

Meta-Variational Quantum Eigensolver: Learning Energy Profiles of Parameterized Hamiltonians for Quantum Simulation


Alba Cervera-Lierta^{1,2,*}, Jakob S. Kottmann^{1,2} and Alán Aspuru-Guzik^{1,2,3,4}

¹*Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Canada*

²*Department of Computer Science, University of Toronto, Canada*

³*Vector Institute for Artificial Intelligence, Toronto, Canada*

⁴*Canadian Institute for Advanced Research (CIFAR), Toronto, Canada*

 (Received 18 October 2020; revised 18 March 2021; accepted 30 April 2021; published 28 May 2021)

We present the meta-variational quantum eigensolver (VQE), an algorithm capable of learning the ground-state energy profile of a parameterized Hamiltonian. If the meta-VQE is trained with a few data points, it delivers an initial circuit parameterization that can be used to compute the ground-state energy of any parameterization of the Hamiltonian within a certain trust region. We test this algorithm with an XXZ spin chain, an electronic H₄ Hamiltonian, and a single-transmon quantum simulation. In all cases, the meta-VQE is able to learn the shape of the energy functional and, in some cases, it results in improved accuracy in comparison with individual VQE optimization. The meta-VQE algorithm introduces both a gain in efficiency for parameterized Hamiltonians in terms of the number of optimizations and a good starting point for the quantum circuit parameters for individual optimizations. The proposed algorithm can be readily mixed with other improvements in the field of variational algorithms to shorten the distance between the current state of the art and applications with quantum advantage.

DOI: [10.1103/PRXQuantum.2.020329](https://doi.org/10.1103/PRXQuantum.2.020329)

I. INTRODUCTION

Variational quantum algorithms (VQAs) are one of the key tools for the era of noisy intermediate-scale quantum (NISQ) computation [1,2] and beyond, due to their natural method of optimization. Their hybrid quantum-classical structure exploits the current advantages of both worlds: a quantum circuit is used to compute the expected values of some observable, whereas a classical subroutine finds the optimal parameterization of this quantum circuit. The continuous parameterization of the quantum gates allows us to adjust their arguments to partially compensate for the effect of noisy qubits and imperfect operations. This approach has opened up the possibility of finding applications of quantum computing in the near term, without the need for quantum error correction.

A VQA can be divided into three principal blocks: the preparation of the initial state, the preparation of a parameterized quantum circuit, and the measurement and

construction of the cost function. The first step is crucial for starting the algorithm in the Hilbert-space region where the solution is likely to be. The second block guides the algorithm around a particular region of the space of quantum states, and therefore a circuit with high expressibility [3] or a physically inspired ansatz is required. The last block computes the expected value of some operator with the final state of the circuit and constructs a cost function that is then minimized with a classical subroutine. The classical minimizer proposes a new set of values for the quantum circuit block, repeating the loop until convergence or the desired precision is achieved.

The first VQA proposed was the variational quantum eigensolver (VQE) [4,5]. This algorithm, originally proposed for molecular systems, tries to find the ground-state energy of a given Hamiltonian by variationally minimizing its expectation value with a parameterized quantum circuit. The cost function of this algorithm is the expected value of the model Hamiltonian. The variational principle states that this value is an upper bound on the ground-state energy, and so everything reduces to minimizing this value by fine-tuning the parameters of the circuit. There are different methods to design both the initial state and the quantum circuit ansatz when the model is a molecule [6–9], an integrable condensed-matter model [10–15], or a more general computational problem [16]. However, there is not a general approach for other kinds of Hamiltonian.

* a.cervera.lierta@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Even in quantum chemistry, where good heuristics for circuit construction and initialization are known, layerwise extension of the circuits can lead to similar initialization problems [17].

There are only a few techniques that try to find quantum circuits for general Hamiltonians. Some proposals focus on finding parameterized circuits that can more efficiently explore a region of the Hilbert space [3,18]. In general, all these circuit ansatz consist of smaller circuit layers that share a common structure, in such a way that these layers are concatenated until the desired precision is achieved. Although this structure may seem efficient, the exponentially large Hilbert space hampers the ability to explore it in a reasonable number of steps. Only if there exists some initial-state guess can the algorithm converge to the correct solution. The lack of this initial estimation sometimes imposes random initialization, which leads to the barren-plateau problem: both the gradient and the variance of the circuit parameters tend to zero exponentially, causing the algorithm to get stuck in some local minimum [19].

