VQE method: a short survey and recent developments

Dmitry A. Fedorov[1,3]* , Bo Peng[2]*, Niranjan Govind[2]* and Yuri Alexeev[3]*

*Correspondence: fedorov@anl.gov;
peng398@pnnl.gov;
niri.govind@pnnl.gov; yuri@anl.gov
[1]Oak Ridge Associated Universities,
Oak Ridge 37830, Tennessee, USA
[2]Physical and Computational
Sciences Directorate, Pacific
Northwest National Laboratory,
Richland 99354, Washington, USA
Full list of author information is
available at the end of the article

Abstract

The variational quantum eigensolver (VQE) is a method that uses a hybrid quantum-classical computational approach to find eigenvalues of a Hamiltonian. VQE has been proposed as an alternative to fully quantum algorithms such as quantum phase estimation (QPE) because fully quantum algorithms require quantum hardware that will not be accessible in the near future. VQE has been successfully applied to solve the electronic Schrödinger equation for a variety of small molecules. However, the scalability of this method is limited by two factors: the complexity of the quantum circuits and the complexity of the classical optimization problem. Both of these factors are affected by the choice of the variational ansatz used to represent the trial wave function. Hence, the construction of an efficient ansatz is an active area of research. Put another way, modern quantum computers are not capable of executing deep quantum circuits produced by using currently available ansatzes for problems that map onto more than several qubits. In this review, we present recent developments in the field of designing efficient ansatzes that fall into two categories—chemistry–inspired and hardware–efficient—that produce quantum circuits that are easier to run on modern hardware. We discuss the shortfalls of ansatzes originally formulated for VQE simulations, how they are addressed in more sophisticated methods, and the potential ways for further improvements.

Keywords: VQE, Chemistry-inspired ansatz, Hardware-efficient ansatz, Unitary coupled cluster, Quantum computing, Quantum chemistry

Introduction

Quantum simulation of chemistry and materials is an important application for classical computing. For example, one can predict the rates of chemical reactions, determine molecular structure, and the properties of materials and molecules. It is achieved through the solution of the electronic structure Hamiltonian, which allows one to describe the properties of interacting electrons in the presence of stationary nuclei. The advent of quantum computing holds promise for significantly speeding up these calculations, which currently are done classically. This idea stems from the Feynman's postulate in the famous paper "Simulating Physics with Computers," published in 1982 (Feynman 1982), that to simulate quantum systems, one would need to build quantum computers to perform quantum computations, and it is much more efficient than doing it classically for

correlated systems. Indeed, one cannot simply scale the calculations on classical computers or apply massive parallelism because of the exponential growth of computational requirements with the size of a chemical system and complexity.

These ideas look attractive in theory, but in reality, there are severe limitations in the computational capabilities of the currently available small quantum devices, which are often referred to as NISQ (noisy intermediate-scale quantum) devices (Preskill 2018). A second important aspect is the inefficiency of existing quantum algorithms in terms of the resources that are needed to solve any useful problem on a quantum computer faster than on a classical computer, which is often discussed in terms of the quantum advantage. For example, the resources estimated to perform chromium dimer calculations, which is a large enough molecular system to demonstrate the quantum advantage on a quantum computer, with existing algorithms require at least 1 million physical qubits (Elfving et al. 2020; Liu et al. 2021) to make 60 virtual high-fidelity qubits. At the same time, these physical qubits need higher fidelity than do existing ones for error correction algorithms to work. This is far from the currently available 71 physical qubit quantum computers and will remain the case for the near future.

Another attractive idea is to use quantum computing on quantum devices to improve quantum technology. Such a use will, in theory, allow systematic improvement of the technology. In fact, it is one of the research objectives of the national quantum center Q-NEXT (https://www.q-next.org/), which is tasked with exploring the potential of this approach and developing new generations of methods.

The first algorithm that was proposed to solve the Schrödinger equation on a quantum computer was the quantum phase estimation (QPE) (Kitaev 1995; Abrams and Lloyd 1999; 1997). It is a fully quantum algorithm that can extract the phase or eigenvalues of a unitary operator. By using the phase kickback trick and inverse quantum Fourier transform (IQFT), (Shor 1994) QPE obtains the binary representation of the phase or eigenvalue. The QPE algorithm is versatile and is a part of other quantum algorithms, such as Shor's algorithm. QPE can achieve exponential speedup in finding the eigenspectrum of unitary operators, such as the electronic Hamiltonian, as long as an appropriate trial state with nonzero overlap with the real solution can be prepared. It is likely that quantum advantage can be demonstrated by using the QPE method when the first big enough fault-tolerant quantum computer is built (Aharonov and Ben-Or 1996). However, the problem with QPE, and IQFT in particular, is that it requires millions of qubits and gates even for relatively small systems, a requirement far beyond the capabilities of present NISQ hardware.

To address this problem, a variational hybrid quantum method VQE was proposed (Peruzzo et al. 2014; McClean et al. 2016; Romero et al. 2018). VQE is designed to utilize both quantum and classical resources to find variational solutions to eigenvalue problems. In a typical setup, the ground state trial wave function of the molecule is constructed from operators generating single- and double-excitation configurations from a Hartree-Fock wave function precomputed on a classical computer (UCCSD (Pal 1984; Hoffmann and Simons 1988; Kutzelnigg 1991; Taube and Bartlett 2006; Sur et al. 2008; Cooper and Knowles 2010; Harsha et al. 2018; Evangelista et al. 2019) ansatz). Next, on a quantum computer, the trial wave function is prepared and the expectation value of the Hamiltonian is measured. Then, the parameters of the trial wave function are

optimized iteratively on a classical computer using the variational principle. Although VQE simulations for small molecules have been performed on various quantum hardware architectures (Peruzzo et al. 2014; Kandala et al. 2017; O'Malley et al. 2016; Nam et al. 2020; Colless et al. 2018; Ryabinkin et al. 2018; McCaskey et al. 2019; Collaborators∗† et al. 2020; Hempel et al. 2018; Shen et al. 2017; Santagati et al. 2018; Gao et al. 2021; Gao et al. 2020; Smart and Mazziotti 2019), major advances are needed to scale this approach to larger systems. VQE also suffers from the large size of the resulting quantum circuits (albeit not as large as QPE circuits) coupled with the need to perform a classical optimization on a large number of variational parameters, which can render calculations intractable.

Trying to solve the electronic Schrödinger equation on a quantum computer translates into massive quantum circuits, which are beyond the reach of NISQ computers. In the rest of the paper, we will look in detail at different approaches, which aim to reduce the size of circuits while maintaining high accuracy. Our paper is not a comprehensive review of all VQE methods. Several comprehensive reviews covering VQE have been published recently (McArdle and Endo 2020; Cao et al. 2019; Cerezo et al. 2021; Bharti et al. 2021). The goal of this work is to present recent advances in methods that produce shorter VQE circuits or provide higher accuracy. In addition, we discuss these methods in more detail from a perspective of how they improve on the standard UCCSD and hardware–efficient ansatzes.
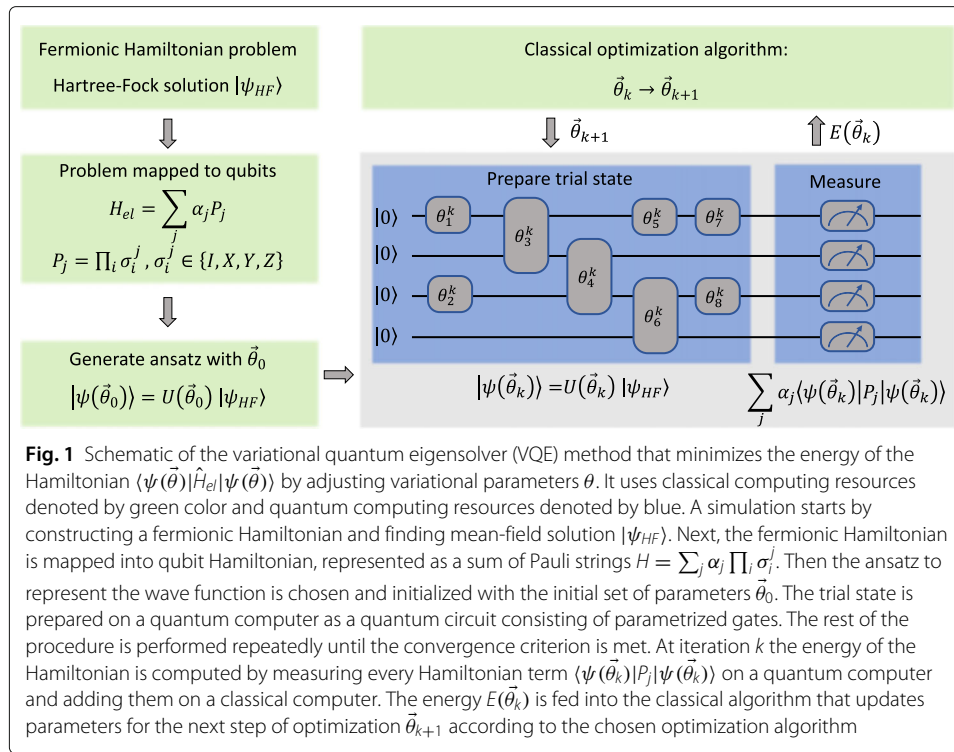
The rest of the paper is organized as follows. In Section 2 we provide a brief description of the VQE method, required for understanding the rest of the paper, and introduce the concepts of chemistry-inspired and hardware-efficient ansatzes. In Section 3 we discuss the chemically inspired ansatzes for VQE simulations. Section 4 is devoted to hardware–efficient ansatzes. In Section 5 the conclusions are presented.

