)	66	0.9999982
PN	[4,4]1372.77 (−3.21)	1480	26	0.99775	[4,4]1376.12 (0.14)	35(10)	66	0.9999997
P_2	[4,4] 798.40 (19.87)	1480	26	0.99988	[4,4] 777.54 (−0.99)	35(10)	66	0.9999974
AsN	[4,4]1055.19 (−33.35)	1480	26	0.99785	[4,4]1088.34 (−0.20)	35(10)	66	0.9999982
As_2	[4,4] 414.10 (−0.18)	1480	26	0.99971	[4,4] 414.31 (0.03)	35(10)	66	0.9999996

[a]Dep denotes the circuit depths.
[b]Nθ denotes the number of tunable circuit parameters.
[c]State fidelity denotes the average state fidelity for the molecular distance points employed to calculate the vibrational frequency.
[d]Rep denotes the number of repetitions of the unit pattern circuit.

two-qubit entangling gates (see Fig. 5). During the procedure of increasing the number of repetitions, the results from the circuits with each number of repetitions will be inspected in order to evaluate and decide what the best circuit depth and number of tunable parameters for the molecule under consideration would be. Meanwhile, the performance of such non-chemistry-inspired ansatz is accomplished through the heuristic search on the parameter space.

In Table III, the results of the harmonic vibrational frequencies between the UCCSD ansatz and the RealAmplitudes ansatz with linear entanglement are compared for the systems whose Mayer bond order indices are larger than 2 and whose required numbers of qubits are not greater than 10 due to the computational time cost consideration. In order to accurately measure the degrees of errors and directly distinguish the accuracies achieved by the UCCSD ansatz and the RealAmplitudes ansatz in VQE, we calculate the state fidelity (defined as $|\langle\Phi_1|\Phi_2\rangle|^2$ for pure states $|\Phi_1\rangle$ and $|\Phi_2\rangle$) to measure how much overlap between the exact wave function obtained from the exact diagonalization method and the wave function obtained from the UCCSD ansatz and from the RealAmplitudes ansatz, respectively. We use 10 repetitions of the unit pattern circuit for six-qubit molecular systems and 25 or 30 repetitions for ten-qubit systems in the RealAmplitudes ansatz to achieve the desired accuracy. As one of the superior advantage of the hardware-efficient ansatz, the circuit depth of the RealAmplitudes ansatz, which is 84 or 99 for ten-qubit systems, is significantly shallower than the circuit depth of the UCCSD ansatz, which is 8460 or 10914 for the corresponding ten-qubit systems. Despite having shallower circuit depths, the RealAmplitudes ansatz could still achieve higher state fidelities than the UCCSD ansatz, and for the cases with a small number of qubits outstanding performance can be achieved. Note that the energy differences between the result obtained by the exact diagonalization and that by the RealAmplitudes ansatz for the molecules are all within the chemical accuracy, which is not true when the UCCSD ansatz is used. This is a clear indication that a heuristic hardware-efficient quantum circuit can span a state space larger than that spanned by the UCCSD method, a sign of polynomial quantum advantage. To the best of our knowledge, our investigation is the first systematical benchmark study to demonstrate that a heuristic hardware-efficient ansatz could outperform a chemistry-inspired UCCSD ansatz in predicting accurate molecular properties by quantum computation.

At this moment, let us recapitulate the performance of quantum computing compared with gold standard method in classical computing. VQE(UCCSD)[IEPA1]-PBE0/Wavelet yields less accurate results compared with CCSD(T)-HF/cc-pVDZ (see Table II and Table I), showing that using the UCCSD ansatz in quantum computation might not be preferable. However, VQE(RealAmplitudes)[IEPA1]-PBE0/Wavelet yields results basically very close or equivalent to the CCSD(T)-HF/cc-pVDZ method. This is especially notable for difficult-case molecules with Mayer bond order indices larger than 2 and whose required numbers of qubits are not greater than 10. In these cases, the quantum computing results could be as good as the best achievable accuracy obtained from the EDQC[IEPA1]-PBE0/Wavelet. Therefore quantum computation could achieve accuracy comparable with CCSD(T)-HF/cc-pVDZ, while, in terms of computational resources, the quantum algorithm benefited from our approach requires much smaller number of Hamiltonian evaluations compared with CCSD(T)-HF/cc-pVDZ.

