Quantum-optimal-control-inspired ansatz for variational quantum algorithms

Alexandre Choquette [1,2] Agustin Di Paolo [1] Panagiotis Kl. Barkoutsos [2] David Sénéchal [1] Ivano Tavernelli,[2] and Alexandre Blais [1,3]

[1]*Institut quantique and Département de physique, Université de Sherbrooke, Sherbrooke J1K 2R1 Quebec, Canada*
[2]*IBM Quantum, Zurich Research Laboratory, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*
[3]*Canadian Institute for Advanced Research, Toronto, Ontario M5G 1M1, Canada*

A central component of variational quantum algorithms (VQAs) is the state-preparation circuit, also known as ansatz or variational form. This circuit is most commonly designed such as to exploit symmetries of the problem Hamiltonian and, in this way, constrain the variational search to a subspace of interest. Here, we show that this approach is not always advantageous by introducing ansatzes that incorporate symmetry-breaking unitaries. This class of ansatzes, that we call quantum-optimal-Control-inspired ansates (QOCA), is inspired by the theory of quantum optimal control and leads to an improved convergence of VQAs for some important problems. Indeed, we benchmark QOCA against popular variational forms applied to the Fermi-Hubbard model at half-filling and show that our variational circuits can approximate the ground state of this model with high accuracy. We also show how QOCA can be used to find the ground state of the water molecule and compare the performance of our ansatz against other common choices used for chemistry problems. This work constitutes a first step towards the development of a more general class of symmetry-breaking ansatzes with applications to physics and chemistry problems.

I. INTRODUCTION

The rise of noisy intermediate-scale quantum processors [1,2] requires us to find novel algorithms designed to attenuate the effects of noise. Variational quantum algorithms (VQAs) are an example of such methods [3,4]. These algorithms make use of a quantum computer and a classical coprocessor to minimize a cost function specified by a problem Hamiltonian $\hat{H}_{\text{prob}}$. This minimization is achieved by preparing a state that approximates the ground state of $\hat{H}_{\text{prob}}$ on the quantum computer using an iterative procedure driven by the classical coprocessor. Importantly, and due to the variational nature of these algorithms, this approach has been shown to potentially be resilient against noise, and well suited to several applications including finance [5], pure mathematics [6], machine learning [7,8], optimization problems [9,10], quantum chemistry and materials [11–15], as well as quantum optics [16].

In VQAs, the state preparation requires the parametrization of a quantum circuit, referred to as the *ansatz* or *variational form*. Recently, a considerable amount of effort has been invested in designing ansatzes that preserve the symmetries of the problem Hamiltonian [17–21]. The goal of symmetry-preserving strategies is to constrain the variational search to a small space of interest, which in principle can improve the probability of convergence to the target state with fewer optimizer iterations.

In this work, we highlight shortcomings of this approach. We then provide an ansatz that goes beyond symmetry-preserving methods by introducing a set of unitaries that break the symmetries of the problem Hamiltonian. To achieve this, we borrow ideas from the theory of quantum optimal control, where fast and high-fidelity operations are achieved through the addition of time-dependent symmetry-breaking terms to the Hamiltonian. Focusing on fermionic systems, we incorporate such terms in a time-evolution-like ansatz [22] to obtain the quantum-optimal-control-inspired ansatz (QOCA). We benchmark this approach against common ansatzes found in the literature for the Fermi-Hubbard model and apply these ideas to the water molecule with minimal modifications. We find that, in most cases, this method produces approximations of the target ground state that are orders of magnitude more accurate then with previous approaches. To understand this improvement, we present evidence that QOCA allows for an exploration in a slightly larger Hilbert space than allowed by structured variational forms.

The paper is organized as follows: in Sec. II we discuss known approaches to ansatz design, while in Sec. III we introduce QOCA after a brief review of quantum optimal control theory. We also elaborate on our strategy for the selection of symmetry-breaking terms. In Sec. IV we explain how these terms can be incorporated into a variational ansatz for the Fermi-Hubbard model. Finally, we compare results obtained with the different approaches in Sec. V.

II. VARIATIONAL ANSATZES

In the VQA framework, a quantum processor stores a quantum state $|\psi(\boldsymbol{\theta})\rangle$ parametrized by a collection of classical variational parameters $\boldsymbol{\theta}$. This state is prepared from a known

initial state, $|\psi_0\rangle$ using a quantum circuit (the ansatz) $\hat{U}(\boldsymbol{\theta})$ such that $|\psi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})|\psi_0\rangle$. The value of $\boldsymbol{\theta}$ is iteratively adjusted by a classical coprocessor with the purpose of minimizing the cost function

$$E[\boldsymbol{\theta}] = \frac{\langle\psi(\boldsymbol{\theta})|\hat{H}_{\text{prob}}|\psi(\boldsymbol{\theta})\rangle}{\langle\psi(\boldsymbol{\theta})|\psi(\boldsymbol{\theta})\rangle}. \quad (1)$$

Numerous variational forms $\hat{U}(\boldsymbol{\theta})$ have been explored in the literature [18,22–27]. Before introducing our approach, in this section we briefly review two widely used ansatzes highlighting their advantages and disadvantages.

A. Hardware-efficient ansatz

The hardware-efficient ansatz (HEA), introduced in Ref. [23], relies on gates that are native to the quantum hardware to produce circuits of high expressibility [28] and low depth. In particular, the HEA requires the application of successive blocks of parametrized single-qubit rotations followed by a generic entangling unitary $\hat{U}_{\text{Ent}}$. An example for N qubits is

$$\hat{U}_{\text{HEA}}(\boldsymbol{\theta}) = \prod_d \hat{U}_{\text{Ent}} \prod_{n=1}^{N} \hat{R}_Z^{(n)}(\theta_{n,d}^Z) \hat{R}_Y^{(n)}(\theta_{n,d}^Y), \quad (2)$$

where $\boldsymbol{\theta} = \{\theta_{n,d}^Z, \theta_{n,d}^Y\}$ groups all the variational parameters and $\hat{R}_{\hat{\sigma}_a}^{(n)}(\theta) = \exp[-i\theta\hat{\sigma}_a/2]$ denotes a single-qubit rotation of angle θ around the $a \in \{x, y, z\}$ axis on qubit n. $\hat{\sigma}_a$ is the corresponding Pauli matrix. The parameter d is the number of layers, or *depth*, of the ansatz. Here and for the rest of this paper, we use the convention $\prod_i^N \hat{U}_i = \hat{U}_N \cdots \hat{U}_1$ for operator multiplication.

A feature of the HEA is that it facilitates a broad exploration of the Hilbert space since it does not purposely favor a particular symmetry sector. This ansatz has already been experimentally implemented to prepare the ground state of small molecules [23], to simulate the folding of a few amino acid polymers [15], and to find the solution of classical optimization problems [10]. However, solving small instances of important problems does not provide a proof of scalability of the method for larger systems. Indeed, there is evidence that sufficiently random parametrized circuits, such as the ones produced by HEA, suffer from an exponentially vanishing gradient with the number of qubits making them harder to converge as the system size grows [29].

B. Variational Hamiltonian ansatz

Ansatzes that leverage the structure of the problem can avoid the aforementioned scalability issues since they do not explore the full Hilbert space. Wecker *et al.* [22] introduced the variational Hamiltonian ansatz (VHA), which consists of a parametrized adaptation of the quantum circuit implementing time evolution under the problem Hamiltonian via Trotterization. In the VHA framework, the state-preparation unitary reads

$$\hat{U}_{\text{VHA}}(\boldsymbol{\theta}) = \prod_d \prod_j e^{i\theta_{j,d}\hat{H}_j}, \quad (3)$$

where $\boldsymbol{\theta} = \{\theta_{j,d}\}$ are the variational parameters and $\hat{H}_{\text{prob}} = \sum_j \hat{H}_j$ is the problem Hamiltonian expressed as the sum of noncommuting groups of terms labeled $\hat{H}_j$. The depth d is associated with each time increment of the Trotterization of the time-evolution operator. With appropriate choices of the grouping of the terms in the Hamiltonian, this approach can be implemented using few variational parameters, therefore simplifying the classical optimization. However, depending on the complexity of the problem, circuits can be considerably deeper as compared to those typically used with the HEA.

Fourier-transformed VHA (FT-VHA)

To further reduce the number of variational parameters, it is possible to take advantage of the fact that most fermionic Hamiltonians can be written as $\hat{H}_{\text{prob}} = \hat{T} + \hat{V}$, where the diagonal bases of $\hat{T}$ and $\hat{V}$ are related through the fermionic Fourier transformation (FT) [30–32]. With the FT-VHA variational form, the FT is used to alternate between these bases at every Trotter step. In the context of quantum chemistry, this is known as the split-operator method [33,34]. This idea was also recently introduced by Babbush *et al.* [35] for the variational quantum simulation of materials. The state-preparation unitary thereby reads

$$\hat{U}_{\text{FT-VHA}}(\boldsymbol{\tau}, \boldsymbol{\nu}) = \prod_d \text{FT}^\dagger \left(\prod_j e^{i\tau_{j,d}\hat{\mathcal{T}}_j}\right) \text{FT}\left(\prod_j e^{i\nu_{j,d}\hat{V}_j}\right), \quad (4)$$

where $\boldsymbol{\tau} = \{\tau_{j,d}\}$ and $\boldsymbol{\nu} = \{\nu_{j,d}\}$ are the parameters associated with $\hat{\mathcal{T}} = \sum_j \hat{\mathcal{T}}_j = \text{FT} \, \hat{T} \, \text{FT}^\dagger$ and $\hat{V}$, respectively. Since now both $\hat{\mathcal{T}}$ and $\hat{V}$ are diagonal, they contain only terms that commute. For this reason, the circuit decomposition of their exponentials can be achieved exactly, which was not the case of $\hat{T}$ in the regular VHA. However, this comes at the cost of the long FT circuit [30,32,35].