In the best-case scenario, with an initial-state preparation and a physically inspired circuit, a VQE run will give only a specific ground-state energy. This is in many cases insufficient when the Hamiltonian depends on some parameters (nuclear coordinates, external field strengths, model-specific parameters) and the goal is to find the configuration that leads to specific properties of the system. Some examples are the lowest ground-state energy with respect to the parameters of the Hamiltonian (in molecular cases this is often referred to as *geometry optimization*), energy gaps between ground and excited states, the convergence behavior of specific parameterized circuits with respect to the parameters of model systems, and, in general, the evolution of some observable with respect to an external parameter such as the strength of an electromagnetic field. Then, one has to run many instances of the VQE to scan over these parameters, inevitably increasing the computational cost. Previous work has explored the possibility of using a VQA to predict the ground state of a Hamiltonian. In particular, Ref. [20] proposed to use adiabatic state preparation to design a circuit ansatz capable of doing that task. However, that proves costly when more than one Trotter step is considered. We aim to generalize this idea further by using short-depth quantum circuits, and analyze different Hamiltonians and encoding strategies.

In this work, we address the general problems stated above at once by proposing the *meta-VQE* algorithm. A meta-VQE encodes the Hamiltonian parameters into the first layers of the quantum circuit, dividing the circuit into two parts: encoding and processing. Next, the meta-VQE is trained with a small set of Hamiltonian parameters by constructing a cost function that is a combination of all expected values. Finally, the meta-VQE has “learned” the Hamiltonian, and we can simply introduce other values of the Hamiltonian parameters into the circuit and obtain a

good estimation of the ground-state energy. If this estimation is not precise enough, we can use the resulting circuit of the meta-VQE as a starting point for a standard VQE, providing a good initial guess and avoiding the random-initialization problem, i.e., barren plateaus [19].

This algorithm is inspired by quantum machine-learning (QML) algorithms and other algorithms that use metatechniques [21–28]. As these algorithms propose, we design a parameterized quantum circuit to be trained with a set of values from a given model, in our case, a physical Hamiltonian. We treat the encoding part of the meta-VQE as a quantum neural network that learns the encoding of the Hamiltonian. The processing part guides the encoded state towards the ground state. We observe two advantages of this algorithm: (i) it can be used to first explore the ground-state energies of the Hamiltonian parameter space with only a few training points and then use the result as an initial state for a precise VQE, and (ii) the encoding in a VQA proves valuable and helps these algorithms to find the ground state. The meta-VQE can be interpreted as a QML application for quantum simulation suited for the NISQ era.

We present the results of this work in the following section. First, we introduce the meta-VQE algorithm from a general perspective. Next, we run a meta-VQE with a spin-Hamiltonian example, the one-dimensional (1D) XXZ model, and then with a molecular-Hamiltonian example, the H_4 complex, consisting of two H_2 molecules with a fixed bond distance of 1.23 Å in a rectangular arrangement with a varying intermolecular distance d (the same system as used in Ref. [17]). We compare the performance of the meta-VQE with that of a standard VQE with random initialization and a standard VQE initialized with the trained parameters obtained from the meta-VQE (*opt-meta-VQE*). Finally, we apply this algorithm to a state-of-the-art application: the simulation of transmon qubits in a quantum computer [29]. We discuss the results and propose further improvements of this algorithm in the Discussion section. Results are obtained via implementation of the meta-VQE protocol using TEQUILA [30] and choosing QULACS [31] as a quantum backend. TEQUILA represents quantum objectives as generalized functions of abstract expectation values, allowing arbitrary transformations and arithmetic operations on those data types to be done in a blackboard fashion. This allows straightforward construction of meta-VQEs.

II. RESULTS

A. Meta-VQE algorithm

The structure of the meta-VQE algorithm is shown diagrammatically in Fig. 1. Given an n -qubit parameterized Hamiltonian of the form $H = H(\lambda)$, where $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_q\}$ are q different parameters, we select M sets