Introduction to VQE

The VQE method was introduced to mitigate the significant hardware demands needed by the QPE approach on NISQ devices. VQE is a hybrid quantum-classical algorithm, where the computational workload is shared between the classical and quantum components of the hardware Fig. 1. It starts with a reasonable assumption about the form of the target wave function. The most common choice is to represent a wave function in a basis of atom-centered Gaussian basis functions. However, plane wave basis sets (Babbush et al. 2018) can be used as well as the recently proposed "basis-set free" approach (Kottmann et al. 2021). A trial wave function or ansatz is constructed with adjustable parameters, followed by the design of a quantum circuit capable of realizing this ansatz. The ansatz parameters are then variationally adjusted until the expectation value of the electronic Hamiltonian

$$E \leq \frac{\left\langle \psi(\vec{\theta}) \left| \hat{H}_{el} \right| \psi(\vec{\theta}) \right\rangle}{\left\langle \psi(\vec{\theta}) \middle| \psi(\vec{\theta}) \right\rangle} \tag{1}$$

is minimized. In Eq. 1 $|\psi(\vec{\theta})\rangle$ is the trial wave function that depends on the vector of variational parameters $\vec{\theta}$; E is the ground state energy of $\hat{H}_{el}$, an electronic Hamiltonian most commonly written in the second quantized form, although the first quantization representation has also been considered (Babbush et al. 2019). In this work, we focus on the second quantized form of the Hamiltonian

Fig. 1 Schematic of the variational quantum eigensolver (VQE) method that minimizes the energy of the Hamiltonian $\langle\psi(\vec{\theta})|\hat{H}_{el}|\psi(\vec{\theta})\rangle$ by adjusting variational parameters θ. It uses classical computing resources denoted by green color and quantum computing resources denoted by blue. A simulation starts by constructing a fermionic Hamiltonian and finding mean-field solution $|\psi_{HF}\rangle$. Next, the fermionic Hamiltonian is mapped into qubit Hamiltonian, represented as a sum of Pauli strings $H = \sum_j \alpha_j \prod_i \sigma_i^j$. Then the ansatz to represent the wave function is chosen and initialized with the initial set of parameters $\vec{\theta}_0$. The trial state is prepared on a quantum computer as a quantum circuit consisting of parametrized gates. The rest of the procedure is performed repeatedly until the convergence criterion is met. At iteration k the energy of the Hamiltonian is computed by measuring every Hamiltonian term $\langle\psi(\vec{\theta}_k)|P_j|\psi(\vec{\theta}_k)\rangle$ on a quantum computer and adding them on a classical computer. The energy $E(\vec{\theta}_k)$ is fed into the classical algorithm that updates parameters for the next step of optimization $\vec{\theta}_{k+1}$ according to the chosen optimization algorithm

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

where $a_p^\dagger$ and a_p are fermionic creation and annihilation operators, which excite and deexcite electrons from orbital p. The first term in Eq. 2 corresponds to single-electron excitations, and the second term corresponds to two-electron excitations; h_{pq} and h_{pqrs} are one- and two-electron integrals that are easily computed on a classical computer. To evaluate the energy on a quantum computer, the Hamiltonian for indistinguishable fermions has to be mapped onto the Hamiltonian of distinguishable qubits by using one of three common mappings: Jordan–Wigner (Jordan and Wigner 1928), parity (Bravyi et al. 2017), or Bravyi–Kitaev (Bravyi and Kitaev 2002). Regardless of the mapping choice, the resulting qubit Hamiltonian can be written as

$$\hat{H} = \sum_j \alpha_j P_j = \sum_j \alpha_j \prod_i \sigma_i^j, \tag{3}$$

where α_j are real scalar coefficients that depend on h_{pq} and h_{pqrs}. P_j are Pauli strings represented by a product of Pauli matrices σ_i^j, where i denotes which qubit the Pauli operator acts on and j denotes the term of the Hamiltonian. After the qubit Hamiltonian is prepared on a classical computer and the ansatz to represent the wave function is chosen, the trial wave function $|\psi(\vec{\theta})\rangle$ is prepared on a quantum computer. Then the quantum computer is used to measure the energy:

$$E(\vec{\theta}) = \sum_j^N \alpha_j \langle\psi(\vec{\theta})| \prod_i \sigma_i^j |\psi(\vec{\theta})\rangle, \tag{4}$$

where N is the number of terms in the Hamiltonian and $\vec{\theta}$ is the vector of variational parameters. Depending on the chosen basis set and mapping type, the Hamiltonian can

contain up to M^4 terms, where M is the number of basis functions. Since these terms represented by the Pauli string operators P_j are non-commutative in general, the state preparation step has to be performed repeatedly for each term that is measured separately. In addition, all the individual terms have to be measured enough times to build up sufficient expectation value statistics. This way of computing energy is known as Hamiltonian averaging (McClean et al. 2014). Thus, the VQE approach trades long circuit depths typically found in QPE for shorter state preparation circuits, at the expense of a greater number of measurements, which scales as $O\left(\frac{1}{\epsilon^2}\right)$, where ϵ is the desired precision. QPE converges quadratically faster with scaling of $O\left(\frac{1}{\epsilon}\right)$ (Wiebe and Granade 2016).

Shorter circuits make VQE more amenable to NISQ devices. Despite these advantages, however, the performance of the algorithm depends largely on the quality and flexibility of the trial ansatz. The approaches for ansatz design can be divided into three categories. Chemistry-inspired ansatzes are designed by using the domain knowledge from traditional quantum chemistry in a way that every term in the ansatz describes a certain electron configuration. So-called hardware-efficient ansatzes are constructed from a limited set of gates that are easy to implement on quantum hardware, but a chemical interpretation of each term is not generally possible. The third kind lies between the chemistry-inspired and hardware-efficient and is called Hamiltonian variational ansatz (Wecker et al. 2015). In the present work, we focus on the first two kinds.

The first VQE experiment by Peruzzo and co-workers (Peruzzo et al. 2014) utilized the commonly used unitary coupled cluster with singles and doubles (UCCSD) ansatz, which is a chemistry-inspired ansatz and represents a unitary version of the classical non-unitary CCSD method. The UCCSD trial state is prepared from the initial state $|\phi\rangle$, usually, a Hartree–Fock mean-field wave function, by applying the exponentiated excitation operator $U(\vec{\theta}) = e^{\hat{T}-\hat{T}^\dagger}$:

$$|\psi(\vec{\theta})\rangle = e^{\hat{T}-\hat{T}^\dagger}|\phi\rangle, \tag{5}$$

where the excitation operator $\hat{T} = \sum_i \hat{T}_i$ is truncated at excitation level i. Truncation at $i = 2$ yields the UCCSD ansatz, which includes single and double excitations:

$$\hat{T}_{UCCSD} = \hat{T}_1 + \hat{T}_2 = \sum_{i\in virt, \alpha\in occ} t_i^\alpha \hat{a}_i^\dagger \hat{a}_\alpha + \sum_{i,j\in virt, \alpha,\beta\in occ} t_{ij}^{\alpha\beta} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_\beta \hat{a}_\alpha. \tag{6}$$

In Eq. 6, $t_{i\alpha}$ and $t_{ij\alpha\beta}$ are cluster amplitudes, *occ* denotes orbitals that are occupied in the reference Hartree–Fock state, and *virt* denotes the virtual (unoccupied) orbitals. The standard version of UCCSD has unfavorable scaling of the number of gates required for implementation with an increasing number of electrons and spin orbitals (Kühn et al. 2019). This scaling is a result of many terms with near-zero contributions to the correlation energy. In this work, we consider some of the methods designed to improve that scaling by choosing excitation operators that constitute the ansatz more efficiently.

Hardware-efficient ansatzes are composed of repeated, dense blocks of a limited selection of parametrized gates that are easy to implement with the available hardware. The main idea behind this approach is to build a trial state that is flexible with as few gates as possible. As a result, they are well adapted to the current quantum hardware. This approach has been used to compute the ground state energies of small molecular systems on quantum hardware Kandala et al. (2017, 2019). Since this approach is agnostic to the chemical nature of the system being simulated, it has some significant drawbacks.

The first problem shown by McClean and co-workers is the "barren plateaus" of the variational parameter landscape (McClean et al. 2018), where the derivative of the objective function is close to zero. In addition, hardware-efficient ansatzes require extra work to enforce the physical symmetries, such as electron parity. Not accounting for the symmetries increases the size of the solution space and complicates the variational parameter optimization. This suggests that an arbitrary, unstructured ansatz can lead to poor convergence of the algorithm. Several approaches have been proposed to mitigate this issue (Barkoutsos et al. 2018; Ganzhorn et al. 2019; Grant et al. 2019; Volkoff and Coles 2021). Further work is needed, however, to improve this approach beyond small systems.

In the next sections, we discuss the recently proposed methods to construct more efficient ansatzes using chemistry-inspired and hardware-efficient approaches. These methods focus on improvements in different areas. They can be centered on constructing circuits with fewer CNOT gates that are easier to run on NISQ hardware, choosing the operators that are added to the ansatz based on their contributions to the correlation energy, or performing simulations with higher precision that goes beyond the standard UCCSD. However, each of them can be considered a way to increase the size of the molecules that can be accurately simulated on modern quantum computers.

Chemistry-Inspired ansatze

In this section, we discuss the ansatzes constructed in the fermionic space that use various techniques to improve the UCCSD method to obtain shorter circuits that are easier to run on NISQ hardware or obtain higher than UCCSD accuracy but with comparable circuit depth.

Unitary pair coupled cluster with generalized singles and doubles product wave functions

One of the chemistry-inspired ansatzes designed to reduce circuit length is the Unitary Pair CC with Generalized Singles and Doubles (k-UpCCGSD) method proposed by Lee et al. (2018). The starting point for k-UpCCGSD is the unitary pair coupled cluster double excitations method (UpCCD), which contains a significantly smaller number of operators compared with UCCSD because it includes only two-electron excitations from occupied spatial orbital i to unoccupied spatial orbital a:

$$\hat{T}_2 = \sum_{ia} t_{i_\alpha i_\beta}^{a_\alpha a_\beta} \hat{a}_{a_\alpha}^\dagger \hat{a}_{a_\beta}^\dagger \hat{a}_{i_\beta} \hat{a}_{i_\alpha}, \tag{7}$$

where excitations are performed from occupied orbitals i to unoccupied orbitals a and α and β represent spin. It suffers less from non-variationality and is able to correctly break single bonds. However, it does not recover the dynamic correlation present in UCCD and loses invariance to rotations in occupied-occupied and virtual-virtual subspaces. To address these problems, in addition to pair double excitations, the full single excitations are included in the UpCCSD. For increased accuracy, generalized single and double excitations are included so that there is no distinction between occupied and virtual orbitals p and q when constructing excitations operators:

$$\hat{T}_{UpCCGSD} = \hat{T}_1 + \hat{T}_2 = \sum_{pq} t_p^q \hat{a}_q^\dagger \hat{a}_p + \sum_{pq} t_{p_\alpha p_\beta}^{q_\alpha q_\beta} \hat{a}_{q_\alpha}^\dagger \hat{a}_{q_\beta}^\dagger \hat{a}_{p_\beta} \hat{a}_{p_\alpha}. \tag{8}$$

We note that UCCGSD always produces more accurate results than UCCSD but at a much higher computational cost. In UpCCGSD the cost is reduced by using only pair

double excitations. UpCCGSD is inferior to UCCGSD because of the large number of missing double-excitation operators. For increased flexibility of ansatz, a product of k unitary operators is included in the k-UpCCGSD ansatz, resulting in the final expression:

$$|\psi_{k-UpCCGSD}\rangle = \prod_{\alpha=1}^{k} \left(e^{\hat{T}^{(\alpha)} - \hat{T}^{(\alpha)\dagger}} \right) |\phi\rangle, \qquad (9)$$

where $|\phi\rangle$ is the reference wave function and each operator $\hat{T}^{(k)}$ contains an independent set of variational parameters for single and paired doubles CC amplitudes. Because of the sparsity of the UpCCGSD operator, the cost to prepare a k-UpCCGSD state scales linearly with the system size with a prefactor k. This structure also provides a way to systematically improve accuracy by increasing k.