Because they are built from the problem Hamiltonian, both VHA and FT-VHA preserve the symmetries of the problem. For example, if no term of $\hat{H}_{\text{prob}}$ allows the number of particles to change, this quantity will be conserved in the variational state $|\psi(\boldsymbol{\theta})\rangle$. This choice restricts the variational search to a relatively small subspace of the Hilbert space which, intuitively, can increase the performance of the VQA. Because of this, the VHA and FT-VHA ansatzes are likely to perform better than HEA for large system sizes. However, as we show in Sec. V, always preserving symmetries of the problem can also be detrimental.

Another approach used in the context of quantum chemistry is the UCCSD ansatz [3]. Although not strictly Hamiltonian-based, this method preserves the parity symmetry of fermions and conserves the number of particles. Despite providing accurate results, the UCCSD ansatz circuits can be very deep, limiting its applicability on near-term quantum devices.

III. QUANTUM-OPTIMAL-CONTROL-INSPIRED ANSATZ (QOCA)

To address the drawbacks of the ansatzes discussed above, we propose an ansatz that borrows ideas from the theory of quantum optimal control [36–39], and which we therefore

dub the quantum-optimal-control-inspired ansatz, or QOCA. The main idea behind QOCA resides in the introduction of carefully chosen symmetry-breaking unitaries into the symmetry-preserving ansatz VHA. Further connections between quantum control and variational quantum algorithms were established in Ref. [40]. In this section, we begin by reviewing some of the central aspects of the theory of quantum optimal control, and then show how these ideas can be incorporated in the design of variational forms.

A. Quantum optimal control

Quantum optimal control (QOC) theory describes methods to optimally steer a quantum system from an initial state to a known final state [41]. Such techniques have been applied to a wide variety of problems including the quantum control of chemical reactions [42,43], spins in nuclear magnetic resonance experiments [38,44] and, more recently, to superconducting qubits [39,45].

In this approach, the control Hamiltonian is specified by a set of time-independent drive Hamiltonians $\{\hat{H}_k\}$ whose amplitudes are parametrized by the time-dependent coefficients $\{c_k(t)\} \in \mathbb{R}$. For example, these can represent the action of electromagnetic control pulses on the system. The total Hamiltonian $\hat{H}(t)$ is then, in general, time-dependent taking the form

$$\hat{H}(t) = \hat{H}_0 + \sum_k c_k(t)\hat{H}_k, \qquad (5)$$

with $\hat{H}_0$ the free, or drift, Hamiltonian of the controlled system. Note that typically $[\hat{H}_0, \hat{H}_k] \neq 0$. Solving the Schrödinger equation of the driven system results in the unitary $\hat{U}(t)$, which can propagate pure states through time as $|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle$.

The system described by the Hamiltonian of Eq. (5) is said to be fully controllable if for any initial state $|\psi(0)\rangle$, there exist a set controls $\{c_k(t)\}$ and a time $T > 0$ for which the state $|\psi(T)\rangle$ can be any target state of the Hilbert space [41].

Quantum optimal control techniques, such as the GRAPE algorithm [38], provide a method for designing the control pulses $c_k(t)$ to achieve a desired state preparation. This is usually realized by seeking the set of controls and time T that optimize a cost function characterizing the state-preparation fidelity, which may include constraints such as the control time and the maximum pulse amplitudes.

In the GRAPE algorithm, time is discretized into N increments, or pixels, of duration Δt such that the total evolution occurs in a time $T = N\Delta t$. Using this discretization, the continuous control fields $c_k(t)$ are now parametrized by the new constant piecewise control fields $\{u_{k,j}\}$ as

$$c_k(t) = \sum_{j=0}^{N-1} u_{k,j} \sqcap_j (t, \Delta t), \qquad (6)$$

where $\sqcap_j(t, \Delta t) \equiv \Theta(t - j\Delta t) - \Theta(t - (j+1)\Delta t)$ with Θ the Heaviside function. The time evolution operator for a time T therefore reads

$$\hat{U}(T) = \prod_{j=0}^{N-1} \exp\left[-i\Delta t\left(\hat{H}_0 + \sum_k u_{k,j}\hat{H}_k\right)\right], \qquad (7)$$

and an optimal control pulse shape is obtained by iteratively tuning the values of the discrete control fields $\{u_{k,j}\}$.

As pointed out in Ref. [40], QOC and VQAs are both quantum-classical optimization routines where minimization is achieved over pulse amplitudes in the former, and gate parameters in the later. In the next section, we tighten this connection in the special case where the VQA ansatz is itself inspired by the methods of QOC.

B. The QOCA variational form

Building on the concept of quantum optimal control, we identify a set of key ingredients that define a QOC-inspired VQA instance (QOCA). Formally, one must specify (1) a problem Hamiltonian $\hat{H}_{\text{prob}}$ whose ground state we wish to prepare, (2) a set of *drive* terms $\{\hat{H}_k\}$, where, by design, $[\hat{H}_{\text{prob}}, \hat{H}_k] \neq 0 \ \forall \ k$, (3) a variational ansatz built from $\hat{H}_{\text{prob}}$ and $\{\hat{H}_k\}$, (4) a cost function, and, finally, (5) an optimization procedure.

More specifically, to build the QOCA variational form we begin by introducing the *control*-like Hamiltonian, similar to Eq. (5), using $\hat{H}_{\text{prob}}$ and $\{\hat{H}_k\}$,

$$\hat{H}_{\text{QOCA}} = \hat{H}_{\text{prob}} + \sum_k \hat{H}_k. \qquad (8)$$

From Eq. (8), we derive the following time-evolution-like parametrized circuit ansatz

$$\hat{U}_{\text{QOCA}}(\boldsymbol{\theta}, \boldsymbol{\delta}) = \prod_d \left(\prod_j e^{i\theta_{j,d}\hat{H}_j} \prod_k e^{i\delta_{k,d}\hat{H}_k} \right), \qquad (9)$$

where $\hat{H}_{\text{prob}} = \sum_j \hat{H}_j$ and $\boldsymbol{\theta} = \{\theta_{j,d}\}$ are the problem Hamiltonian parameters. Similar to Eqs. (6) and (7), $\boldsymbol{\delta} = \{\delta_{k,d}\}$ are the variational parameters associated with the *drive* Hamiltonians $\{\hat{H}_k\}$. Again, d is the depth of the ansatz and is analog to the steps in the time evolution of Eq. (7).

In a purely QOC setting, the cost function is often chosen to be the state or process fidelity. This choice is however not appropriate for VQAs since the target ground state is generally unknown. Instead, we use the expectation value of the problem Hamiltonian as in Eq. (1) which corresponds to an energy measurement in a quantum physics or chemistry problem.

While in QOC optimization is on the amplitude of the control fields, for QOCA we directly optimize the gate angles associated with the different terms of Eq. (8). Moreover, note that in QOC we do not have control over the drift Hamiltonian; however, in the QOCA setting the optimization controls both the problem and drive Hamiltonians.

The key intuition behind QOCA is that the problem Hamiltonian $\hat{H}_{\text{prob}}$ helps in constraining the variational search to the relevant symmetry sector of the Hilbert space, while the drives $\{\hat{H}_k\}$ potentially allow the ansatz to take shortcuts by temporarily exiting this sector. This concept is schematically illustrated in Fig. 1(b) where we illustrate possible paths in the Hilbert space for the HEA, VHA, and QOCA variational forms. We provide numerical evidence for this intuition in Sec. V.

In principle, one has the freedom to select any drive Hamiltonians that do not commute with $\hat{H}_{\text{prob}}$. QOCA can therefore
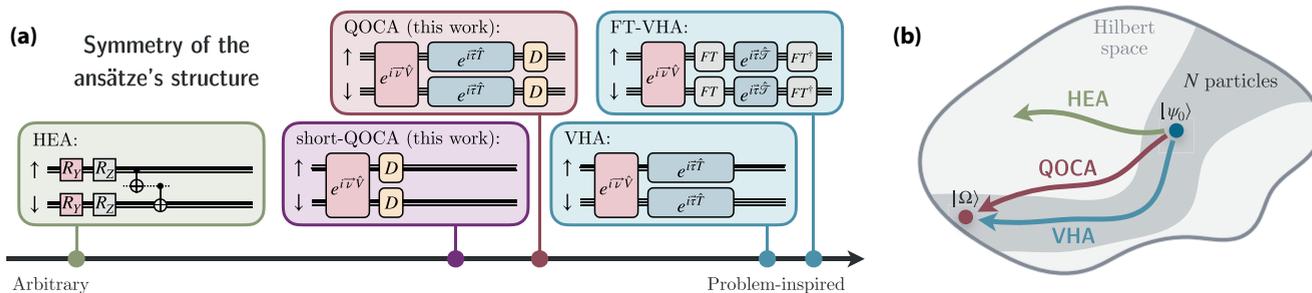
FIG. 1. (a) Single circuit layer of the ansatzes studied in this work arranged according to the symmetry of their structure. Here high symmetry implies that the ansatz is completely built around the problem Hamiltonian, while the lack of symmetries reflects the arbitrariness of its circuit. We show the hardware-efficient ansatz (HEA), the variational Hamiltonian ansatz (VHA), the Fourier-transformed VHA (FT-VHA), the quantum-optimal-control-inspired ansatz (QOCA), along with a shallower version of QOCA, the short-QOCA ansatz. The horizontal lines represent the qubit registers that encode the spin orbitals associated with the ↑ or ↓ spins. For HEA, the entangling block is a stair of CNOTs indicated by the dashed qubit line. For all other ansatzes, $\hat{T}$ and $\hat{V}$ are, respectively, the kinetic and interaction parts of the problem Hamiltonian and $\{\tau, \nu\}$ are their associated variational parameters. For FT-VHA, we have that $\hat{\mathcal{T}} = \text{FT} \, \hat{T} \, \text{FT}^\dagger$. The subcircuits D represent the implementation of the drives. (b) Possible paths in the Hilbert space for the HEA, VHA, and QOCA variational forms. The initial state $|\psi_0\rangle$ and the target state $|\Omega\rangle$ are in the same symmetry sector containing N particles. Since HEA does not preserve the symmetries of $\hat{H}_{\text{prob}}$, its path easily escapes from the fixed particle number subspace, while VHA is restricted to it. By introducing symmetry-breaking terms, QOCA may have the ability to find shortcuts in Hilbert space, therefore avoiding local extrema.

be seen as a QOC experiment that is simulated on a quantum computer, where the drives need not be related to the physical interactions with the system of interest. However, it is not straightforward to predict which choice of $\{\hat{H}_k\}$ will have the most positive impact on the outcome of the VQA. One option is to use an adaptive approach such as the one described in Refs. [46,47]. In the next section we show how simple considerations can help to bound the number of interesting drive operators, and suggest which of these could be most effective.