We, however, note here that for the molecular systems using the same number of qubits, the number of tunable parameters for the RealAmplitudes ansatz is more than that for the UCCSD ansatz. For example, for the ten-qubit systems shown in Table III, the number of tunable parameters for the RealAmplitudes ansatz is 260 or 310, while it is 92 or 117 for the UCCSD ansatz. Besides, the state fidelities of the RealAmplitudes ansatz become slightly lower when the number of qubits becomes larger. For a system using a large number of qubits for quantum computation, it is necessary to increase the number of repeated circuit layers of the parameterized and entanglement gates in order to obtain sufficient expressibility. This could lead to a circuit with many parameters, making the optimization difficult as the initial values of the parameters would potentially critically affect the optimization result. McClean *et al.* [73] have shown that the optimization process of random initialization would be stuck in the local trap due to the barren plateaus. Contrarily, the zero initialization (all the values of the tunable parameters are initialized to zeros) would give better results in most situations. Specifically for the optimization procedure, we apply a combination of the SLSQP and L-BFGS-B optimizers, where the former converges faster but less accurate than the latter. In the first stage, we consider the zero initialization of the tunable parameters and then also add small random numbers to them for the SLSQP optimizer to reach the converged parameters. Then, in the second stage, the converged parameters with relatively high state fidelities from SLSQP are taken as the initial parameters for L-BFGS-B to find the optimal parameter values.

Since we aim to calculate the vibrational frequencies derived from the curvature around the equilibrium geometry of the PEC, we need to calculate the molecular ground state energies on a set of points (different distances between the atoms) around the equilibrium (bond-length) point to construct the PEC. Note that in order to obtain accurate vibrational frequency for a molecule, an extremely high state fidelity might not be required to obtain a very accurate ground state energy; instated, the state fidelities or more precisely the ground state energies should have consistent correlated accuracy on all the points, which would give a parallel constant energy shift with respect to the reference PEC. If achieving chemical accuracy is the only requirement, then the circuit could become even shallower to yield results within it. However, if the state fidelities are not high enough, one of the drawback of the non-chemistry-inspired ansatz is that each point is optimized independently so the optimized energy points on the PEC behave like no appreciable correlation. To enhance the correlation between different molecular distance points, after the second optimization stage, if the optimization on one of the molecular distance points yields a distinctively lower molecular ground state energy, its converged parameters are taken as the initial parameters for the other points to optimize further until the variance of the energy for each point is small. Therefore the decision on the choice of the number of the repeated circuit layers for the RealAmplitudes ansatz depends not only on the state fidelity of a single point but also on the

variances of energies on all the points used to calculate the vibrational frequency.

We remark here that even for the same size of the RealAmplitudes circuit, different molecular systems (or different Hamiltonians) yield different degrees of state fidelities. This comes from the differences in structure and coefficients of the weighted Pauli terms in the Hamiltonians so that the optimizer may favor some cases. On the other hand, in addition to the number of repeated circuit layers, the types of the entangling gates or/and entanglement circuit structures would also affect the results. The advantage and disadvantage of more (or less) entanglement for different Hamiltonians remain to be clarified. The approach we present here is just a way to achieve high state fidelities, so potentially more efficient methods for large systems should be studied in the future.