The k-UpCCGSD method has been tested on H_4, H_2O, and N_2 molecules. Compared with the UCCSD method, k-UpCCGSD offers higher accuracy for the ground state energy with a smaller number of operators in the ansatz for all studied molecules. k-UpCCGSD can achieve chemical accuracy with $k = 2$ for H_4 (STO-3G and 6-31G basis sets) and H_2O (STO-3G basis) molecules. UCCGSD is numerically exact for these systems, but it requires more operators in the ansatz. For N_2 molecule with STO-3G basis, k-UpCCGSD requires $k=4$ to achieve chemical accuracy, whereas UCCSD with more operators fails to come close to chemical accuracy and UCCGSD is within chemical accuracy but requires double the number of operators compared with k-UpCCGSD.

Orbital optimized UCC (OO-UCC) ansatz

OO-UCC is another variant of the UCC approach (Mizukami et al. 2020) where a single-particle orbital rotation operator $\hat{\kappa} = \sum_{pq} \kappa_{pq} \left(a_p^\dagger a_q - a_q^\dagger a_p \right)$ is introduced to the UCC energy functional,

$$E(\hat{\tau}, \hat{\kappa}) = \langle\Phi|e^{-\hat{\kappa}}\hat{H}e^{\hat{\kappa}}|\Phi\rangle = \langle\phi|e^{-\hat{\tau}}e^{-\hat{\kappa}}\hat{H}e^{\hat{\kappa}}e^{\hat{\tau}}|\phi\rangle, \qquad (10)$$

(here $\hat{\tau} = \hat{T} - \hat{T}^\dagger$ is the UCC cluster operator) to variationally determine the coupled cluster amplitudes and also molecular orbital coefficients. Note that optimizing $\hat{\kappa}$ is essentially equivalent to minimizing a wave function with respect to orbital rotation parameters. When $\hat{\tau}$ is fixed, $\hat{\kappa}$ can be obtained through solving a linear equation, $\mathbf{H}\kappa = -\mathbf{g}$, where $\mathbf{H}$ and $\mathbf{g}$ are the electronic Hessian and gradient, respectively. In practice, the OO-UCC ansatz can be easily derived from the UCCSD-alike ansatz (also similar to the DUCC ansatz discussed in the following section), except that only doubles are included,

$$|\Phi\rangle_{\text{UCCSD}} = e^{\hat{\tau}_1 + \hat{\tau}_2}|\phi\rangle \Rightarrow |\Phi\rangle_{\text{UCCSD}'} = e^{\hat{\tau}_2}e^{\hat{\tau}_1}|\phi\rangle \Rightarrow |\Phi\rangle_{\text{OO-UCCD}} = e^{\hat{\tilde{\tau}}_2}|\tilde{\phi}\rangle, \qquad (11)$$

where $\hat{\tau}_1$ and $\hat{\tau}_2$ don't commute, and the singles part $e^{\hat{\tau}_1}$ takes the role of carrying out orbital rotations that would be variationally optimized to obtain $|\tilde{\phi}\rangle = e^{\hat{\tau}_1}|\phi\rangle$ by using a classical computer, while the doubles part, $e^{\hat{\tilde{\tau}}_2}$, is optimized by the VQE. We note that the variational orbitals obtained from the OO-UCCD approach coincide with the commonly known Brueckner orbitals. The reported numerical testing on small molecules has shown that in comparison with conventional UCCSD ansatzes, the qubitization of the OO-UCCD ansatz allows a slight reduction in the number of VQE parameters and the circuit depth. For example, for the NH_3 molecule with the STO-3G basis set, the number of the VQE parameters and the circuit depth required in the OO-UCCD approach were

120 and 2,720, respectively, while those required by the conventional UCCSD were 135 and 2,780, respectively. Regarding accuracy, OO-UCCD maintains a similar level of accuracy to that of UCCSD for most of the small molecules except for LiH, where OO-UCCD results are closer to the full configuration interaction (FCI) results than to UCCSD.

Double unitary coupled cluster (DUCC) ansatz

In modeling second-quantized problems, the major challenge comes from the sizable number of qubits that scales linearly with the size of the basis. To address this issue and enable more realistic simulations on NISQ computers, active space approximations are often employed. Along this line, Metcalf et al. (2020) recently reported VQE applications based on DUCC ansatz that originate from the subsystem embedding sub-algebras coupled cluster studies (Kowalski 2018) to constitute the effective form of the system Hamiltonian

$$\bar{H}^{\mathrm{eff}} = e^{-\sigma_{\mathrm{ext}}} H e^{\sigma_{\mathrm{ext}}} \approx H + [H_{\mathrm{N}}, \sigma_{\mathrm{ext}}] + \frac{1}{2}[[F_{\mathrm{N}}, \sigma_{\mathrm{ext}}], \sigma_{\mathrm{ext}}], \tag{12}$$

where F_{N} and H_{N} are normal product forms of Fock and Hamiltonian operators, respectively, and $\sigma_{\mathrm{ext}} = \hat{T}_{\mathrm{ext}} - \hat{T}_{\mathrm{ext}}^{\dagger}$ is the anti-Hermitian cluster operator constructed from the cluster operators that produce excited configuration outside of the active space when acting on the reference $|\phi\rangle$. After the projection, the energy of the whole system, E, can be directly obtained by diagonalizing the effective Hamiltonian

$$\bar{H}^{\mathrm{eff}} e^{\sigma_{\mathrm{int}}} |\phi\rangle = E e^{\sigma_{\mathrm{int}}} |\phi\rangle \tag{13}$$

with the eigenvector being the excited configurations within the active space.

Note that the DUCC formalism, just like the conventional CC formalism, is formally exact and independent of the choice of the active space, thus providing a systematically improvable hierarchy for implementation. Practically, the accuracy of the DUCC ansatzes depends on the approximate level of the effective Hamiltonian $\bar{H}^{\mathrm{eff}}$ that can be controlled through either the length of the similarity expansion $e^{-\sigma_{\mathrm{ext}}} H e^{\sigma_{\mathrm{ext}}}$ or the many-body terms included in σ_{ext} (and $\hat{T}_{\mathrm{ext}}$), or both. Also, when working with the VQE algorithm, another source of approximation comes from the unitary ansatz for the active space.

For example, the reported DUCC-VQE algorithms include a DUCC effective Hamiltonian truncated at the second-order (see Eq. (12)) and a UCCSD ansatz acting on the active space. In preliminary DUCC-VQE calculations on H_2, Li_2, and BeH_2 molecules (Metcalf et al. 2020), the DUCC Hamiltonians have been shown to significantly outperform the bare Hamiltonian on the same active space in terms of approaching closer to the FCI results.

Quantum subspace expansion (QSE)

The QSE scheme is similar to the generalized eigenvalue problem that is often encountered in quantum chemistry, where the Hamiltonian is diagonalized in a general non-orthogonal basis of many-body states, (McClean et al. 2017; Colless et al. 2018; Takeshita et al. 2020; Motta et al. 2019; Huggins et al. 2020; Parrish and McMahon 2019; Ollitrault et al. 2020; Stair et al. 2020)

$$\mathbf{H}\mathbf{c} = \mathbf{S}\mathbf{c}E. \tag{14}$$

Here the elements of the overlap matrix (**S**) and Hamiltonian matrix (**H**) in the non-orthogonal basis $\{\Phi_i, i = 1, 2, \cdots\}$ are given by

$$S_{ij} = \langle\Phi_i|\Phi_j\rangle,$$
$$\hat{H}_{ij} = \langle\Phi_i|H|\Phi_j\rangle. \tag{15}$$

Compared to the fully classical analog, the QSE quantum simulations utilize quantum devices to construct and store the arbitrarily complex states and measure the overlapping and Hamiltonian matrix elements, while leaving the corresponding eigenvalue problem to the classical machinery.

The many-body basis required for the QSE quantum simulations can be generated in many different ways. For example, analogous to the classical truncated configuration interaction expansion, McClean and co-workers proposed diagonalizing the Hamiltonian in the basis of states $a_i^\dagger a_j|\phi\rangle$ with $|\phi\rangle$ being a reference state obtained from a VQE run (McClean et al. 2017; Colless et al. 2018; Takeshita et al. 2020). The Hamiltonian elements in this basis are the three- and four-body density matrices. Following this line, the many-body basis can be employed to construct a Krylov subspace to improve efficiency and accuracy (Motta et al. 2019; Stair et al. 2020). One of the proposed approaches is the quantum Lanczos (QLanczos) algorithm (Motta et al. 2019), where the basis is the imaginary-time evolution of a single reference state sampled at regular intervals in imaginary time. In QLanczos, for a qubit-encoded Hamiltonian $H = \sum_j H_j$ (here $H_j = \alpha_j P_j$ with α_j complex scalars and P_j Pauli string), the infinitesimal imaginary-time propagator $e^{\Delta_\tau H_j}$ is mirrored by a unitary evolution $e^{i\Delta_\tau A_j}$ acting on properly normalized states. One can show from the Taylor expansion of the time propagator that for an infinitesimal time step, the propagator would span a classical Krylov space. Similarly, via real-time evolution of a set of reference states, Stair et al. (2020) also proposed a multireference selected quantum Krylov (MRSQK) algorithm based on QSE as an alternative to the quantum phase estimation algorithm and have shown that the proposed approach is able to capture the important multideterminantal features (if any) of the wave function and predict the energy of strongly correlated target states.