C. Which drive Hamiltonians are useful for fermionic problems?

With the objective of applying QOCA to fermionic systems, we consider the time-dependent fermionic Hamiltonian

$$
\hat{H}_f(t) = \sum_j [\alpha_j(t)\hat{a}_j + \alpha_j^*(t)\hat{a}_j^\dagger] \\
+ \sum_{i,j} \beta_{ij}(t)(\hat{a}_i^\dagger \hat{a}_j + \hat{a}_j^\dagger \hat{a}_i) + \sum_{i,j} \gamma_{ij}(t)\hat{a}_i^\dagger \hat{a}_i \hat{a}_j^\dagger \hat{a}_j,
$$

(10)

where $\hat{a}_j^\dagger$ and $\hat{a}_j$ are fermionic ladder operators of spin orbital j respecting the anticommutation relations $\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}$ and $\{\hat{a}_i, \hat{a}_j\} = \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0$. Importantly, $\hat{H}_f(t)$ is fully controllable in the sense that any unitary matrix can be generated by solving its Schrödinger equation for a finite time [48,49].

We note that while the first term of $\hat{H}_f(t)$ is unphysical since it breaks the parity symmetry of fermions, the quadratic and quartic terms occur in many physical models. This makes $\hat{H}_f(t)$ attractive for designing driven physically inspired ansatzes, as terms of the form $\alpha(t)\hat{a} + \alpha^*(t)\hat{a}^\dagger$ do not commute with the physical problem Hamiltonian. Interestingly, the use of such terms has been proposed in the

context of variational error suppression [4] as they may allow a variational state to "re-enter" a particular symmetry sector to correct for the effect of symmetry-breaking errors.

As a first example of QOCA variational forms, we therefore propose to use the first term of $\hat{H}_f(t)$ as symmetry-breaking Hamiltonian. This choice is, however, not restrictive, and future work will investigate a broader class of drive Hamiltonians. We also note that our approach can easily be extended to the simulation of nonfermionic Hamiltonians.

IV. QOCA FOR THE FERMI-HUBBARD MODEL

For completeness, we start this section by reviewing the Fermi-Hubbard model and explain how we use the QOCA variational form to approximate its ground state. We motivate our choice of initial state, and elaborate on the selection and circuit decomposition of the drive terms. Finally, we introduce short-QOCA, a variant of QOCA that yields shorter circuits by dropping some terms of $\hat{H}_{\text{prob}}$ from the Hamiltonian that generates the regular QOCA variational form.

A. The Fermi-Hubbard model (FHM)

The Fermi-Hubbard model is an iconic model in the study of strongly correlated materials [50]. It describes interacting spin-$\frac{1}{2}$ fermions on a lattice where each site can be occupied by up to two particles of opposite spins. The Hamiltonian of the FHM for L lattice sites takes the form

$$
\hat{H}_{\text{FHM}} = \underbrace{-t \sum_{\langle i,j \rangle, \sigma} \hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma}}_{\equiv \hat{T}} + \underbrace{U \sum_{i=1}^L \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} - \mu \sum_{i,\sigma} \hat{n}_{i\sigma}}_{\equiv \hat{V}}, \quad (11)
$$

where i, j are the lattice-site indices, and $\sigma = \{\uparrow, \downarrow\}$ labels the spin degree of freedom. In the first term, $\langle i, j \rangle$ denotes a sum over nearest-neighbor sites, and $\hat{n}_{i\sigma} = \hat{a}_{i\sigma}^\dagger \hat{a}_{i\sigma}$ is the occupation operator of the spin orbital labeled $i\sigma$.

The first term of Eq. (11) represents hopping between neighboring sites with amplitude $-t$ and will be referred to as $\hat{T}$. This term is diagonal in momentum space if periodic boundary conditions are imposed, and its ground state consists of delocalized plane waves. The second term is a nonlinear, on-site Coulomb repulsion of strength U, while the last term is the chemical potential. These last two terms are diagonal in the position basis and, taken together, are denoted $\hat{V}$. The ground state of $\hat{V}$ is described by wave functions localized on the sites.

A particularly interesting instance of the FHM is the half-filling regime (which occurs for $\mu = U/2$) at intermediate coupling, $U/t \sim 4$. In this regime, both $\hat{T}$ and $\hat{V}$ contribute significantly to the system's energy, thus creating competition between the localized and delocalized states of the electrons, leading to rich physics such as the Mott transition. Because it becomes impossible to accurately treat either part of the Hamiltonian perturbatively, exact numerical diagonalization of the FHM is difficult beyond 24 lattice sites at half-filling [51]. To benchmark our variational form, we work in this particularly challenging regime.

Despite its apparent simplicity, the Fermi-Hubbard model has been used to study systems ranging from heavy fermions [52] to high-temperature superconductors [53,54]. As a result, it is an interesting problem to benchmark near-term quantum computers [55], and a useful performance test for variational ansatzes. For these reasons, variational quantum algorithms have already been used to find the ground state of the FHM, for example, using the HEA variational form [56], the VHA [21,22,57,58], and other symmetry-preserving ansatzes [21,24,55,59,60].

B. Encoding, parametrization and optimization of the ansatzes

We use the Jordan-Wigner (JW) transformation to encode fermionic Fock states into qubit registers, as detailed in Appendix A. Moreover, we work in real space and order the basis vectors for the $2L$ spin orbitals as $|f_{1\uparrow} \cdots f_{L\uparrow}; f_{1\downarrow} \cdots f_{L\downarrow}\rangle$ with $f_p \in \{0, 1\}$ the occupation of orbital p.

Using this purely conventional choice, in Fig. 1(a) we schematically draw one layer of the circuits implementing the different ansatzes discussed above and arranged by the symmetry of their structure. A highly symmetric ansatz is completely built around $\hat{H}_{\text{prob}}$ while a weakly symmetric construction is arbitrary with respect to the problem.

To parametrize these circuits, we consider two possible strategies: one corresponding to full parametrization of the single- and two-qubit gates and the other having a number of parameters that grows only with the depth of the ansatz, but not with the number of qubits. Whenever used, the latter is specified with the label *scalable*. Both strategies are elaborated on in Appendix B, and details of the numerical simulation are presented in Appendix C.

In most cases, the parameters are initialized at zero, and we use the COBYLA optimizer [61–63], which operates deterministically. The reproducibility of our simulations is therefore guaranteed over repeated runs, avoiding the necessity to average over many simulations. Moreover, we use the state-vector simulator from Qiskit [64], which outputs the full

wave function and consequently no measurement statistics is needed to analyze our results.

C. Initial state

In general, the performance of VQAs strongly depends on the choice of initial state and value of the variational parameters. The initial state acts as an educated guess to the target state and is often chosen such as to be easily computable classically. Moreover, because the initialization stage of a variational algorithm should be straightforward for the quantum processor or otherwise be treated as a separate routine [65], we are interested in benchmarking the performance of the QOCA variational form for the N-qubit simple initial state

$$|\psi_0\rangle = H^{\otimes N}|0\rangle = |+\rangle^{\otimes N}, \tag{12}$$

where H is the Hadamard gate. In addition to being easy to prepare, this initial state corresponds to half-filling and zero total spin, placing it in the same symmetry sector as the target state.

While this choice allows us to demonstrate the usefulness of the QOCA variational form given unstructured, simple initial conditions, we also show how the convergence can be improved further by using the ground state of the noninteracting FHM fixing $U = \mu = 0$ in Eq. (11) as initial state. More details on how to prepare this more complex state are provided in Appendix D.