IV. CONCLUSION AND OUTLOOK

We propose a quantum computational approach that combines KS MOs expanded in the Daubechies wavelet basis set and an optimal active space determined by IEPA1 energy criterion, resulting in a significantly reduced qubit number requirement (2 to 12 versus 20 to 60 required by cc-pVDZ with frozen core approximation) while maintaining excellent accuracy compared to the experimental data. We validate the approach by benchmarking its performance on the harmonic vibrational frequencies of 43 neutral closed-shell diatomic molecules. The RMSD is small by using the Daubechies wavelets basis set and the error is further decreased by using KS MOs with higher-level XC functionals. The best approach here is EDQC[IEPA1]-PBE0/Wavelet with performance comparable to CASCI[NOON]-MP2NO/cc-pVDZ and CCSD(T)-HF/cc-pVDZ. The results obtained by this EDQC[IEPA1]-PBE0/Wavelet approach, considered as the best achievable results by quantum computation, are in great agreement with the experimental data. In contrast, for the traditional basis set on this closed-shell dataset, the same approach, e.g., CASCI[IEPA1]-PBE0/cc-pVDZ, could not provide noticeable improvements.

For larger systems, the exact diagonalization method is unfeasible due to the exponential scaling of the problem with the system size and thus a quantum computing approach is required. So we conduct VQE quantum computations of the vibrational frequencies of the 43 neutral closed-shell diatomic molecules using the Hamiltonians constructed from the approach of [IEPA1]-PBE0/Wavelet with the chemistry-inspired UCCSD ansatz. The results shows that the VQE(UCCSD) approach can be as accurate as the exact diagonalization method except for systems whose Mayer bond order indices are larger than 2. Then for those systems, we demonstrate that a hardware-efficient RealAmplitudes ansatz can provide significant improvements over the UCCSD ansatz. This indicates that the hardware-efficient ansatz can avoid the restriction on the accessible Hilbert space by the chemistry-inspired UCCSD ansatz. At the same time, the appealing feature of the shallow circuit of the hardware-efficient ansatz will make quantum computation of accurate vibrational frequencies on the near-term NISQ devices realizable.

Based on the improvement of the results with the increasing hierarchy of the DFT XC functionals in this work, it is reasonable to expect that more promising results by quantum computing would be obtained by using higher-level XC functionals while keeping the number of qubits significant reduced. On the other hand, it is interesting to note that the XC functionals used here are developed for DFT and it might be worth trying to design XC functionals for KS MOs for the WFT calculation such that the accuracy and the efficiency of the calculation could be further improved reaching another level.

In summary, this benchmark study validates novel means to achieve highly-accurate calculations of molecular properties on quantum computers with significantly reduced qubit resources. Furthermore, VQE calculations with chemistry-inspired UCCSD and heuristic hardware-efficient ansatzes are compared to demonstrate the advantage of the heuristic ansatz in complex chemical bonding systems. Our calculations show that a quantum computer capable of carrying out calculations on $\leqslant 10$ qubits with circuit depth < 100 can accurately predict the vibrational frequencies of neutral closed-shell diatomic molecules, and these quantum resource requirements should be able to be achieved on near-term NISQ devices. In fact, according to the so-called 100×100 Challenge in the announcement of the IBM Quantum Summit 2022, IBM Quantum plans to offer a tool able to calculate unbiased (noiseless) observables of circuits with 100 qubits and depth-100 gate operations in a reasonable runtime in 2024 [50]. Our benchmark investigation here provides a critical assessment on the power of quantum computation of molecular properties and insights on further improvements.

ACKNOWLEDGMENTS

J.P.C. gratefully acknowledges the financial support from the National Science and Technology Council, Taiwan (NSTC 109-2112-M-018-008-MY3). Y.C.C. thanks the National Science and Technology Council, Taiwan (Grant No. NSTC 112-2119-M-002-018 and NSTC 111-2113-M-002-017), Physics Division, National Center for Theoretical Sciences (Grant No. 110-2124-M-002-012), and National Taiwan University (Grant No. 111L894603) for financial support. Y.C.C. is grateful to Computer and Information Networking Center, National Taiwan University for the support of high-performance computing facilities. A.H. gratefully acknowledges the sponsorship from Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11200120), City University of Hong Kong (Projects No. 7005615, No. 7006103), CityU Seed Fund in Microelectronics (Project No. 9229135), and Hong Kong Institute for Advanced Study, City University of Hong Kong (Project No. 9360157). H.-S.G. acknowledges support from the National Science and Technology Council, Taiwan under Grants No. NSTC 112-2119-M-002-014, No. NSTC 111-2119-M-002-006-MY3, No. NSTC 111-2119-M-002-007, No. NSTC 110-2627-M-002-002, No. NSTC 111-2627-M-002-001, and No. NSTC 111-2627-M-002-006, from the US Air Force Office of Scientific Research under Award Number FA2386-23-1-4052, and from the National Taiwan University under Grant No. NTU-CC-112L893404. H.-S.G. and Y.C.C. are also grateful for the support from the "Center for Advanced Computing and