An alternative approach for creating and utilizing the Krylov subspace was recently proposed by (Kowalski and Peng 2020; Peng and Kowalski 2021). They found that the connected moments expansion, proposed and intensively developed in the 1980s and 1990s (Horn and Weinstein 1984; Cioslowski 1987; Peeters and Devreese 1984; Soldatov 1995), can be employed to re-engineer the energy functional to return a better energy estimate than the straightforward expectation value of the Hamiltonian operator for some trial wave function. Preliminary results on simple molecules and models exhibit high accuracy at finding the ground and excited states and their energies through the rotation of the trial wave function of modest quality, (Kowalski and Peng 2020) and potential capability to circumvent the 'barren plateau' problem (Peng and Kowalski 2021).

We note that despite their higher accuracy and fewer numerical parameters to optimize, QSE methods also expose some practical issues. For example, almost all the above-mentioned QSE methods suffer from the linearly dependent basis that is generated from the procedure, which would cause numerical instabilities when solving the eigenvalue problem. When accounting for hardware noise, the instability would be potentially amplified, leading to worse gate fidelity and larger measurement errors. Furthermore, some QSE methods, in particular, the ones generating bases via real-time propagation,

(Huggins et al. 2020; Parrish and McMahon 2019) may also require extra resources for the evaluation of off-diagonal matrix elements.

Quantum anti-Hermitian contracted schrödinger equation

Contracted Schrödinger equation (CSE) is an approach from the classical electronic structure theory that is based on the reduced density matrix (RDM) theory and solves the contracted eigenvalue equation. The N-electron Schrödinger equation is contracted onto the space of two electrons (Mazziotti 1998; 2002). The anti-Hermitian part of CSE, known as ACSE, has been used to find energies of the ground and excited states in strongly correlated systems on classical computers (Mazziotti 2004; 2020; 2006; 2007; Gidofalvi and Mazziotti 2009; Mukherjee and Kutzelnigg 2001).

The anti-Hermitian part of the CSE reads:

$$\langle \psi | [\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l, \hat{H}] | \psi \rangle = 0. \tag{16}$$

To solve the ACSE, the variational wave function ansatz is constructed iteratively by adding unitary two-body exponential operators:

$$|\psi_{n+1}\rangle = e^{\epsilon \hat{A}_n}|\psi_n\rangle, \tag{17}$$

where $\hat{A}_n$ is anti-hermitian two-body operator

$$\hat{A}_n = \sum_{pqrs} A_n^{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r. \tag{18}$$

The energy at iteration $n+1$ is expressed as

$$E_{n+1} = E_n + \epsilon \langle \psi_n | [\hat{H}, \hat{A}_n] | \psi_n \rangle + O(\epsilon^2). \tag{19}$$

From Eq. 19 it is easy to derive the gradient of the energy with respect to variational parameter A_n^{pqrs}

$$\frac{\partial E}{\partial (A_n^{pqrs})} = -\epsilon \langle \psi_n | [\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r, \hat{H}] | \psi_n \rangle + O(\epsilon^2). \tag{20}$$

The gradient in Eq. 20 is the residual of the ACSE. It vanishes if and only if ACSE is satisfied, meaning that the wave function at iteration n has converged and the minimum of energy is found. On a classical computer, the solution of ACSE for the 2-particle RDM is indeterminant without the storage or reconstruction of 3-RDM, for which the cost scales exponentially. The quantum algorithm for solving the ACSE, (Smart and Mazziotti 2021) however, can be used to solve for 2-RDM without 3-RDM reconstruction. In QACSE the auxiliary states $|\Lambda_n^\pm\rangle = e^{\pm i\delta \hat{H}}|\Psi_n\rangle$ are prepared on a quantum computer. Then, the 2-RDMs of the auxiliary states are measured on a quantum computer to construct the A matrix:

$$A_n^{pqrs} = \frac{1}{2i\delta} \left(\langle \Lambda_n^+ | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r | \Lambda_n^+ \rangle - \langle \Lambda_n^- | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r | \Lambda_n^- \rangle \right) + O\left(\delta^2\right) \tag{21}$$

This measurement of A matrix on a quantum computer allows to avoid 3-RDM reconstruction. Thus, the QACSE algorithm has a potentially exponential speed-up compared to the classical one with full RDM reconstruction. The QACSE algorithm has been applied to H_2, H_3, and C_6H_4 molecules on quantum hardware (Smart and Mazziotti 2021; Smart et al. 2021).

It is important to note that the ACSE wave function also served as an inspiration for the development of the ADAPT-VQE ansatz and its different forms discussed in this article.

Equation 20 is central in constructing ADAPT-VQE ansatz as it allows to find the operator from the pool, which will result in the largest contribution to correlation energy.

Adaptive derivative-Assembled pseudo-Trotter ansatz variational quantum eigensolver (ADAPT-VQE)

Grimsley et al. introduced the Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Quantum Eigensolver, or ADAPT-VQE in short (Grimsley et al. 2019). The key idea of the method is to construct an ansatz that would recover the most correlation energy with the least number of fermionic operators and variational parameters. It is inspired by the iterative algorithm for the ACSE solution (see previous section for details). The first step is to define a pool of operators that contains all fermionic excitation operators that can be added to the ansatz. In principle, this pool can be constructed from any set of operators, but the most straightforward choice is to use operators generated by the unitary coupled cluster ansatz. The use of the generalized version of UCCSD allows achieving higher than UCCSD accuracy with a smaller number of operators. In the generalized version (UCCGSD) (Lee et al. 2018) the operators are formed through single and double excitations over all occupied and virtual orbitals, which are not distinguished in this approach (Eq. 7). Potentially, higher-order excitations can be included. After performing standard computations of one- and two-electron integrals along with reference Hartree–Fock wave function $|\phi\rangle$ on a classical computer the algorithm starts to gradually grow the variational ansatz with operators that would contribute the most correlation energy. Such operators are found through the computation of the following energy derivatives with respect to the variational parameters:

$$\frac{\partial E}{\partial \theta_i} = \langle \psi | [\hat{H}, \hat{\tau}_i] | \psi \rangle, \tag{22}$$

where $\hat{\tau}_i$ is a sum of excitation and de-excitation operators, for example, $\hat{\tau}_{pq}^{rs} = \hat{t}_{pq}^{rs} - \hat{t}_{rs}^{pq}$. Equation 22 for computation of the gradient is equivalent to Eq. 20 derive in the previous section of this article that describes ACSE. At every ADAPT-VQE step, one operator is added and full VQE optimization of all parameters is performed. The process is repeated until the convergence criterion is met when the norm of the gradient vector becomes smaller than the predefined threshold:

$$\| \overrightarrow{g} \| = \sqrt{\sum_i \left(\frac{\partial E}{\partial \theta_i} \right)^2}. \tag{23}$$

When convergence is achieved, the algorithm produces the following ansatz:

$$|\psi^{ADAPT}\rangle = \left(e^{\hat{\tau}_N} \right) \left(e^{\hat{\tau}_{N-1}} \right) ... \left(e^{\hat{\tau}_2} \right) \left(e^{\hat{\tau}_1} \right) |\phi\rangle. \tag{24}$$

The ADAPT-VQE method has advantages compared with UCCSD for performing simulations on NISQ hardware. Compared with the UCCSD ansatz that produces many redundant terms contributing little to the correlation energy, ADAPT-VQE contains a much smaller number of operators. As a result, the quantum circuits are much shorter, which can enable simulations on quantum hardware for larger molecules than what is possible with UCCSD ansatz. In addition, it addresses perhaps the weakest quality of the VQE method, namely, that classical parameter optimization in VQE can become intractable because of barren plateaus (McClean et al. 2018) or simply because the number of variational parameters is too large. ADAPT-VQE adds operators only with the

highest contribution to the correlation energy and avoids the problem of having to optimize a large number of parameters with near-zero contributions. Another advantage is that parameters are optimized one at a time, so every iteration of ADAPT-VQE is started with all parameters preoptimized, except for the one newly added, which is likely to accelerate the rate of convergence. Another benefit of ADAPT-VQE is that the accuracy of the method can be controlled by adjusting the convergence criterion. When higher accuracy is needed, one can simply add operators with smaller contributions in the order of decreasing gradients (Eq. 22) systematically converging to the solution provided by the full operator pool.

When considering the cost of simulations using a certain algorithm on NISQ hardware, among the most critical metrics are circuit depth, which is limited because of the short coherence times and large error rates, and the total shot count, which defines the time to solution. ADAPT-VQE is successful in identifying short and accurate ansatzes. However, the total shot count can drastically exceed the shot count for UCCSD ansatz. The reason is mainly that the number of energy gradients to be computed at each ADAPT iteration is equal to the number of operators in the pool, which grows quickly with system size. Another factor that increases the time to solution is that the number of VQE optimizations is equal to the final number of operators in the ansatz. Therefore, the time to solution can become unfeasible for large molecules, requiring a large number of operators to converge. We point out that for NISQ devices with limited coherence times, shorter circuits are more critical than total shot count and can be a defining factor if the simulation can be performed at all.

The examples of ADAPT-VQE implementation of ADAPT-VQE include potential energies of LiH, BeH$_2$, and highly correlated H$_6$ molecules using the STO-3G basis set. For the LiH molecule, the UCCSD ansatz produces results within chemical accuracy across the whole PES. ADAPT-VQE produces the same or better results with around 20 parameters instead of 60 required by UCCSD. With extra parameters, the errors can be reduced further. With an additional 5 parameters, the error is consistently under 10^{-3} kcal/mol. For the BeH$_2$ molecule, UCCSD is within chemical accuracy around equilibrium; but for longer internuclear distances, the error gets larger than 1 kcal/mol. ADAPT-VQE achieves accuracies around 0.1 kcal/mol across the whole PES using around 30 parameters instead of 120 in UCCSD. For the H$_6$ molecule, ADAPT-VQE also outperforms UCCSD. The number of required parameters grows significantly for longer distances to describe this highly correlated system.