D. Drive Hamiltonians

To reduce the number of variational parameters, we take $\alpha_j(t) = 1 + i$ in Eq. (10) leading to the drive Hamiltonians

$$\hat{H}_1 = \sum_{j=1}^{L} (\hat{a}_j^\dagger + \hat{a}_j), \tag{13}$$

$$\hat{H}_2 = \sum_{j=1}^{L} i(\hat{a}_j^\dagger - \hat{a}_j). \tag{14}$$

We use these drives on all sites, and for both up and down spins. Performing the JW transformation on Eqs. (13) and (14) leads to

$$\hat{H}_1 \mapsto \sum_{j=1}^{L} \hat{X}_j \bigotimes_{l<j} \hat{Z}_l, \tag{15}$$

$$\hat{H}_2 \mapsto \sum_{j=1}^{L} \hat{Y}_j \bigotimes_{l<j} \hat{Z}_l, \tag{16}$$

where $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ are Pauli matrices. To incorporate these expressions into the QOCA variational form of Eq. (9), we perform a first-order Trotter-Suzuki decomposition, arriving at the drive circuit equation for the dth layer of the ansatz,

$$\prod_{k=1,2} e^{i\delta_{k,d}\hat{H}_k} \approx \prod_{j=1}^{L} \exp\left[i\delta_{1,d}\,\hat{X}_j \bigotimes_{l<j} \hat{Z}_l\right]$$

$$\times \exp\left[i\delta_{2,d}\,\hat{Y}_j \bigotimes_{l<j} \hat{Z}_l\right], \tag{17}$$
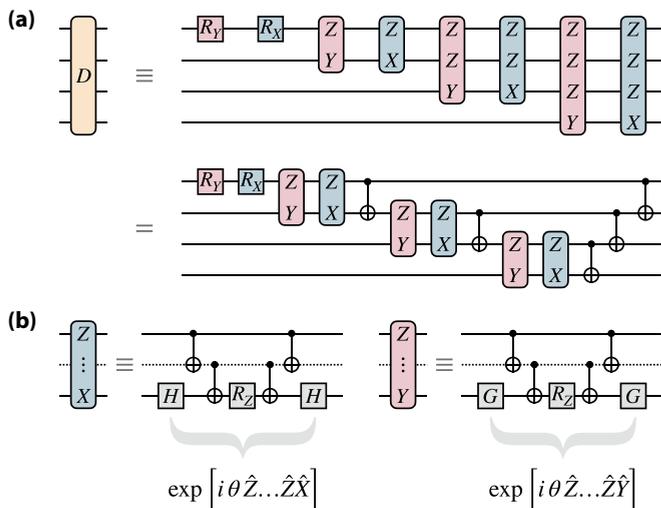
FIG. 2. (a) Circuit decomposition of the drive of Eq. (17) used for QOCA. This circuit generalizes to any number of qubits by appending additional $\hat{Z} \ldots \hat{Z}\hat{Y}$ and $\hat{Z} \ldots \hat{Z}\hat{X}$ multiqubit gates at the end. We also show the circuit compiled to one- and two-qubit (CNOT) gates. (b) Decomposition of the multiqubit gates based on a conventional approach to decompose exponentials of Pauli strings into circuits of CNOTs described in [66]. The transformation $H = (\hat{X} + \hat{Z})/\sqrt{2}$ is the Hadamard gate which changes between the $\hat{X}$ and $\hat{Z}$ bases and $G = (\hat{Y} + \hat{Z})/\sqrt{2}$ is the equivalent transformation between the $\hat{Y}$ and $\hat{Z}$ bases. The angles of the rotations $R_a(\theta) = \exp[-i\theta\hat{\sigma}_a/2]$ are the variational parameters, where $\hat{\sigma}_a$ is a Pauli matrix. The dashed qubit line may comprise multiple qubits on which are applied descending and ascending stairs of CNOTs on either side of the R_Z rotations.

where $\{\delta_{k,d}\}$ are the variational parameters associated with the kth drive term of that layer. A schematic of the circuit implementing Eq. (17) for four qubits is illustrated in Fig. 2, where we also show a compiled version of the circuit.

E. The short-QOCA variational form

Depending on the form of the drive, the quantum circuits corresponding to QOCA can be long. There exists, however, a practical approach to reduce the circuit depth without compromising the performance substantially.

Because the drive D in Fig. 2 and the kinetic part of the FHM (11) are both block-diagonal in the spin degree of freedom, we chose to remove the latter term from the Hamiltonian that generates the ansatz (8), arriving at the simplified form

$$\hat{U}_{\mathrm{sQOCA}}(\boldsymbol{\nu}, \boldsymbol{\delta}) = \prod_d \left(\prod_j e^{i\nu_{j,d}\hat{V}_j} \prod_k e^{i\delta_{k,d}\hat{H}_k} \right), \qquad (18)$$

where $\hat{V} = \sum_j \hat{V}_j$ is the on-site interaction part of the Fermi-Hubbard Hamiltonian (11), and $\boldsymbol{\nu} = \{\nu_{j,d}\}$ are the associated variational parameters. Since implementing $e^{i\theta\hat{T}}$ is more costly in terms of two-qubit gates than $e^{i\theta\hat{V}}$, this simplification reduces the circuit depth by a factor of two. We refer to this simplified version of the QOCA variational form as *short-QOCA*; see Fig. 1.

V. NUMERICAL RESULTS

We now present results obtained from numerical simulations of QOCA and short-QOCA for the Fermi-Hubbard model, and contrast these results with those obtained with the other possible ansatzes. As an illustration of the use of QOCA beyond the Fermi-Hubbard model, we also present a comparison of the performance of this ansatz over a hardware-efficient approach and the UCCSD ansatz for a 12-qubit representation of the H_2O molecule.

While our VQA cost function is the energy of the variational state $|\psi(\boldsymbol{\theta})\rangle$ given in Eq. (1), throughout this section we rather report the fidelity

$$\text{Fidelity} = |\langle\psi(\boldsymbol{\theta})|\Omega\rangle|^2, \qquad (19)$$

of the prepared state with respect to the target $|\Omega\rangle$. This metric is a more appropriate measure of the quality of the trial wave function than the energy.

A. Fermi-Hubbard model

We consider 2×2 (8 qubits) and 2×3 (12 qubits) lattices of the Fermi-Hubbard model at half-filling with open boundary conditions. As the former configuration can be seen as a periodic 1×4 chain, we also explore this case using the FT-VHA variational form which assumes periodic boundary conditions. Importantly, we find that for smaller systems such as the four-qubit 2×1 dimer, all ansatzes converge in a few tens of iterations on the ground-state energy with a precision $<10^{-7}$ using a single ansatz layer, $d = 1$, except for the HEA which requires two layers.

1. Comparing the ansatzes

For systems with four and six fermionic sites, we observe important variations in the ability of the different ansatzes to converge to the ground state energy. This is illustrated in Fig. 3, which shows, for all approaches, the final state infidelity as a function of the number of ansatz layers, d, initialized with the simple half-filled state of Eq. (12). The maximum fidelities achieved for all ansatzes are reported in Table I along with resource counts using a circuit compilation in terms of CNOTs. The gate count for the different wave function ansätze reported in Table I indicates a clear tradeoff between the circuit depth and the achievable accuracy. For the QOCA ansatzes, the number of CNOTs required is above the capabilities of current, state-of-the-art, noisy hardware (limited by gate fidelities and decoherence time), but the obtained level of accuracy nevertheless places QOCA among the most favorable ansatzes for future near-term quantum computers.

We first note that VHA and FT-VHA perform poorly for both system sizes and that their performance does not improve with the addition of more entangling layers, i.e., by increasing d. Because these ansatzes are fermion-number preserving, this observation suggests that VHA and FT-VHA may not efficiently search over all states of fixed particle number in the variational landscape. Moreover, since FT-VHA performs similarly to VHA for the 2×2 system, we also conclude that alternating bases with the fermionic Fourier transform does not yield superior results for these lattice sizes.

TABLE I. Maximum fidelities with respect to the ground state of the FHM, attained for d ansatz layers, requiring a number n_θ of variational parameters. The total gate depth is also indicated. The latter estimate assumes an all-to-all connectivity, and the same compiling procedure is used for all ansatzes.

Hubbard model	Maximum fidelity	d	n_θ	Gate depth
2×2(8 qubits)				
HEA	0.9876	9	144	72
VHA	0.1343	8	64	320
FT-VHA	0.1315	7	56	588
QOCA	0.9999	4	64	272
QOCA (scalable)	0.9992	10	50	680
Short-QOCA	0.9999	9	108	279
2×3(12 qubits)				
HEA	0.7276	10	240	120
VHA	0.0804	10	130	740
QOCA	0.9965	9	225	1098
QOCA (scalable)	0.8822	10	60	1220
Short-QOCA	0.7476	8	144	408

Interestingly, QOCA systematically reaches the ground state of the Fermi-Hubbard model with significantly more accuracy than VHA for both system sizes, indicating that the additional symmetry-breaking terms help the convergence. This advantage persists even when reducing the number of variational parameters from 16 to 5 per layer in the case of the scalable parametrization of QOCA, which converged with 0.9992 fidelity at $d = 10$ for the 2×2 system. The hardware-efficient approach also performs better than VHA, although it uses considerably more parameters than all other ansatzes. The increase in the number of variational parameters as a

function of the system size for such heuristic ansatzes can be seen as a potential limiting factor for the application of QOCA. However, there is evidence that when a small error in the total energy is tolerated, the number of parameters scales only polynomially [27], allowing for an efficient optimization towards an approximated ground-state wave function in near-term quantum devices.

Data obtained with the short-QOCA variational form shows that the QOCA circuits can be substantially shortened by removing more than half of the two-qubit gates at every step without much compromise on the performance for small systems. In fact, for the 2×2 Hubbard model, a fidelity of 0.9999 is achieved with nine layers of this ansatz.

With improved fidelities for shallower circuits which use fewer variational parameters than standard approaches, we find that QOCA provides significant gain with respect to other common ansatzes.

2. The benefits of breaking symmetries

Figura 4 shows the evolution of the average number of particles per lattice site (top panel) and the infidelity of the variational state with respect to the target state (bottom panel) throughout the optimization process for the same simulations as in Fig. 3.

Focusing first on the top panel, we first note that, because the initial state $|+\rangle^{\otimes N}$ is half-filled, all variational states begin in the correct particle-number symmetry sector of the Hilbert space with $\langle \hat{N} \rangle / L = 1$, where L is the number of fermionic sites. Because VHA does not contain terms that allow the particle number to change, this quantity is constant throughout the optimization. We hypothesize that the poor performance of this ansatz in reaching the ground state is caused by the inability of this variational form to overcome local minima in parameter space.