Imaging in Biomedicine (NTU-112L900702)" through The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE), Taiwan, and the support from the Physics Division, National Center for Theoretical Sciences, Taiwan.

[1] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya *et al.*, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).

[2] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[3] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke *et al.*, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[4] W. Kutzelnigg, Pair correlation theories, in *Methods of Electronic Structure Theory*, Modern Theoretical Chemistry, Vol. 3, edited by H. F. Schaefer (Plenum Press, New York, 1977), pp. 129–188.

[5] M. R. Hoffmann and J. Simons, A unitary multiconfigurational coupled-cluster method: Theory and applications, J. Chem. Phys. **88**, 993 (1988).

[6] R. J. Bartlett, S. A. Kucharski, and J. Noga, Alternative coupled-cluster ansätze II. The unitary coupled-cluster method, Chem. Phys. Lett. **155**, 133 (1989).

[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] M.-H. Yung, J. Casanova, A. Mezzacapo, J. Mcclean, L. Lamata, A. Aspuru-Guzik, and E. Solano, From transistor to trapped-ion computers for quantum chemistry, Sci. Rep. **4**, 3589 (2014).

[9] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, Quantum Sci. Technol. **4**, 014008 (2018).

[10] A. Anand, P. Schleich, S. Alperin-Lea, P. W. Jensen, S. Sim, M. Díaz-Tinoco, J. S. Kottmann, M. Degroote, A. F. Izmaylov, and A. Aspuru-Guzik, A quantum computing view on unitary coupled cluster theory, Chem. Soc. Rev. **51**, 1659 (2022).

[11] P. Jordan and E. Wigner, Ìber das paulische Äquivalenzverbot, Z. Phys. **47**, 631 (1928).

[12] A. G. Taube and R. J. Bartlett, New perspectives on unitary coupled-cluster theory, Int. J. Quantum Chem. **106**, 3393 (2006).

[13] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, A fifth-order perturbation comparison of electron correlation theories, Chem. Phys. Lett. **157**, 479 (1989).

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

[15] 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).

[16] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo *et al.*, Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions, Phys. Rev. A **98**, 022322 (2018).

[17] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm, npj Quantum Inf. **6**, 10 (2020).

[18] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio *et al.*, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

[19] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev, VQE method: A short survey and recent developments, Mater. Theory **6**, 2 (2022).

[20] 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).

[21] W. J. Hehre, R. F. Stewart, and J. A. Pople, Self-consistent molecular-orbital methods. I. Use of gaussian expansions of slater-type atomic orbitals, J. Chem. Phys. **51**, 2657 (1969).

[22] T. H. Dunning, Jr., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. **90**, 1007 (1989).

[23] D. E. Woon and T. H. Dunning, Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon, J. Chem. Phys. **98**, 1358 (1993).

[24] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton, J. Chem. Phys. **110**, 7667 (1999).

[25] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson, Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg, Theor. Chem. Acc. **128**, 69 (2011).

[26] R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan, and G. Beylkin, Multiresolution quantum chemistry: Basic theory and initial applications, J. Chem. Phys. **121**, 11587 (2004).

[27] R. J. Harrison *et al.*, MADNESS: A multiresolution, adaptive numerical environment for scientific simulation, SIAM J. Sci. Comput. **38**, S123 (2016).