As with any adaptive method, estimating the resources required to perform a simulation is difficult. It strongly depends on the chemical system, initial operator pool, and specifics of the implementation. In the original implementation, (Grimsley et al. 2019) operators are added one by one, and all parameters are reoptimized at every ADAPT-VQE iteration. Adding operators in batches can help reduce the cost, as can freezing some of the parameters for a certain number of iterations. Another way to reduce the shot count is to pre-screen the operators in the pool by using various techniques, for example, MP2 amplitude screening. Additionally, to reduce the time to solution, one can take advantage of the fact that the gradient calculations are independent and can be parallelized by using multiple quantum computers. Because of the limited availability of quantum hardware, however, this can still be a problem.

qubit-ADAPT-VQE

Although the ADAPT-VQE algorithm allows one to significantly reduce the circuit depth compared with the UCCSD ansatz while achieving higher accuracy, the resulting quantum circuits are still beyond the reach of NISQ devices. Particularly problematic are the multi-qubit (e.g., CNOT) gates since they tend to have much higher error rates compared with 1-qubit gates. To reduce the number of CNOT gates in the circuits, Tang et al. proposed the qubit-ADAPT-VQE algorithm (Tang et al. 2021). In this algorithm, the general concept is the same as in the original ADAPT-VQE, but instead of fermionic operators, the individual Pauli strings form the operator pool $\hat{\tau} = \hat{P} = i \prod_i p_i, p_i \in \{X, Y, Z\}$. These strings are obtained from the strings that are generated in the fermionic pool. This ansatz yields shallower circuits with fewer CNOT gates compared with the fermionic ADAPT-VQE, but at a cost of a larger number of variational parameters. Essentially, it is a way of offloading the computational effort from a quantum processor to a classical computer motivated by the limitations of NISQ hardware. The qubit-ADAPT-VQE approach uses the same procedure to add operators from the operator pool to the ansatz by computing gradients of energy with respect to the variational parameters associated with the operators. Because of the larger number of parameters compared with the fermionic ADAPT, the "qubit pool" requires more gradient calculations on a quantum computer. The number of operators in the pool scales exponentially as $2^n - 1$. Therefore, it is crucial to eliminate redundant operators from the pool to reduce its size. First, the strings with an even number of Y operators are eliminated to ensure that the fermionic operator $\hat{P} = i \prod_i p_i, p_i \in \{X, Y, Z\}$ is real. In addition, chains of Z gates are removed from the pool since they do not affect the performance of the method according to numerical simulations. However, the size of this "qubit pool" remains large. Also, it is important to note that the chains of Z gates are responsible for enforcing the anti-symmetry of the electronic wave function. The removal of these gates did not affect the numerical results for small test molecules but the effect of this symmetry removal on the accuracy of description of larger molecules is still unknown. It has been shown that the size of the qubit-ADAPT-VQE pool can be reduced dramatically without sacrificing accuracy, and it has been analytically proven that complete pools of size $2n - 2$ exist for any n (Grimsley et al. 2019). The recipe for constructing such a minimal complete pool is a direction for future research. If the extra measurement overhead in qubit-ADAPT-VQE can be reduced by an efficient algorithm for finding a complete pool, this method will be a good fit for applications on NISQ hardware, as long as the classical optimization part can be efficiently solved. The qubit-ADAPT-VQE method was tested on H_4, H_6, and *LiH* molecules using STO-3G basis set. It can achieve the same accuracy as does the fermionic ADAPT with fewer CNOT gates but with a larger number of ADAPT iterations because of the larger number of variational parameters.

Hardware-Efficient ansatzes

In this section, we discuss hardware-efficient ansatzes that produce quantum circuits that are easier to run on modern quantum hardware.

Symmetry-Preserving state preparation

The idea of imposing symmetries associated with particle number, spin, and time-reversal symmetries goes along with the approach that bases the VQE ansatzes on the capabilities

of the hardware and performs state preparation through combining parameterized gates available on the processor (Bian et al. 2019; Gard et al. 2020).

The ansatzes in this category have the advantage of being compatible with the capabilities of the hardware, but the ad hoc ansatzes of this type can also cause the so-called "barren plateaus" problem (McClean et al. 2018), where gradients vanish exponentially in sufficiently expressive parameterized quantum circuits, which in turn requires an exponentially large precision to navigate through the "barren plateaus" landscape. Similar problems (i.e., gradient-vanishing) existed in the early studies of deep neural networks (Bradley et al. 2009; Shalev-Shwartz et al. 2017; Kremer and Kolen 2001) with mitigation techniques proposed later on (LeCun et al. 2015; Ioffe and Szegedy 2015; Hinton et al. 2006; He et al. 2016). Nevertheless, for hardware-based ansatzes to be successful in solving the problems of interest, most of the approaches have focused on assuring that the hardware-based ansatzes span the part of the Hilbert space that includes the true solution.

So far, two ways have been employed for facilitating the access of the parameterized quantum circuits to the "right" part of the Hilbert space. On the one hand, penalty terms can be implemented in the VQE energy function for symmetry violations (McClean et al. 2016; Ryabinkin et al. 2019). On the other hand, the state preparation circuits need to be carefully designed to preserve appropriate symmetries regardless of variational parameter rotation. An early attempt in this direction (Wang et al. 2009), in comparison with preparing a general state in the full Hilbert space of n qubits that requires $\mathcal{O}(2^n)$ controlled-NOT gates, shows the actual number of the controlled-NOT gates required for exploring a smaller Hilbert space, where the true solution scales only polynomially with the number of qubits, which can still be greatly reduced by several orders of magnitude by properly designing the quantum algorithm that accounts for additional symmetries. Similar advantages have also been reported by Barkoutsos et al. (2018), who found the reformulation of the molecular Hamiltonian in second quantization using the particle-hole picture in conjunction with a parameterized particle-conserving exchange-type gate (Ganzhorn et al. 2019; Roth et al. 2017; Egger et al. 2019; Sagastizabal et al. 2019) is able to improve the computational efficiency and accuracy for quantum chemistry simulations. Nevertheless, important open questions remain, including how other symmetries can be built into the circuits and whether more efficient circuits exist that contain the minimal number of parameters necessary to span the symmetry subspace. To encode other symmetries into the circuits, while still balancing the efficiency and accuracy, Gard et al. (2020) introduced generalized state preparation circuits that accommodate well-defined symmetries (including particle number, total spin, spin projection, and time reversal) and require a minimal number of parameters to directly target the appropriate symmetry subspace. Although tested only for H_2 and LiH molecules, it is shown that the circuits are able to locate the true ground state within the subspace of states spanned by the circuit and reduce the complexity of the classical optimization step of the VQE algorithm, thus outperforming the standard state preparation ansatze.

Despite these efforts, as well as more recent ones (Fontana et al. 2020; Anand et al. 2021; Arrasmith et al. 2021; Uvarov and Biamonte 2021; Zhang et al. 2020; Pesah et al. 2021), it is still too early to claim success since additional difficulties can come from hardware noise that can potentially modify the cost landscape associated with the parameter space. For example, a recent study (Fontana et al. 2020) found that the hardware noise (specifically non-unital noise) can break the underlying symmetries in parameterized

quantum circuits and lift the degeneracy of minima that falsifies local minima as global minima.

Qubit coupled cluster method

The qubit coupled cluster (QCC) method introduced by Ryabinkin et al. (2018) resembles the structure of coupled cluster; but instead of using fermionic excitations, the ansatz is built directly in the qubit space. The QCC wave function is of the form

$$|\Psi(\tau, \Omega)\rangle = \hat{U}(\tau)|\Omega\rangle. \tag{25}$$

The mean-field wave function $|\Omega\rangle$ is a product of single-qubit states

$$|\Omega\rangle = \prod_{j=1}^{n}|\Omega_j\rangle, \tag{26}$$

where n is the number of spin-orbitals. Each $|\Omega_j\rangle$ is parameterized with Bloch angles ϕ and θ:

$$|\Omega_j\rangle = \cos\left(\frac{\theta_j}{2}\right)|\uparrow_j\rangle + e^{i\phi_j}\sin\frac{\theta_j}{2}|\downarrow_j\rangle, \tag{27}$$

where $|\uparrow_j\rangle$ and $|\downarrow_j\rangle$ are eigenstates of $\hat{z}_j$. Entanglement is introduced by the multi-qubit rotations with real amplitudes $\tau = \{\tau_k\}$:

$$\hat{U}(\tau) = \prod_{k=1}^{N_{ent}} e^{-i\tau_k \hat{P}_k/2}, \tag{28}$$

where $\hat{P}_k$ are Pauli strings of length from 2 to a number of qubits N_q and the number of entanglers N_{ent} is less than or equal to the number of $\hat{P}_k$. The ground state energy of the system is obtained through minimization of the Hamiltonian expectation value

$$E_{QCC} = \min_{\Omega,\tau}\langle\Psi(\tau,\Omega)|\hat{H}|\Psi(\tau,\Omega)\rangle = \min_{\Omega,\tau}\langle\Omega|U(\tau)^{\dagger}\hat{H}U(\tau)|\Omega\rangle. \tag{29}$$

Similar to other hardware-efficient methods working directly in the qubit space, QCC ansatz can result in unphysical results, for example breaking of symmetries, such as non-conservation of the total number of particles or obtaining a state with a wrong electronic spin. This problem has been shown (Ryabinkin et al. 2019) to be more a rule than an exception. Therefore, additional constraints need to be used to ensure that results do not violate any physical laws. Yen et al. (2019) proposed symmetry projectors that can enforce the symmetries and implemented these projectors with the QCC method.

The total number of operators in the transformed Hamiltonian $U(\tau)^{\dagger}\hat{H}U(\tau)$ scales exponentially and contains $3^{N_{ent}}$ parameters. Therefore, the operators with the largest contribution to the correlation energy need to be chosen to avoid exponential scaling. In the QCC framework such screening of operators is performed by computing the energy derivative with respect to τ_k at $\tau_k = 0$:

$$\left|\frac{dE[\hat{P}_k]}{d\tau_k}\right|_{\tau_k=0} = \left|\frac{d}{d\tau_k}\min_{\Omega}\langle\Omega|e^{i\tau_k\hat{P}_k/2}\hat{H}e^{-i\tau_k\hat{P}_k/2}|\Omega\rangle\right|_{\tau_k=0}$$
$$= \left|\langle\Omega_{\min}|-\frac{i}{2}[\hat{H},\hat{P}_k]|\Omega_{\min}\rangle\right| > 0, \tag{30}$$

where $|\Omega_{\min}\rangle$ is the qubit mean field (QMF) wave function at the point of minimum QMF energy. This expression for the gradient is analogous to one that is used in the ADAPT-VQE approach.