In contrast, both parametrizations of QOCA allow the average site occupancy to deviate from $\langle \hat{N} \rangle / L = 1$ as the drive angles are being tuned away from zero by the optimizer. As seen in Fig. 4, this can lead to the sharp features observed in the first few $\sim 10^2$ iterations as the classical optimizer can
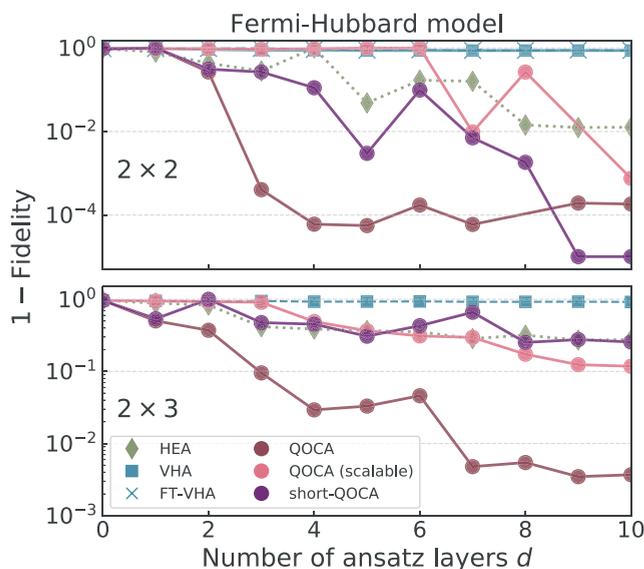
FIG. 3. Final variational state infidelities with respect to the target state as a function of the number of layers d of the variational forms of this work. Top panel is for a 2×2 plaquette, while the bottom panel is a 2×3 system. The initial state is $|+\rangle^{\otimes N}$ for all cases. Data at $d = 0$ correspond to the initial state alone, which has a fidelity of 0.035 with the target state. Unless specified, all ansatzes are fully parametrized according to Appendix B 1.
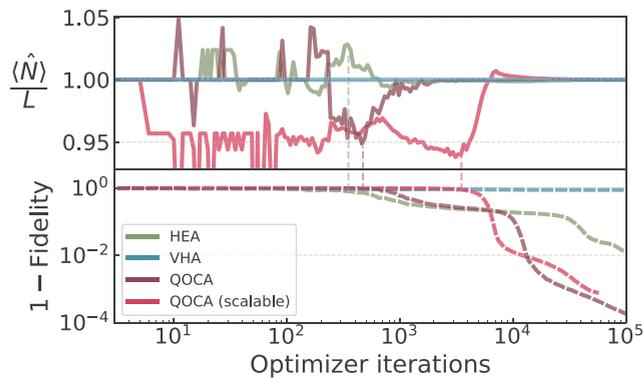
FIG. 4. Top: Average number of particles per lattice site in the variational state at every iteration of the VQA routine. $\langle \hat{N} \rangle = \sum_{i,\sigma} \langle \hat{n}_{i\sigma} \rangle$ is the total occupation and L is the number of sites. Bottom: Corresponding variational state infidelity, $1 - |\langle \psi(\boldsymbol{\theta})|\Omega \rangle|^2$, with respect to the ground state of the Fermi-Hubbard model $|\Omega\rangle$. The results are for a 2×2 system and the initial state is $|+\rangle^{\otimes N}$ for all ansatzes. Runs for ansatz depth $d = 9$ was used for HEA and $d = 10$ for the others, but this behavior is observed for most d.

initially overweight the value of drive terms. Over the full optimization, the number of particles deviates only slightly from the target value $\langle \hat{N} \rangle / L = 1$ with changes of only $\sim 5\%$ of the site occupancy. This is an indication that the symmetry-breaking terms in QOCA allow the ansatz to explore a Hilbert space that is slightly larger than the manifold of fixed particle number. Nevertheless, we find that these relatively small excursions out of the target symmetry sector can significantly ease convergence of the VQA. Indeed, we observe that the onset of the return to the target symmetry sector, as indicated by the vertical dashed lines in Fig. 4 is often associated with the abrupt descents in the infidelity, which may indicate that regions of steep gradients in parameter space are found.

This behavior is similarly observed for the hardware-efficient ansatz of Eq. (2) which also does not preserve the symmetries of $\hat{H}_{\text{prob}}$. This phenomenon is not particular to the realizations displayed in the figure, and it is generally observed for other system sizes and initial states.

We note, however, that these desired regions in parameter space would never be found if an error-mitigation technique based on symmetry verification were employed [19,67]. Indeed, in these schemes the variational states are postselected after the energy measurements only if they preserve desired symmetries of the target state. However, other strategies for error mitigation remain applicable [68–70].

3. Initial state

Because it provides a simple setting to benchmark the performance of the different ansatzes, we have so far considered only the easily prepared initial state of Eq. (12). Improved approximation to the ground state can be obtained if a more structured initial state is considered, although at the price of more complex state preparation circuits.

In Fig. 5 we compare the performance of the VHA and QOCA variational forms on the 2×2 lattice with the following initial states of increasing complexity: (1) the simple state
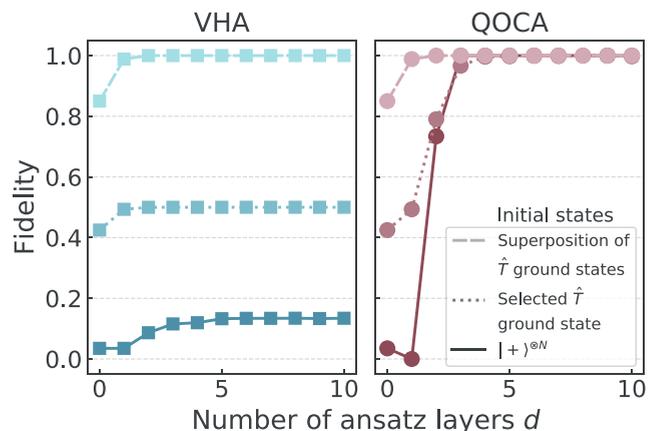
FIG. 5. Variational state fidelities with respect to the ground state of the 2×2 FHM as a function of the number of ansatz layers, d, for the VHA and QOCA variational forms. Results with three initial states are presented: (solid) Hadamard gates on every qubit $|+\rangle^{\otimes N}$, (dotted) a selected ground state of $\hat{T}$ corresponding to $|\Omega_T^{(1)}\rangle$ of Appendix D, and (dashed) a superposition of ground states of $\hat{T}$ corresponding to $|\Omega_T\rangle$ of Appendix D.

$|+\rangle^{\otimes N}$ already used above, (2) one of the degenerate ground states of $\hat{T}$ labeled $|\Omega_T^{(1)}\rangle$ in Appendix D, and (3) the superposition of ground states of $\hat{T}$ labeled $|\Omega_T\rangle$ in Appendix D.

Remarkably, while the final variational state obtained with VHA strongly depends on the initial state, QOCA systematically achieves convergence with fidelity > 0.9999 regardless of the initialization choice. Again because of its ability to move between symmetry sectors, these results illustrate QOCA's robustness to simple, unstructured, initial conditions that can have very small overlaps with the target ground state. For both variational forms, using a superposition of the degenerate ground states of $\hat{T}$ as initial state (dashed lines) leads to convergence with fewer entangling layers. This, however, comes at the cost of significantly increasing the complexity of the initialization stage of the VQA (see Appendix D).

B. Proof-of-principle implementation of the H$_2$O molecule

The previous section illustrates how QOCA can approximate the ground state of the FHM with systematically more accuracy than previously introduced ansatzes, even when faced with unstructured initial conditions. In order to investigate the broader applicability of this method, we now benchmark the QOCA variational form on a quantum chemistry problem. As a proof-of-principle test, we consider the H_2O molecule in its equilibrium configuration. Because we disregard the degrees of freedom corresponding to the core orbitals, this problem maps to 12 qubits using the STO3G basis set [71,72]. The Hamiltonian is obtained using the PySCF driver as provided by Qiskit Chemistry [64].

We compare the performance of QOCA against HEA together with the well-known chemistry-inspired UCCSD ansatz parametrized at the fermionic level [3]. Depending on which molecular orbitals are considered for the single and double excitations, the UCCSD ansatz can take multiple forms with greatly varying circuit depth. To ensure a fair comparison

TABLE II. Maximum fidelities obtained with d entangling layers, n_θ variational parameters and different initial states for the QOCA, HEA, and UCCSD variational forms applied to the water molecule. The total gate depth reported assumes an all-to-all connectivity, and the same compiling procedure is used for all ansatzes. The initial states are either the Hartree-Fock (HF) approximation to the ground state or the equal superposition of all basis states $|+\rangle^{\otimes N}$. Depending on which molecular orbitals are considered, two instances of UCCSD are presented: the full ansatz (full), and one with a reduced number of excitations (red.); see Appendix E for more UCCSD results.

Water molecule	(12 qubits)				
	Initial state	Maximum fidelity	d	n_θ	Gate depth
QOCA	$|+\rangle^{\otimes N}$	0.9742	1	23	71
	$|+\rangle^{\otimes N}$	0.9955	7	161	497
	$|+\rangle^{\otimes N}$	0.9969	10	230	710
	HF	0.9735	1	23	71
	HF	0.9917	7	161	497
HEA	$|+\rangle^{\otimes N}$	0.9820	8	192	96
UCCSD (red.)	HF	0.9813	1	8	477
UCCSD (full)	HF	0.9999	1	92	6878

with QOCA, we therefore highlight the maximum fidelity acquired with the full UCCSD ansatz (all orbitals considered) and one with a reduced number of excitations such that the circuit depth matches that of QOCA (see also Appendix E for all possible constructions of UCCSD). Because the Hamiltonian describing the water molecule has significantly more terms than the FHM, directly implementing Hamiltonian-based ansatzes as it is done above would lead to very long circuits. Therefore, we do not consider VHA for this problem.