[28] J. S. Kottmann, F. A. Bischoff, and E. F. Valeev, Direct determination of optimal pair-natural orbitals in a real-space representation: The second-order Moller–Plesset energy, J. Chem. Phys. **152**, 074105 (2020).

[29] J. S. Kottmann, P. Schleich, T. Tamayo-Mendoza, and A. Aspuru-Guzik, Reducing qubit requirements while maintaining numerical precision for the variational quantum eigensolver: A basis-set-free approach, J. Phys. Chem. Lett. **12**, 663 (2021).

[30] C.-L. Hong, T. Tsai, J.-P. Chou, P.-J. Chen, P.-K. Tsai, Y.-C. Chen, E.-J. Kuo, D. Srolovitz, A. Hu, Y.-C. Cheng, and H.-S.

Goan, Accurate and efficient quantum computations of molecular properties using daubechies wavelet molecular orbitals: A benchmark study against experimental data, PRX Quantum **3**, 020360 (2022).

[31] I. Daubechies, *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, 1992).

[32] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, and R. Schneider, Daubechies wavelets as a basis set for density functional pseudopotential calculations, J. Chem. Phys. **129**, 014109 (2008).

[33] L. Genovese, B. Videau, M. Ospici, T. Deutsch, S. Goedecker, and J.-F. Méhaut, Daubechies wavelets for high performance electronic structure calculations: The BIGDFT project, C.R. Mec. **339**, 149 (2011).

[34] S. Mohr, L. E. Ratcliff, P. Boulanger, L. Genovese, D. Caliste, T. Deutsch, and S. Goedecker, Daubechies wavelets for linear scaling density functional theory, J. Chem. Phys. **140**, 204110 (2014).

[35] L. E. Ratcliff, W. Dawson, G. Fisicaro, D. Caliste, S. Mohr, A. Degomme, B. Videau, V. Cristiglio, M. Stella, M. D'Alessandro *et al.*, Flexibilities of wavelets as a computational basis set for large-scale electronic structure calculations, J. Chem. Phys. **152**, 194110 (2020).

[36] I. Shavitt, The method of configuration interaction, in *Methods of Electronic Structure Theory*, Modern Theoretical Chemistry, Vol. 3, edited by H. F. Schaefer (Plenum Press, New York, 1977), pp. 189–275.

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.5.043216 for supplemental data on harmonic vibrational frequencies calculated by the minimal basis (MB) set approach and density functional theory (DFT)-XC/Wavelet with XC functional being HF, PBE or PBE0, on equilibrium bond lengths calculated by all the benchmark methods, and on the Mayer bond order indices calculated by DFT-XC/cc-pVDZ.

[38] A. Khedkar and M. Roemelt, Active space selection based on natural orbital occupation numbers from *n*-electron valence perturbation theory, J. Chem. Theory Comput. **15**, 3522 (2019).

[39] S. Grimme, Density functional calculations with configuration interaction for the excited states of molecules, Chem. Phys. Lett. **259**, 128 (1996).

[40] S. Grimme and M. Waletzke, A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods, J. Chem. Phys. **111**, 5645 (1999).

[41] P. Bouř, Configuration interaction with Kohn–Sham orbitals and their relation to excited electronic states, Chem. Phys. Lett. **345**, 331 (2001).

[42] L. Veseth, Molecular excitation energies computed with Kohn–Sham orbitals and exact exchange potentials, J. Chem. Phys. **114**, 8789 (2001).

[43] G. J. O. Beran, S. R. Gwaltney, and M. Head-Gordon, Approaching closed-shell accuracy for radicals using coupled cluster theory with perturbative triple substitutions, Phys. Chem. Chem. Phys. **5**, 2488 (2003).

[44] J. Kim, K. Hong, S. Choi, S.-Y. Hwang, and W. Y. Kim, Configuration interaction singles based on the real-space numerical grid method: Kohn–Sham versus Hartree–Fock orbitals, Phys. Chem. Chem. Phys. **17**, 31434 (2015).