The entanglers responsible for the largest energy reduction are ranked in order of decreasing first energy derivatives. If the first derivative vanishes, the second tier of entanglers can be formed by performing ranking using the second derivatives. Thus, the QCC ansatz can be expanded systematically based on this ranking, with more terms recovering a larger amount of correlation energy. By including less important terms, one can systematically approach the exact solution. However, the computational complexity of this screening procedure is exponential with respect to the number of terms in the Hamiltonian. The screening procedure can be improved by taking advantage of the Hamiltonian symmetry and splitting operators into groups, which results in the polynomial complexity for screening the operators (Ryabinkin et al. 2020).

Another shortcoming of QCC, an exponential number of operators, which restricts the application of QCC to small systems, was addressed in the iterative version of the QCC method (iQCC) (Ryabinkin et al. 2020). In this method, the "dressed" canonically transformed Hamiltonian is considered:

$$\hat{H}_d = \hat{U}^\dagger(\tau)\hat{H}\hat{U}(\tau). \tag{31}$$

The dressed Hamiltonian is evaluated recursively according to

$$
\begin{aligned}
\hat{H}_d^{(k)}(\tau_k, ..., \tau_1) &= e^{i\tau_k\hat{P}_k/2}\hat{H}_d^{(k-1)}\left(\tau_{k-1}, ..., \tau_1\right)e^{-i\tau_k\hat{P}_k/2} \\
&= \hat{H}_d^{(k-1)} - \frac{i}{2}\left[\hat{H}_d^{(k-1)}, \hat{P}_k\right]\sin\tau_k \\
&\quad + \frac{1}{2}\hat{P}_k\left[\hat{H}_d^{(k-1)}, \hat{P}_k\right](1 - \cos\tau_k),
\end{aligned}
\tag{32}
$$

where $k = 1, ..., N_{ent}$. Instead of optimizing all N_{ent} amplitudes at once, the amplitudes can be optimized sequentially by introducing one or more operators at each iteration. During each iQCC iteration $N_{ent} \geq 1$ amplitudes are optimized using the Hamiltonian from the previous step. Then the optimized amplitudes are used to construct the dressed Hamiltonian (Eq. 29) for the next iteration according to Eq. (32).

In general, this procedure introduces $3^{N_{ent}}$ distinct operators and shows the exponential complexity of the QCC ansatz. It has been demonstrated that if the amplitudes τ in Eq. (32) are fixed, the complexity of transforming the dressed Hamiltonian in Eq. (31) is $\sim M(3/2)^{N_{ent}}$, where M is the number of terms in the Hamiltonian. The $(3/2)^{N_{ent}}$ factor comes from the canonical transformation of Hamiltonian Eq. (31) at each step of iQCC and for larger systems can make simulation prohibitively expensive. This problem can be partially solved by "compressing" the dressed Hamiltonian. This procedure essentially removes the terms from the Hamiltonian with little contribution to the ground state energy in a controlled manner so that the energy changes by no more than desired accuracy ϵ. Numerical examples have shown that the compressing procedure reduces the $(3/2)^{N_{ent}}$ factor. The benefits from this procedure increase for larger systems where the fraction of terms with a small contribution to energy in the Hamiltonian becomes large. In a more recent effort, aiming at reducing the number of operators in the dressed Hamiltonian, Eq. (32), Izmaylov group introduced involutory linear combinations (ILC) of anti-commuting Pauli products (i.e. entanglers) in the QCC framework, (Lang et al. 2021) and the resulting QCC-ILC unitary dressing gives an exact quadratic truncation of the Baker-Campbell-Hausdorff expansion which, in comparison to the random QCC transformation, only results in quadratic growth of the number of Pauli strings in the dressed Hamiltonian while still yields accurate energy estimates in the strongly correlated regime.

QCC has been benchmarked on LiH, H_2O, and N_2 molecules using the STO-3G basis. Chemical accuracy was achieved for these small molecules. However, performing QCC calculations on large molecules would require further improvements to the algorithm, in particular, the $(3/2)^{N_{ent}}$ factor appearing in the dressed Hamiltonian transformation.

Conclusions

The main shortcoming of the standard UCCSD ansatz for VQE simulations is rooted in the overall number of excitations. Although each excitation operator generally adds only one variational parameter, the total number of excitations grows rapidly with system size. A large number of amplitudes increases the length of circuits, thus increasing the load on the quantum computer. At the same time, a larger number of parameters have to be optimized within the classical VQE loop. Moreover, circuits to implement fermionic excitations contain a large number of multi-qubit gates, which are a problem for simulations on NISQ hardware because of the large error rates of multi-qubit gates. The methods described in this paper are designed to construct short-depth ansatzes that are easier to run on NISQ hardware. In general, chemistry-inspired ansatzes have a goal of a more careful choice of excitations to include in the ansatz, while hardware-efficient approaches focus on the reduction of the number of multi-qubit gates and using various techniques to choose compact blocks of entangling gates with single-qubit rotational gates more efficiently.

In adaptive methods that include ADAPT, qubit-ADAPT-VQE, and QCC methods, the addition of the operators to the ansatz is based on the calculation of the gradient of energy with respect to the parameter associated with this operator, where the largest gradient corresponds to the largest contribution to the correlation energy. It is straightforward to ensure that an ansatz does not contain redundant terms that complicate classical optimization and make quantum circuits longer. Doing so, however, comes at the cost of extra measurements associated with gradient calculations used to choose the operators contributing the most amount of correlation energy. QCC and qubit-ADAPT-VQE work directly in the qubit space and are considered more "hardware-efficient" because they are generally more compact and easier to run on NISQ hardware. However, they require additional work to ensure the preservation of all physical symmetries. UCC-based methods such as ADAPT-VQE, *k*-UpCCGSD, DUCC, OO-UCC, and ansatz are constructed from fermionic operators and therefore avoid the problems with symmetry preservation. However, the number of multi-qubit gates remains high in this method, which requires the error rates on quantum hardware to significantly improve before such methods can be used. On the other hand, QSE-based approaches exhibit higher accuracy from ansatzes prepared with relatively low circuit depth and require fewer numerical parameters to be optimized, but they need extra measurements for Hamiltonian powers, overlap matrix, and/or energy gradient, and may suffer from the linearly dependent many-body basis. Thus many issues still are left to be resolved, and significant progress is expected in these directions in the post-VQE era.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Oak Ridge Associated Universities, Oak Ridge 37830, Tennessee, USA. [2]Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland 99354, Washington, USA. [3]Computational Science Division, Argonne National Laboratory, Lemont 60439, Illinois, USA.

References

D. S. Abrams, S. Lloyd, Simulation of many-body fermi systems on a universal quantum computer. Phys. Rev. Lett. **79**, 2586–2589 (1997). https://doi.org/10.1103/PhysRevLett.79.2586

D. S. Abrams, S. Lloyd, Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. Phys. Rev. Lett. **83**, 5162–5165 (1999). https://doi.org/10.1103/PhysRevLett.83.5162

D. Aharonov, M. Ben-Or, Fault Tolerant Quantum Computation with Constant Error. arXiv (1996). quant-ph/9611025. 03 Oct 2021

A. Anand, M. Degroote, A. Aspuru-Guzik, Natural evolutionary strategies for variational quantum computation. Mach. Learn. Sci. Technol. **2**(4), 045012 (2021). https://doi.org/10.1088/2632-2153/abf3ac

A. Arrasmith, M. Cerezo, P. Czarnik, L. Cincio, P. J. Coles, Effect of barren plateaus on gradient-free optimization. Quantum. **5**, 558 (2021). https://doi.org/10.22331/q-2021-10-05-558

R. Babbush, D. W. Berry, J. R. McClean, H. Neven, Quantum simulation of chemistry with sublinear scaling in basis size. NPJ Quantum Inf. **5**(1), 92 (2019). https://doi.org/10.1038/s41534-019-0199-y. 1807.09802

R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, G. K.-L. Chan, Low-depth quantum simulation of materials. Phys. Rev. X. **8**, 011044 (2018). https://doi.org/10.1103/PhysRevX.8.011044

P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, S. Filipp, I. Tavernelli, Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions. Phys. Rev. A. **98**(2), 022322 (2018). https://doi.org/10.1103/physreva.98.022322. 1805.04340

K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, A. Aspuru-Guzik, Noisy intermediate-scale quantum (NISQ) algorithms (2021). 2101.08448. Accessed 03 Oct 2021

T. Bian, D. Murphy, R. Xia, A. Daskin, S. Kais, Quantum computing methods for electronic states of the water molecule. Mol. Phys. **117**(15–16), 2069–2082 (2019). https://doi.org/10.1080/00268976.2019.1580392. https://doi.org/10.1080/00268976.2019.1580392

D. M. Bradley, J. A. Bagnell, Y. Bengio, M. Hebert, F. De, L. Torre, Learning in modular systems. Technical report (2009)

S. B. Bravyi, A. Y. Kitaev, Fermionic quantum computation. Ann. Phys. **298**(1), 210–226 (2002). https://doi.org/10.1006/aphy.2002.6254

S. Bravyi, J. M. Gambetta, A. Mezzacapo, K. Temme, Tapering off qubits to simulate fermionic Hamiltonians. arXiv (2017). 1701.08213. Accessed 03 Oct 2021

Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferova, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing. Chem. Rev. **119**(19), 10856–10915 (2019). https://doi.org/10.1021/acs.chemrev.8b00803

M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, P. J. Coles, Variational quantum algorithms. Nat. Rev. Phys. **3**(9), 625–644 (2021). https://doi.org/10.1038/s42254-021-00348-9

J. Cioslowski, Connected moments expansion: a new tool for quantum many-body theory. Phys. Rev. Lett. **58**(2), 83 (1987)

G.gleA.I.Q.uantum.a.nd. Collaborators∗†, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Y. Niu, T. E. O'Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh, A. Zalcman, Hartree-Fock on a superconducting qubit quantum computer. Science. **369**(6507), 1084–1089 (2020). https://doi.org/10.1126/science.abb9811. 2004.04174