In consequence, as a simple implementation of QOCA for a quantum chemistry problem, we use a variation of the ansatz based on the 12-qubit Hamiltonian of an open 1×6 Fermi-Hubbard chain with the drive terms of Eqs. (13) and (14). Although the water molecule Hamiltonian describes a richer set of fermionic interactions than the FHM, this choice of ansatz offers one of the simplest construction that captures electron-electron correlations and is therefore a good starting point. Note that further improvement to the method could be realized by considering an ansatz that better captures the chemistry of the water molecule and by using other drive terms. Moreover, the ansatz is fully parametrized as before and the simulations are performed under the same numerical conditions. Again, we use the energy of the variational state as the cost function of the minimization procedure, but we report the final state fidelity with respect to the target ground state as a mean of comparing the ansatzes.

The maximum fidelities achieved for QOCA, HEA, and UCCSD are reported in Table II for different numbers of ansatz layers d and initial states, which are either the Hartree-Fock (HF) approximation to the ground state or the equal superposition of all basis states $|+\rangle^{\otimes N}$. Both initial states require one layer of single-qubit gates to prepare. The number of variational parameters n_θ and total gate depth are also indicated. For UCCSD (both full and reduced instances), we use $d = 1$, as it is known to be enough for the simulation of chemical systems [25,73].
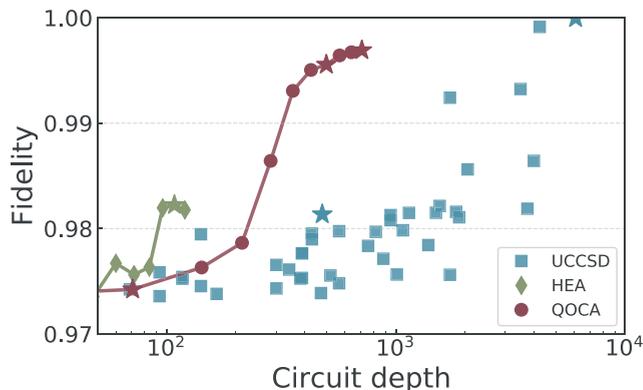
FIG. 6. Final VQA state fidelity with respect to the ground state of the water molecule in its equilibrium geometry for the UCCSD, HEA, and QOCA variational forms. Each UCCSD data point corresponds to an ansatz built from a unique combination of molecular orbitals, and all possible instances are shown. Data for QOCA and HEA come from simulations using up to $d = 10$ ansatz layers. The starred data are also reported in Table II.

From Table II we observe that a $d = 7$ QOCA circuit produces a better approximation to the target ground state (fidelity of 0.9955) than the reduced UCCSD instance (fidelity of 0.9813) while requiring a comparable gate depth. In order to outperform the best QOCA results (fidelity of 0.9969 with $d = 10$), UCCSD requires circuits that are almost an order of magnitude deeper; see Fig. 6 in Appendix E. When considering all molecular orbitals, the full UCCSD ansatz can reach fidelities of > 0.9999, but this comes at the cost of extremely deep circuits with almost 10^4 gates. On the other hand, the HEA approach with $d = 8$ performs similarly than the reduced UCCSD with circuits that are five times shallower. However, HEA requires considerably more parameters, which may compromise its performance for larger systems.

Interestingly, the $|+\rangle^{\otimes N}$ initial state, which has a very small overlap of $\sim 10^{-4}$ with the target state, yields better fidelities for QOCA than the Hartree-Fock initial state, which has a 0.9735 overlap. This, again, emphasizes the fact that QOCA is robust to unstructured initial conditions.

It is important to highlight that while QOCA uses more variational parameters than UCCSD, the circuit depth required to achieve similar state fidelities is significantly smaller, suggesting that QOCA can be useful for solving quantum-chemistry problems in noisy-intermediate-scale quantum devices for which circuit depth is more limiting than the number of parameters. Although the circuit depth is reduced, it is still prohibitively large for application in current devices; nonetheless it constitutes a clear step forward, providing a systematic approach for the construction of a class of quantum circuits that will enable the accurate simulation of complex chemical systems in the near-term quantum computers. Moreover, modifying the QOCA circuit to better reproduce the interactions between the molecular orbitals of the water molecule could lead to further improvements in performance.

VI. CONCLUSION

We introduced the quantum-optimal-control-inspired ansatz by providing a framework to design quantum variational ansatzes that include symmetry-breaking drive terms, similar to the methods of quantum optimal control. We applied QOCA to the half-filled Fermi-Hubbard model and found that, in most cases, the QOCA variational form yields a far faster and more accurate convergence than standard approaches, even when using unstructured initial states having little overlap with the target ground state. We showed evidence that this improved convergence is a consequence of the symmetry-breaking terms of the ansatz which allow for small excursions of the trial wave function outside of the target symmetry sector. Moreover, with minimal modifications of the variational form over our FHM implementation, we used QOCA to prepare the ground state of the water molecule and showed that it can perform equivalently well to the commonly used UCCSD ansatz with shorter circuits.

Its broader applicability and the flexibility in choosing drive terms make QOCA a promising approach to tackle a wide range of quantum chemistry and materials problems on near-term quantum computers. Our work represents a first step towards the development of a more general class of symmetry-breaking ansatzes for variational quantum algorithms.

Note added. Recently we became aware of related work [74].

ACKNOWLEDGMENTS

APPENDIX A: JORDAN-WIGNER FERMIONIC ENCODING

In the Jordan-Wigner transformation, each fermionic site is encoded into the state of two qubits with the mapping $(0, \uparrow, \downarrow, \uparrow\downarrow) \mapsto (00, 01, 10, 11)$. Moreover, the fermionic ladder operators take the form

$$
\hat{a}_p \mapsto \hat{\sigma}_p \bigotimes_{l<p} \hat{Z}_l,
$$
$$
\hat{a}_p^\dagger \mapsto \hat{\sigma}_p^\dagger \bigotimes_{l<p} \hat{Z}_l, \qquad (A1)
$$

where $\hat{\sigma} = |0\rangle\langle 1|$, $\hat{Z}$ is the Pauli-Z operator and the indices denote the spin orbitals or qubits. For a lattice of L sites, we arrange the $N = 2L$ spin orbitals as $|f_{1\uparrow} \dots f_{L\uparrow}; f_{1\downarrow} \dots f_{L\downarrow}\rangle$ with $f_p \in \{0, 1\}$ the occupation of spin orbital p.

TABLE III. Asymptotic scaling of the number of variational parameters of the ansatzes of this work for the full and scalable parametrization strategies. These numbers are for periodic η-dimensional Fermi-Hubbard systems of L lattice sites. d is the number of layers of the ansatzes.

	Full parametrization	Scalable parametrization
HEA	$2Ld$	–
VHA	$(\eta + 1)Ld$	$(2\eta + 1)d$
FT-VHA	$(\eta + 1)Ld$	$(\eta + 1)d$
QOCA	$(\eta + 3)Ld$	$(2\eta + 3)d$
sQOCA	$3Ld$	$3d$

With this mapping, hopping terms between spin orbitals p and q with $p < q$ transform as

$$
\hat{a}_p^\dagger \hat{a}_q + \hat{a}_q^\dagger \hat{a}_p \mapsto \frac{1}{2}(\hat{X}_p \hat{X}_q + \hat{Y}_p \hat{Y}_q) \bigotimes_{l=p+1}^{q-1} \hat{Z}_l, \qquad (A2)
$$

where $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ are Pauli matrices. The product of $\hat{Z}$ operators, referred to as the JW string, vanishes when $q = p + 1$. Moreover, the number operator on spin orbital p, and therefore the onsite Coulomb interaction between spin orbitals p and q take the form

$$
\hat{n}_p = \hat{a}_p^\dagger \hat{a}_p \mapsto \frac{1}{2}(\hat{I} - \hat{Z}_p),
$$
$$
\hat{n}_p \hat{n}_q \mapsto \frac{1}{4}(\hat{I} - \hat{Z}_p - \hat{Z}_q + \hat{Z}_p \hat{Z}_q). \qquad (A3)
$$

At half-filling, the single $\hat{Z}$s coming from the onsite interaction terms are canceled by similar terms arising from the chemical potential, leading to a simple expression for the potential

$$
\hat{V} \mapsto \frac{U}{4} \sum_{i=1}^{L} \hat{Z}_{i\uparrow} \hat{Z}_{i\downarrow}, \qquad (A4)
$$

which is diagonal in the computational basis.

APPENDIX B: PARAMETRIZATION OF THE ANSATZES

1. Full parametrization

This strategy corresponds to taking all (or almost all) gate angles as variational parameters. This gives the classical optimizer enough freedom to explore the Hilbert space spanned by the ansatz at the cost of a longer optimization time. We note that the HEA has, by default, a *fully parametrized* configuration since all single-qubit gates are parametrized. Moreover, the same strategy for VHA consists of assigning one parameter to every $\hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma}$ + H.c. hopping terms and duplicating the parameter to take into account the two spin orientations. This is because at half-filling and zero total spin, there is a spin-inversion symmetry which removes the need to treat spins up and down differently. Additionally, every term of the on-site interaction is associated with a variational parameter. The asymptotic scaling of number of variational parameters for all ansatzes is summarized in Table III for both parametrization strategies.