[45] J. Lim, S. Choi, J. Kim, and W. Y. Kim, Outstanding performance of configuration interaction singles and doubles using exact exchange Kohn-Sham orbitals in real-space numerical grid method, J. Chem. Phys. **145**, 224309 (2016).

[46] L. W. Bertels, J. Lee, and M. Head-Gordon, Polishing the gold standard: The role of orbital choice in CCSD(T) vibrational frequency prediction, J. Chem. Theory Comput. **17**, 742 (2021).

[47] O. Sinanoğlu, Many-electron theory of atoms, molecules and their interactions, in *Advances in Chemical Physics*, Vol. VI (John Wiley & Sons, Ltd, 1964), pp. 315–412.

[48] R. K. Nesbet, Electronic correlation in atoms and molecules, in *Advances in Chemical Physics*, Vol. IX (John Wiley & Sons, Ltd, 1965), pp. 321–363.

[49] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, Nature (London) **618**, 500 (2023).

[50] IBM Quantum, The IBM Quantum State of the Union 2022: The 100 × 100 Challenge, https://research.ibm.com/blog/next-wave-quantum-centric-supercomputing (2022).

[51] I. Mayer, Charge, bond order and valence in the AB initio SCF theory, Chem. Phys. Lett. **97**, 270 (1983).

[52] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).

[53] C. Adamo and V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, J. Chem. Phys. **110**, 6158 (1999).

[54] A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. **98**, 5648 (1993).

[55] S. Goedecker, M. Teter, and J. Hutter, Separable dual-space gaussian pseudopotentials, Phys. Rev. B **54**, 1703 (1996).

[56] C. Hartwigsen, S. Goedecker, and J. Hutter, Relativistic separable dual-space Gaussian pseudopotentials from H to Rn, Phys. Rev. B **58**, 3641 (1998).

[57] M. Krack, Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals, Theor. Chem. Acc. **114**, 145 (2005).

[58] B. Natarajan, L. Genovese, M. E. Casida, T. Deutsch, O. N. Burchak, C. Philouze, and M. Y. Balakirev, Wavelet-based linear-response time-dependent density-functional theory, Chem. Phys. **402**, 29 (2012).

[59] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Inc. Mineola, New York, 1996).

[60] C. Møller and M. S. Plesset, Note on an approximation treatment for many-electron systems, Phys. Rev. **46**, 618 (1934).

[61] J. Gambetta *et al.*, Qiskit/qiskit: QISKIT 0.24.0, https://zenodo.org/records/4582797 (2021).

[62] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, Ann. Phys. **298**, 210 (2002).

[63] J. T. Seeley, M. J. Richard, and P. J. Love, The Bravyi-Kitaev transformation for quantum computation of electronic structure, J. Chem. Phys. **137**, 224109 (2012).

[64] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, Tapering off qubits to simulate fermionic Hamiltonians, arXiv:1701.08213.

[65] P. Virtanen *et al.*, SCIPY 1.0: Fundamental algorithms for scientific computing in python, Nature Methods **17**, 261 (2020).

[66] Q. Sun *et al.*, PYSCF: The python-based simulations of chemistry framework, WIREs Comput. Mol. Sci. **8**, e1340 (2018).

[67] B. O. Roos, P. R. Taylor, and P. E. Sigbahn, A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach, Chem. Phys. **48**, 157 (1980).

[68] P.-O. Löwdin, Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction, Phys. Rev. **97**, 1474 (1955).

[69] R. D. Johnson III, NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, http://cccbdb.nist.gov (2022).

[70] M. K. Kesharwani, B. Brauer, and J. M. Martin, Frequency and zero-point vibrational energy scale factors for double-hybrid density functionals (and other selected methods): Can anharmonic force fields be avoided?, J. Phys. Chem. A **119**, 1701 (2015).

[71] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, Generalized unitary coupled cluster wave functions for quantum computation, J. Chem. Theory Comput. **15**, 311 (2019).

[72] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms, Adv. Quantum Technol. **2**, 1900070 (2019).

[73] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 4812 (2018).