J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, I. Siddiqi, Computation of molecular spectra on a quantum processor with an error-resilient algorithm. Phys. Rev. X. **8**, 011021 (2018). https://doi.org/10.1103/PhysRevX.8.011021

B. Cooper, P. J. Knowles, Benchmark studies of variational, unitary and extended coupled cluster methods. J. Chem. Phys. **133**(23), 234102 (2010). https://doi.org/10.1063/1.3520564

D. J. Egger, M. Ganzhorn, G. Salis, A. Fuhrer, P. Müller, P. K. Barkoutsos, N. Moll, I. Tavernelli, S. Filipp, Entanglement generation in superconducting qubits using holonomic operations. Phys. Rev. Applied. **11**, 014017 (2019). https://doi.org/10.1103/PhysRevApplied.11.014017

V. E. Elfving, B. W. Broer, M. Webber, J. Gavartin, M. D. Halls, K. P. Lorton, A. Bochevarov, How will quantum computers provide an industrially relevant computational advantage in quantum chemistry? arXiv preprint arXiv:2009.12472 (2020)

F. A. Evangelista, G. K.-L. Chan, G. E. Scuseria, Exact parameterization of fermionic wave functions via unitary coupled cluster theory. J. Chem. Phys. **151**(24), 244112 (2019). https://doi.org/10.1063/1.5133059

R. P. Feynman, Simulating physics with computers. Int. J. Theor. Phys. **21**(6/7), 467–488 (1982). https://doi.org/10.1007/BF02650179

E. Fontana, M. Cerezo, A. Arrasmith, I. Rungger, P. J. Coles, Optimizing parametrized quantum circuits via noise-induced breaking of symmetries. arXiv (2020). 2011.08763

M. Ganzhorn, D. J. Egger, P. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, M. Roth, A. Fuhrer, P. Mueller, S. Woerner, I. Tavernelli, S. Filipp, Gate-efficient simulation of molecular eigenstates on a quantum computer. Phys. Rev. Applied. **11**, 044092 (2019). https://doi.org/10.1103/PhysRevApplied.11.044092

Q. Gao, G. O. Jones, M. Motta, M. Sugawara, H. C. Watanabe, T. Kobayashi, E. Watanabe, Y.-y. Ohnishi, H. Nakamura, N. Yamamoto, Applications of Quantum Computing for Investigations of Electronic Transitions in Phenylsulfonyl-carbazole TADF Emitters (2020). http://arxiv.org/abs/2007.15795

Q. Gao, H. Nakamura, T. P. Gujarati, G. O. Jones, J. E. Rice, S. P. Wood, M. Pistoia, J. M. Garcia, N. Yamamoto, Computational investigations of the lithium superoxide dimer rearrangement on noisy quantum devices. J. Phys. Chem. A. **125**(9), 1827–1836 (2021). https://doi.org/10.1021/acs.jpca.0c09530. PMID: 33635672. https://doi.org/10.1021/acs.jpca.0c09530

B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, E. Barnes, Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm. NPJ Quantum Inf. **6**(1), 10 (2020). https://doi.org/10.1038/s41534-019-0240-1

G. Gidofalvi, D. A. Mazziotti, Direct calculation of excited-state electronic energies and two-electron reduced density matrices from the anti-hermitian contracted schrödinger equation. Phys. Rev. A. **80**, 022507 (2009). https://doi.org/10.1103/PhysRevA.80.022507

E. Grant, L. Wossnig, M. Ostaszewski, M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits. Quantum. **3**, 214 (2019). https://doi.org/10.22331/q-2019-12-09-214

H. R. Grimsley, S. E. Economou, E. Barnes, N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer. Nat. Commun. **10**(1), 3007 (2019). https://doi.org/10.1038/s41467-019-10988-2

G. Harsha, T. Shiozaki, G. E. Scuseria, On the difference between variational and unitary coupled cluster theories. J. Chem. Phys. **148**(4), 044107 (2018). https://doi.org/10.1063/1.5011033

K. He, X. Zhang, S. Ren, J. Sun, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Deep residual learning for image recognition, (2016), pp. 770–778. https://doi.org/10.1109/CVPR.2016.90

C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, C. F. Roos, Quantum chemistry calculations on a trapped-ion quantum simulator. Phys. Rev. X. **8**, 031022 (2018). https://doi.org/10.1103/PhysRevX.8.031022

G. E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets. Neural Comput. **18**(7), 1527–1554 (2006). https://doi.org/10.1162/neco.2006.18.7.1527. PMID: 16764513. https://doi.org/10.1162/neco.2006.18.7.1527

M. R. Hoffmann, J. Simons, A unitary multiconfigurational coupled-cluster method: Theory and applications. J. Chem. Phys. **88**(2), 993–1002 (1988). https://doi.org/10.1063/1.454125

D. Horn, M. Weinstein, The t expansion: a nonperturbative analytic tool for hamiltonian systems. Phys. Rev. D. **30**(6), 1256 (1984)

W. J. Huggins, J. Lee, U. Baek, B. O'Gorman, K. B. Whaley, A non-orthogonal variational quantum eigensolver. New J. Phys. **22**(7), 073009 (2020). https://doi.org/10.1088/1367-2630/ab867b

S. Ioffe, C. Szegedy, in *Proceedings of the 32nd International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 37*. ed. by F. Bach, D. Blei, Batch normalization: Accelerating deep network training by reducing internal covariate shift (PMLR, Lille, 2015), pp. 448–456. http://proceedings.mlr.press/v37/ioffe15.html

P. Jordan, E. Wigner, Über das Paulische Äquivalenzverbot. Z. Phys. **47**(9-10), 631–651 (1928). https://doi.org/10.1007/BF01331938

A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature. **549**(7671), 242–246 (2017). https://doi.org/10.1038/nature23879

A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor. Nature. **567**(7749), 491–495 (2019). https://doi.org/10.1038/s41586-019-1040-7

A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem. arXiv (1995). quant-ph/9511026. Accessed 03 Oct 2021

J. S. Kottmann, P. Schleich, T. Tamayo-Mendoza, 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**(1), 663–673 (2021). https://doi.org/10.1021/acs.jpclett.0c03410. PMID: 33393305. https://doi.org/10.1021/acs.jpclett.0c03410

K. Kowalski, Properties of coupled-cluster equations originating in excitation sub-algebras. J. Chem. Phys. **148**(9), 094104 (2018). https://doi.org/10.1063/1.5010693. https://doi.org/10.1063/1.5010693

K. Kowalski, B. Peng, Quantum simulations employing connected moments expansions. J. Chem. Phys. **153**(20), 201102 (2020). https://doi.org/10.1063/5.0030688. https://doi.org/10.1063/5.0030688

S. C. Kremer, J. F. Kolen, *Field Guide to Dynamical Recurrent Networks*, 1st edn. (Wiley-IEEE Press, 2001). https://www.wiley.com/en-us/A+Field+Guide+to+Dynamical+Recurrent+Networks-p-9780780353695

M. Kühn, S. Zanker, P. Deglmann, M. Marthaler, H. Weiß, Accuracy and resource estimations for quantum chemistry on a near-term quantum computer. J. Chem. Theory Comput. **15**(9), 4764–4780 (2019). https://doi.org/10.1021/acs.jctc.9b00236. PMID: 31403781. https://doi.org/10.1021/acs.jctc.9b00236

W. Kutzelnigg, Error analysis and improvements of coupled-cluster theory. Theo. Chim. Acta. **80**(4), 349–386 (1991). https://doi.org/10.1007/BF01117418

R. A. Lang, I. G. Ryabinkin, A. F. Izmaylov, Unitary transformation of the electronic hamiltonian with an exact quadratic truncation of the baker-campbell-hausdorff expansion. J. Chem. Theory Comput. **17**(1), 66–78 (2021). https://doi.org/10.1021/acs.jctc.0c00170

Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature. **521**(7553), 436–444 (2015). https://doi.org/10.1038/nature14539

J. Lee, W. J. Huggins, M. Head-Gordon, K. B. Whaley, Generalized unitary coupled cluster wave functions for quantum computation. J. Chem. Theory Comput. **15**(1), 311–324 (2018). https://doi.org/10.1021/acs.jctc.8b01004

H. Liu, G. H. Low, D. S. Steiger, T. Häner, M. Reiher, M. Troyer, Prospects of Quantum Computing for Molecular Sciences (2021). 2102.10081. Accessed 03 Oct 2021

D. A. Mazziotti, Contracted schrödinger equation: Determining quantum energies and two-particle density matrices without wave functions. Phys. Rev. A. **57**, 4219–4234 (1998). https://doi.org/10.1103/PhysRevA.57.4219

D. A. Mazziotti, Variational method for solving the contracted Schrödinger equation through a projection of the N -particle power method onto the two-particle space. J. Chem. Phys. **116**(4), 1239–1249 (2002). https://doi.org/10.1063/1.1430257

D. A. Mazziotti, Exactness of wave functions from two-body exponential transformations in many-body quantum theory. Phys. Rev. A. **69**, 012507 (2004). https://doi.org/10.1103/PhysRevA.69.012507

D. A. Mazziotti, Anti-hermitian contracted schrödinger equation: Direct determination of the two-electron reduced density matrices of many-electron molecules. Phys. Rev. Lett. **97**, 143002 (2006). https://doi.org/10.1103/PhysRevLett.97.143002

D. A. Mazziotti, Anti-hermitian part of the contracted schrödinger equation for the direct calculation of two-electron reduced density matrices. Phys. Rev. A. **75**, 022505 (2007). https://doi.org/10.1103/PhysRevA.75.022505

D. A. Mazziotti, Exact two-body expansion of the many-particle wave function. Phys. Rev. A. **102**, 030802 (2020). https://doi.org/10.1103/PhysRevA.102.030802

S. McArdle, S. Endo, Quantum computational chemistry. Rev. Mod. Phys. **92**(1), 015003 (2020). https://doi.org/10.1103/revmodphys.92.015003

A. J. McCaskey, Z. P. Parks, J. Jakowski, S. V. Moore, T. D. Morris, T. S. Humble, R. C. Pooser, Quantum chemistry as a benchmark for near-term quantum computers. NPJ Quantum Inf. **5**(1), 99 (2019). https://doi.org/10.1038/s41534-019-0209-0