2. Scalable parametrization

In a scalable parametrization strategy, we employ a number of variational parameters that is independent of the system size. Because there are fewer parameters, we expect the optimization to be faster, but larger circuit depths might be necessary to achieve the same accuracy as full parametrization.

Although it is less clear how one would achieve a scalable parametrization for hardware-efficient approaches, a simple strategy exists for physics-inspired ansatzes such as QOCA. It consists in grouping the individual terms of the Hamiltonian into a constant number of sets containing commuting terms. For example, a common way of grouping the different terms of the FHM on a 2D lattice is

$$\hat{H}_{\text{FHM}} = \hat{H}_{h,\text{even}} + \hat{H}_{h,\text{odd}} + \hat{H}_{v,\text{even}} + \hat{H}_{v,\text{odd}} + \hat{H}_U, \quad \text{(B1)}$$

where the first four terms now group the even and odd, vertical and horizontal hopping terms, while $\hat{H}_U$ collects the on-site interaction terms. Note that for the 3D FHM, two additional sets of hopping terms covering the third dimension would be necessary.

APPENDIX C: NUMERICAL SIMULATIONS

All simulations are done using Qiskit Aqua's VQA tools [64]. Because noise is not considered, a unitary state-vector simulator is used. For simplicity, we also assumed all-to-all connectivity of the qubits, although this is not strictly needed. We chose the COBYLA [61–63] method as the classical optimizer with a maximum number of function evaluation of $\sim 10^5$. This number was justified as being reasonable in [21] using experimentally realistic arguments.

Whenever possible, we initialize all variational parameters to zero. With this choice, Hamiltonian-based ansatzes implement the identity operator at the start of the optimization routine and the variational search begins from the initial state. In contrast to a random initialization of the parameters, this strategy also avoids the need of doing repeated VQA runs and postselecting the best results. However, in the case of short-QOCA, this strategy results in premature convergence of the optimizer into states close to the initial guess, forcing us to use a random initialization of the parameters. Interestingly, even without postselection, this did not hinder the convergence capability due to the robustness of QOCA regarding initial conditions.

Finally, all layers of the ansatzes are optimized simultaneously. Further improvement can potentially be achieved by adopting a layer-by-layer optimization strategy as in Ref. [22].

For the simulation of the water molecule, we use the PySCF driver to obtain the Hamiltonian as provided

APPENDIX D: INITIAL STATES

In most quantum simulations of the FHM reported in the literature [21,22,24,57,58,60], the initial state is the ground state of the noninteracting FHM, i.e., fixing $U = \mu = 0$ in Eq. (11). Because the resulting Hamiltonian is diagonal in Fourier space, this is a convenient choice because the ground state is readily computed classically. However, preparing this

on a quantum computer generally requires very long quantum circuits as it involves the fermionic Fourier transformation. Current implementations of this transformation [30,32,35] are defined only for periodic systems, which limits this initial state's applicability. To the best of our knowledge, no implementation of an open-boundary-conditions fermionic Fourier transformation has been developed to date. Furthermore, the ground state of the noninteracting FHM can be degenerate which makes it difficult to choose which one or superposition thereof to use. This challenge is often pointed out as an open problem [21,60], since in most VQA realization, prior knowledge of the target state is used to find the initial state that maximizes the fidelity. It is unclear how one would make this choice as systems grow computationally intractable.

1. The noninteracting Fermi-Hubbard model

To see how this degeneracy arises, we consider the 1D noninteracting FHM ($U = \mu = 0$) with L sites and periodic boundary conditions. In momentum space, the Hamiltonian is given by a collection of free fermionic modes

$$\hat{\mathcal{T}} = \text{FT}\,\hat{T}\,\text{FT}^\dagger = \sum_{k,\sigma=\{\uparrow,\downarrow\}} \varepsilon_k \hat{c}_{k\sigma}^\dagger \hat{c}_{k\sigma}, \quad \text{(D1)}$$

where the energy spectrum is

$$\varepsilon_k = -2t \cos\left(\frac{2\pi k}{L}\right). \quad \text{(D2)}$$

In the above Hamiltonian, $\hat{c}_{k\sigma}^\dagger$ and $\hat{c}_{k\sigma}$ are respectively the creation and annihilation fermionic operators of momentum k and spin σ. They are obtained from the real-space ladder operators $\hat{a}_{k\sigma}^\dagger$ and $\hat{a}_{k\sigma}$ and the fermionic Fourier transformation as

$$\hat{c}_{k\sigma}^\dagger = \text{FT}\,\hat{a}_{k\sigma}^\dagger\,\text{FT}^\dagger = \frac{1}{\sqrt{L}} \sum_{j=0}^{L-1} e^{-i\frac{2\pi k}{L}j} \hat{a}_{j\sigma}^\dagger, \quad \text{(D3)}$$

$$\hat{c}_{k\sigma} = \text{FT}\,\hat{a}_{k\sigma}\,\text{FT}^\dagger = \frac{1}{\sqrt{L}} \sum_{j=0}^{L-1} e^{i\frac{2\pi k}{L}j} \hat{a}_{j\sigma}. \quad \text{(D4)}$$

Because k can take only discrete values, one notices that a degeneracy appears when there are energy levels at $\varepsilon_k = 0$ since these levels could be occupied or empty without affecting the ground state energy. It is straightforward to see from Eq. (D2) that this can happen only when $L = 4l$, with l an integer. In this case, there are two values of k (corresponding to $k = l$ and $k = 3l$) which leads to $\varepsilon_k = 0$. The degeneracy is therefore $4^2 = 16$ since each momentum mode can be empty, occupied by a $\uparrow$ or $\downarrow$ spin, or both. In the half-filled symmetry sector, the degeneracy is reduced to $\binom{4}{2} = 6$. Note that in the case $L \neq 4l$, the ground state of the noninteracting FHM is not degenerate and is a simple basis state in momentum space.

As mentioned above, this occasional degeneracy makes it difficult to guess which basis state (or superposition thereof) is the best initial state to use in a VQA. However, one can select states that respect certain desired properties such as particle number, total spin, and total momentum.

Typically, the degeneracy at $L = 4l$ can be lifted by applying a small perturbative Coulomb interaction U. In this case, the ground state of the noninteracting FHM becomes a superposition of basis states in Fourier space. One must apply the $FT^\dagger$ in order to transform this initial state into real space for the VQA.

2. Choosing and preparing the initial states

In the case of $L = 4$ (or 2×2), we computed the fidelity of the 16 degenerate ground states of Eq. (D1) with respect to the target ground state and postselected the ones leading to the highest fidelity. This strategy is, of course, not scalable and therefore it remains unclear how one would proceed in practice in the case where the fidelity with the target ground state cannot be computed beforehand.

In the present case, this strategy yields two ground states with a fidelity of ≈ 0.425 with respect to the ground state of the full model. Labeling the spin orbitals $|f_{1\uparrow} \dots f_{L\uparrow}; f_{1\downarrow} \dots f_{L\downarrow}\rangle$, these two states in real space are

$$|\Omega_T^{(1)}\rangle = FT^\dagger |1100 ; 1100\rangle, \tag{D5}$$

$$|\Omega_T^{(2)}\rangle = FT^\dagger |1001 ; 1001\rangle. \tag{D6}$$

Preparing these two states requires applying Pauli-X gates on selected qubits followed by the fermionic Fourier transformation, something which requires long quantum circuits [30,32,35].

Adding a small perturbation $U = 1 \times 10^{-5} t$, we find that the following superposition of $|\Omega_T^{(1)}\rangle$ and $|\Omega_T^{(2)}\rangle$ yields a significantly larger fidelity to the true ground state of ≈ 0.85:

$$\begin{aligned}|\Omega_T\rangle &= \frac{|\Omega_T^{(1)}\rangle - |\Omega_T^{(2)}\rangle}{\sqrt{2}} \\ &= FT^\dagger \frac{|1100 ; 1100\rangle - |1001 ; 1001\rangle}{\sqrt{2}}. \end{aligned} \tag{D7}$$

This, however, increases the complexity of the initial state preparation.

APPENDIX E: SIMULATION OF THE WATER MOLECULE

In order to vary the circuit depth of the UCCSD ansatz and provide a fair comparison against the QOCA variational form, we consider all possible combinations of excitations for the 12-qubit water molecule. The varied number of molecular orbitals involved in the excitations yields circuits with different gate depths, according to the construction rules of the UCCSD ansatz [18,25]. Furthermore, this choice of excitations participating in the generation of the ansatz might affect the quality of the final VQA state. In Fig. 6 we compare the performance of the UCCSD, HEA, and QOCA variational forms in preparing the ground state of the water molecule in its equilibrium geometry.

Each data point in Fig. 6 obtained for the UCCSD ansatz represents a unique combination of spin orbitals considered in the cluster operators, starting from one single excitation and going up to all possible single and double excitations. We observe that the number of excitations considered play a critical role in the final state fidelity, which can reach 0.9999 when the full UCCSD ansatz is taken into account. Importantly, the circuit depth required for this approach is significantly larger than the one of QOCA for the same accuracy. In contrast, the number of variational parameters is smaller for UCCSD.

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

[2] K. Wright, K. M. Beck, S. Debnath, J. M. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. C. Pisenti, M. Chmielewski, C. Collins *et al.*, Benchmarking an 11-qubit quantum computer, Nat. Commun. **10**, 5464 (2019).

[3] 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).

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

[5] C. Zoufal, A. Lucchi, and S. Woerner, Quantum generative adversarial networks for learning and loading random distributions, npj Quantum Inf. **5**, 103 (2019).

[6] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles, Variational quantum linear solver: A hybrid algorithm for linear systems, arXiv:1909.05820.

[7] M. Schuld and N. Killoran, Quantum Machine Learning in Feature Hilbert Spaces, Phys. Rev. Lett. **122**, 040504 (2019).

[8] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, Nature (London) **567**, 209 (2019).

[9] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[10] P. K. Barkoutsos, G. Nannicini, A. Robert, I. Tavernelli, and S. Woerner, Improving variational quantum optimization using CVaR, Quantum **4**, 256 (2020).

[11] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, Elucidating reaction mechanisms on quantum computers, Proc. Natl. Acad. Sci. USA **114**, 7555 (2017).

[12] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, Solving strongly correlated electron models on a quantum computer, Phys. Rev. A **92**, 062318 (2015).

[13] J. Olson, Y. Cao, J. Romero, P. Johnson, P.-L. Dallaire-Demers, N. Sawaya, P. Narang, I. Kivlichan, M. Wasielewski, and A. Aspuru-Guzik, Quantum information and computation for chemistry, arXiv:1706.05413.

[14] 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).

[15] A. Robert, P. K. Barkoutsos, S. Woerner, and I. Tavernelli, Resource-efficient quantum algorithm for protein folding, npj Quantum Inf. **7**, 38 (2021).

[16] A. Di Paolo, P. K. Barkoutsos, I. Tavernelli, and A. Blais, Variational quantum simulation of ultrastrong light-matter coupling, Phys. Rev. Res. **2**, 033364 (2020).

[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] 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).

[19] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. Rol, C. Bultink, X. Fu, C. Price, V. Ostroukh, N. Muthusubramanian, A. Bruno *et al.*, Error mitigation by symmetry verification on a variational quantum eigensolver, Phys. Rev. A **100**, 010302 (2019).

[20] M. Ganzhorn, D. J. Egger, P. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, M. Roth, A. Fuhrer, P. Mueller, S. Woerner *et al.*, Gate-Efficient Simulation of Molecular Eigenstates on a Quantum Computer, Phys. Rev. Appl. **11**, 044092 (2019).

[21] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, Strategies for solving the Fermi-Hubbard model on near-term quantum computers, Phys. Rev. B **102**, 235122 (2020).

[22] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).

[23] 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).

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

[25] 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).

[26] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, Phys. Rev. Lett. **120**, 110501 (2018).

[27] A. J. Woitzik, P. K. Barkoutsos, F. Wudarski, A. Buchleitner, and I. Tavernelli, Entanglement production and convergence properties of the variational quantum eigensolver, Phys. Rev. A **102**, 042402 (2020).

[28] 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).

[29] 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).

[30] F. Verstraete, J. I. Cirac, and J. I. Latorre, Quantum circuits for strongly correlated quantum systems, Phys. Rev. A **79**, 032316 (2009).

[31] A. J. Ferris, Fourier Transform for Fermionic Systems and the Spectral Tensor Network, Phys. Rev. Lett. **113**, 010401 (2014).

[32] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions, Phys. Rev. Appl. **9**, 044036 (2018).

[33] J. A. Fleck, J. Morris, and M. Feit, Time-dependent propagation of high energy laser beams through the atmosphere, Appl. Phys. **10**, 129 (1976).

[34] M. Feit, J. Fleck Jr, and A. Steiger, Solution of the Schrödinger equation by a spectral method, J. Comput. Phys. **47**, 412 (1982).

[35] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, Low-Depth Quantum Simulation of Materials, Phys. Rev. X **8**, 011044 (2018).

[36] D. Dong and I. R. Petersen, Quantum control theory and applications: A survey, IET Control Theor. Appl. **4**, 2651 (2010).

[37] M. James, Quantum control theory, arXiv:1406.5260.

[38] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, J. Magn. Reson. **172**, 296 (2005).

[39] F. Motzoi, J. M. Gambetta, S. Merkel, and F. K. Wilhelm, Optimal control methods for rapidly time-varying Hamiltonians, Phys. Rev. A **84**, 022307 (2011).

[40] A. B. Magann, C. Arenz, M. D. Grace, T.-S. Ho, R. L. Kosut, J. R. McClean, H. A. Rabitz, and M. Sarovar, From pulses to circuits and back again: A quantum optimal control perspective on variational quantum algorithms, PRX Quantum **2**, 010101 (2021).

[41] D. d'Alessandro, *Introduction to Quantum Control and Dynamics* (CRC Press, Boca Raton, FL, 2007).

[42] S. A. Rice, Interfering for the good of a chemical reaction, Nature (London) **409**, 422 (2001).

[43] M. Shapiro and P. Brumer, Quantum control of bound and continuum state dynamics, Phys. Rep. **425**, 195 (2006).

[44] N. Khaneja, T. Reiss, B. Luy, and S. J. Glaser, Optimal control of spin dynamics in the presence of relaxation, J. Magn. Reson. **162**, 311 (2003).

[45] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Implementing a universal gate set on a logical qubit encoded in an oscillator, Nat. Commun. **8**, 94 (2017).

[46] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, Adapt-vqe: An exact variational algorithm for fermionic simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).

[47] H. L. Tang, E. Barnes, H. R. Grimsley, N. J. Mayhall, and S. E. Economou, qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor, arXiv:1911.10205.

[48] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Quantum algorithms for fermionic simulations, Phys. Rev. A **64**, 022319 (2001).

[49] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Erratum: Quantum algorithms for fermionic simulations [Phys. Rev. A **64**, 022319 (2001)], Phys. Rev. A **65**, 029902(E) (2002).

[50] J. Hubbard, Electron correlations in narrow energy bands, Proc. R. Soc. London A **276**, 238 (1963).

[51] S. Yamada, T. Imamura, and M. Machida, 16.447 tflops and 159-billion-dimensional exact-diagonalization for trapped fermion-Hubbard model on the earth simulator, in *SC'05: Proceedings of the 2005 ACM/IEEE Conference on Supercomputing* (IEEE, Seattle, WA, USA, 2005), pp. 44–44.

[52] K. Masuda and D. Yamamoto, Variational cluster approach to *s*-wave pairing in heavy-fermion superconductors, Phys. Rev. B **91**, 104508 (2015).

[53] M. Guillot, Compétition entre l'antiferromagnétisme et la supraconductivité dans le modèle et Hubbard appliqué aux cuprates, M.Sc. thesis, Université de Sherbrooke, 2007.

[54] J. Kaczmarczyk, J. Spałek, T. Schickling, and J. Bünemann, Superconductivity in the two-dimensional Hubbard model: Gutzwiller wave function solution, Phys. Rev. B **88**, 115127 (2013).

[55] P.-L. Dallaire-Demers, M. Stechly, J. F. Gonthier, N. T. Bashige, J. Romero, and Y. Cao, An application benchmark for fermionic quantum simulations, arXiv:2003.01862.

[56] M. Wilson, R. Stromswold, F. Wudarski, S. Hadfield, N. M. Tubman, and E. G. Rieffel, Optimizing quantum heuristics with meta-learning, Quantum Machine Intelligence, **3**, 13 (2021).

[57] J.-M. Reiner, F. Wilhelm-Mauch, G. Schön, and M. Marthaler, Finding the ground state of the Hubbard model by variational methods on a quantum computer with gate errors, Quantum Sci. Technol. **4**, 035005 (2019).

[58] G. Verdon, M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni, Learning to learn with quantum neural networks via classical neural networks, arXiv:1907.05415.

[59] I. O. Sokolov, P. K. Barkoutsos, P. J. Ollitrault, D. Greenberg, J. Rice, M. Pistoia, and I. Tavernelli, Quantum orbital-optimized unitary coupled cluster methods in the strongly correlated regime: Can quantum algorithms outperform their classical equivalents? J. Chem. Phys. **152**, 124107 (2020).

[60] L. Xu, J. T. Lee, and J. Freericks, Test of the unitary coupled-cluster variational quantum eigensolver for a simple strongly correlated condensed-matter system, Modern Phys. Lett. B **34**, 2040049 (2020).

[61] M. J. Powell, A direct search optimization method that models the objective and constraint functions by linear interpolation, in *Advances in Optimization and Numerical Analysis* edited by S. Gomez and J.-P. Hennart (Springer, 1994), pp. 51–67.

[62] M. J. Powell, Direct search algorithms for optimization calculations, Acta Numer. **7**, 287 (1998).

[63] M. J. Powell, A view of algorithms for optimization without derivatives, Math. Today-Bull. Inst. Math. Appl. **43**, 170 (2007).

[64] H. Abraham *et al.*, Qiskit: An open-source framework for quantum computing (2019), https://github.com/Qiskit.

[65] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, Algorithms **12**, 34 (2019).

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

[67] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).

[68] K. Temme, S. Bravyi, and J. M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[69] S. Endo, S. C. Benjamin, and Y. Li, Practical Quantum Error Mitigation for Near-Future Applications, Phys. Rev. X **8**, 031027 (2018).

[70] A. Kandala, K. Temme, A. D. Corcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Extending the computational reach of a noisy superconducting quantum processor, Nature (London) **567**, 491 (2019).

[71] 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).

[72] J. B. Collins, P. von R. Schleyer, J. S. Binkley, and J. A. Pople, Self-consistent molecular orbital methods. XVII. Geometries and binding energies of second-row molecules. A comparison of three basis sets, J. Chem. Phys. **64**, 5142 (1976).

[73] 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 *et al.*, Scalable Quantum Simulation of Molecular Energies, Phys. Rev. X **6**, 031007 (2016).

[74] N. Vogt, S. Zanker, J.-M. Reiner, T. Eckl, A. Marusczyk, and M. Marthaler, Preparing symmetry broken ground states with variational quantum algorithms, arXiv:2007.01582.