J. R. McClean, R. Babbush, P. J. Love, A. Aspuru-Guzik, Exploiting Locality in Quantum Computation for Quantum Chemistry. J. Phys. Chem. Lett. **5**(24), 4368–4380 (2014). https://doi.org/10.1021/jz501649m

J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. Phys. Rev. A. **95**, 042308 (2017). https://doi.org/10.1103/PhysRevA.95.042308

J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, H. Neven, Barren plateaus in quantum neural network training landscapes. Nat. Commun. **9**(1), 4812 (2018). https://doi.org/10.1038/s41467-018-07090-4. 1803.11173

J. R. McClean, J. Romero, R. Babbush, A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms. New J. Phys. **18**(2), 023023 (2016). https://doi.org/10.1088/1367-2630/18/2/023023

M. Metcalf, N. P. Bauman, K. Kowalski, W. A. de Jong, Resource-efficient chemistry on quantum computers with the variational quantum eigensolver and the double unitary coupled-cluster approach. J. Chem. Theory Comput. **16**(10), 6165–6175 (2020). https://doi.org/10.1021/acs.jctc.0c00421. PMID: 32915568. https://doi.org/10.1021/acs.jctc.0c00421

W. Mizukami, K. Mitarai, Y. O. Nakagawa, T. Yamamoto, T. Yan, Y.-y. Ohnishi, Orbital optimized unitary coupled cluster theory for quantum computer. Phys. Rev. Research. **2**, 033421 (2020). https://doi.org/10.1103/PhysRevResearch.2.033421

M. Motta, C. Sun, A. T. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. Brandão, G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. Nat. Phys. **16**(2) (2019). https://doi.org/10.1038/s41567-019-0704-4

D. Mukherjee, W. Kutzelnigg, Irreducible brillouin conditions and contracted schrödinger equations for n -electron systems. i. the equations satisfied by the density cumulants. J. Chem. Phys. **114**(5), 2047–2061 (2001). https://doi.org/10.1063/1.1337058

Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, K. M. Beck, A. Blinov, V. Chaplin, M. Chmielewski, C. Collins, S. Debnath, K. M. Hudek, A. M. Ducore, M. Keesan, S. M. Kreikemeier, J. Mizrahi, P. Solomon, M. Williams, J. D. Wong-Campos, D. Moehring, C. Monroe, J. Kim, Ground-state energy estimation of the water molecule on a trapped-ion quantum computer. npj Quantum Inf. **6**(1), 33 (2020). https://doi.org/10.1038/s41534-020-0259-3

P. J. Ollitrault, A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, I. Tavernelli, Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor. Phys. Rev. Research. **2**, 043140 (2020). https://doi.org/10.1103/PhysRevResearch.2.043140

P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, J. M. Martinis, Scalable quantum simulation of molecular energies. Phys. Rev. X. **6**, 031007 (2016). https://doi.org/10.1103/PhysRevX.6.031007

S. Pal, Use of a unitary wavefunction in the calculation of static electronic properties. Theo. Chim. Acta. **66**(3), 207–215 (1984). https://doi.org/10.1007/BF00549670

R. M. Parrish, P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation. arXiv (2019). 1909.08925. Accessed 03 Oct 2021

F. Peeters, J. Devreese, Upper bounds for the free energy. A generalisation of the Bogolubov inequality and the feynman inequality. J. Phys. A. **17**(3), 625 (1984)

B. Peng, K. Kowalski, Variational quantum solver employing the PDS energy functional. Quantum. **5**, 473 (2021). https://doi.org/10.22331/q-2021-06-10-473

A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. **5**(1), 4213 (2014). https://doi.org/10.1038/ncomms5213

A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, P. J. Coles, Absence of Barren Plateaus in Quantum Convolutional Neural Networks. Phys. Rev. X. **11**(4), 041011 (2021). https://doi.org/10.1103/PhysRevX.11.041011, https://link.aps.org/doi/10.1103/PhysRevX.11.041011

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

J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. Quantum Sci. Technol. **4**(1), 014008 (2018). https://doi.org/10.1088/2058-9565/aad3e4

M. Roth, M. Ganzhorn, N. Moll, S. Filipp, G. Salis, S. Schmidt, Analysis of a parametrically driven exchange-type gate and a two-photon excitation gate between superconducting qubits. Phys. Rev. A. **96**, 062323 (2017). https://doi.org/10.1103/PhysRevA.96.062323

I. G. Ryabinkin, S. N. Genin, A. F. Izmaylov, Constrained variational quantum eigensolver: Quantum computer search engine in the Fock space. J. Chem. Theory Comput. **15**(1), 249–255 (2019). https://doi.org/10.1021/acs.jctc.8b00943

I. G. Ryabinkin, R. A. Lang, S. N. Genin, A. F. Izmaylov, Iterative qubit coupled cluster approach with efficient screening of generators. J. Chem. Theory Comput. **16**(2), 1055–1063 (2020). https://doi.org/10.1021/acs.jctc.9b01084

I. G. Ryabinkin, T.-C. Yen, S. N. Genin, A. F. Izmaylov, Qubit coupled cluster method: A systematic approach to quantum chemistry on a quantum computer. J. Chem. Theory Comput. **14**(12), 6317–6326 (2018). https://doi.org/10.1021/acs.jctc.8b00932

R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, L. DiCarlo, Experimental error mitigation via symmetry verification in a variational quantum eigensolver. Phys. Rev. A. **100**, 010302 (2019). https://doi.org/10.1103/PhysRevA.100.010302

R. Santagati, J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, D. P. Tew, X. Zhou, J. L. O'Brien, M. G. Thompson, Witnessing eigenstates for quantum simulation of hamiltonian spectra. Sci. Adv. **4**(1) (2018). https://doi.org/10.1126/sciadv.aap9646. https://advances.sciencemag.org/content/4/1/eaap9646.full.pdf

S. Shalev-Shwartz, O. Shamir, S. Shammah, in *Proceedings of the 34th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 70.* ed. by D. Precup, Y. W. Teh, Failures of gradient-based deep learning (PMLR, International Convention Centre, Sydney, Australia, 2017), pp. 3067–3075. http://proceedings.mlr.press/v70/shalev-shwartz17a.html

Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, K. Kim, Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. Phys. Rev. A. **95**, 020501 (2017). https://doi.org/10.1103/PhysRevA.95.020501

P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, Algorithms for quantum computation: discrete logarithms and factoring, (1994), pp. 124–134. https://doi.org/10.1109/SFCS.1994.365700

S. E. Smart, J.-N. Boyn, D. A. Mazziotti, Resolving Correlated States of Benzyne on a Quantum Computer with an Error-Mitigated Quantum Contracted Eigenvalue Solver (2021). 2103.06876. Accessed 15 Mar 2021

S. E. Smart, D. A. Mazziotti, Quantum-classical hybrid algorithm using an error-mitigating n-representability condition to compute the mott metal-insulator transition. Phys. Rev. A. **100**, 022517 (2019). https://doi.org/10.1103/PhysRevA.100.022517

S. E. Smart, D. A. Mazziotti, Quantum Solver of Contracted Eigenvalue Equations for Scalable Molecular Simulations on Quantum Computing Devices. Phys. Rev. Lett. **126**(7), 070504 (2021). https://doi.org/10.1103/physrevlett.126.070504. 2004.11416

A. Soldatov, Generalized variational principle in quantum mechanics. Int. J. Mod. Phys. B. **9**(22), 2899–2936 (1995)

N. H. Stair, R. Huang, F. A. Evangelista, A multireference quantum Krylov algorithm for strongly correlated electrons. J. Chem. Theory Comput. **16**(4), 2236–2245 (2020). https://doi.org/10.1021/acs.jctc.9b01125. PMID: 32091895. https://doi.org/10.1021/acs.jctc.9b01125

C. Sur, R. K. Chaudhuri, B. K. Sahoo, B. P. Das, D. Mukherjee, Relativistic unitary coupled cluster theory and applications. J. Phys. B. **41**(6), 065001 (2008). https://doi.org/10.1088/0953-4075/41/6/065001

T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, J. R. McClean, Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources. Phys. Rev. X. **10**, 011004 (2020). https://doi.org/10.1103/PhysRevX.10.011004

A. G. Taube, R. J. Bartlett, New perspectives on unitary coupled-cluster theory. Int. J. Quantum Chem. **106**(15), 3393–3401 (2006). https://doi.org/10.1002/qua.21198

H. L. Tang, V. O. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, S. E. Economou, Qubit-ADAPT-VQE: An Adaptive Algorithm for Constructing Hardware-Efficient Ansätze on a Quantum Processor. PRX Quantum. **2**(2), 020310 (2021). https://doi.org/10.1103/PRXQuantum.2.020310. https://link.aps.org/doi/10.1103/PRXQuantum.2.020310

A. V. Uvarov, J. D. Biamonte, On barren plateaus and cost function locality in variational quantum algorithms. J. Phys. A Math. Theor. **54**(24), 245301 (2021). https://doi.org/10.1088/1751-8121/abfac7

T. Volkoff, P. J. Coles, Large gradients via correlation in random parameterized quantum circuits. Quantum Sci. Technol. **6**(2), 025008 (2021). 10.1088/2058-9565/abd891. 2005.12200

H. Wang, S. Ashhab, F. Nori, Efficient quantum algorithm for preparing molecular-system-like states on a quantum computer. Phys. Rev. A. **79**, 042335 (2009). https://doi.org/10.1103/PhysRevA.79.042335

D. Wecker, M. B. Hastings, M. Troyer, Progress towards practical quantum variational algorithms. Phys. Rev. A. **92**, 042303 (2015). https://doi.org/10.1103/PhysRevA.92.042303

N. Wiebe, C. Granade, Efficient Bayesian phase estimation. Phys. Rev. Lett. **117**, 010503 (2016). https://doi.org/10.1103/PhysRevLett.117.010503

T.-C. Yen, R. A. Lang, A. F. Izmaylov, Exact and approximate symmetry projectors for the electronic structure problem on a quantum computer. J. Chem. Phys. **151**(16), 164111 (2019). https://doi.org/10.1063/1.5110682. 1905.08109

K. Zhang, M.-H. Hsieh, L. Liu, D. Tao, Toward trainability of quantum neural networks. arXiv (2020). 2011.06258. Accessed 03 Oct 2021

Publisher's Note