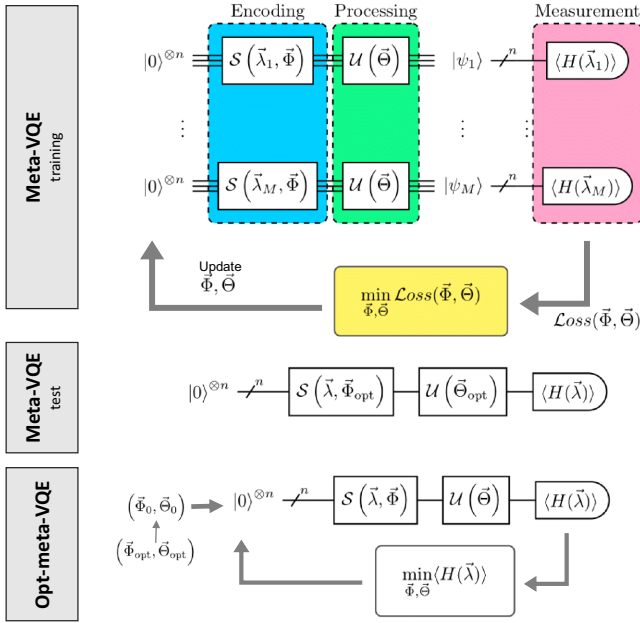


FIG. 1. Schematic illustration of the meta-VQE algorithm. The top diagram represents the training part of the algorithm, where the Hamiltonian parameters $\vec{\lambda}$ are encoded together with the variational parameters $\vec{\Phi}$ and $\vec{\Theta}$. By computing the expected value of the Hamiltonian for multiple values of λ , we design a cost function \mathcal{L}_{oss} to be minimized. Once the algorithm converges to $\vec{\Phi}_{\text{opt}}$ and $\vec{\Theta}_{\text{opt}}$, we can use these values to obtain any $\langle H(\vec{\lambda}) \rangle$ (meta-VQE test) or as initial values of a standard VQE algorithm (opt-meta-VQE).

of $\vec{\lambda}$ that we will use as a training set. The circuit is initialized in the $|0\rangle^{\otimes n}$ state, or in the Hartree-Fock state for molecules (see next section). The first part of the circuit is the encoding unitary \mathcal{S} containing parameterized gates, the arguments of which include the Hamiltonian parameters of one of the training points $\vec{\lambda}_i$ and the variational parameters $\vec{\Phi}$. The second unitary of the circuit, \mathcal{U} , also contains parameterized gates that depend on variables $\vec{\Theta}$, but not the Hamiltonian parameters. The final state of the circuit can be written as

$$|\psi_i\rangle \equiv |\psi(\vec{\lambda}_i, \vec{\Phi}, \vec{\Theta})\rangle = \mathcal{U}(\vec{\Theta})\mathcal{S}(\vec{\lambda}_i, \vec{\Phi})|0\rangle^{\otimes n}. \quad (1)$$

The meta-VQE is optimized over a set of training points λ_i by minimizing a cost function that depends on all expected values of the Hamiltonians $H(\lambda_i)$. Here we employ a simple cost function that is the sum of all expected values,

$$\mathcal{L}_{\text{oss}}(\vec{\Phi}, \vec{\Theta}) = \sum_{i=1}^M \langle \psi_i | H(\lambda_i) | \psi_i \rangle, \quad (2)$$

but other more sophisticated or problem-dependent ones can be developed and conveniently implemented within TEQUILA. Finally, we minimize this loss function with a

gradient-based method, obtaining the optimal values for the variational parameters, $\vec{\Phi}_{\text{opt}}$ and $\vec{\Theta}_{\text{opt}}$.

Once the meta-VQE circuit is trained, we can proceed to test its performance. We compute the expected value of the Hamiltonian with other values of $\vec{\lambda}$ by running the meta-VQE circuit with the trained parameters $\vec{\Phi}_{\text{opt}}$ and $\vec{\Theta}_{\text{opt}}$ and the corresponding $\vec{\lambda}$ values for the test point. The results show that a meta-VQE is capable of learning the profile of the ground-state energy as a function of the parameters $\vec{\lambda}$, but its accuracy depends strongly on the ansatz for the encoding and processing circuit. For that reason, we also propose to use the result from the meta-VQE as an initial-state guess for a standard VQE, giving what we call *opt-meta-VQE*. In the following subsections, we present particular examples of meta-VQE circuits and test them with different kinds of physical Hamiltonian.

B. Spin Hamiltonian: 1D XXZ model

Let us start with the 1D antiferromagnetic XXZ spin Hamiltonian with an external magnetic field,

$$\mathcal{H} = \sum_{i=1}^n \sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z + \lambda \sum_{i=1}^n \sigma_i^z, \quad (3)$$

where Δ is the anisotropy parameter and λ the transverse field strength. We also consider periodic boundary conditions. For $\lambda = 0$, this model contains two quantum phase transitions, at $\Delta = \pm 1$. Its ground state is a product state for $\Delta < -1$ and is highly entangled in the critical region, i.e., for $-1 < \Delta \leq 1$. The introduction of the external field λ moves the $\Delta = -1$ phase transition to higher values of Δ , increasing the region with a product state as its ground state [32].

We use this model as a test for various reasons. First, this is a nontrivial model, with a highly entangled ground state, a property that motivates the use of a quantum computer. Second, there is no known quantum circuit capable of computing the ground state exactly, in contrast with other integrable models [10–15], although there has been recent work on the preparation of Bethe-ansatz eigenstates using projective algorithms [33]. Third, the energy profile as a function of Δ for a nonzero value of λ has a nontrivial shape with a peak (see Fig. 2). Finally, some condensed-matter models, such as the Haldane-Shastry model [34] and some electronic Hamiltonians, are related to this model.

As stated above, no known quantum circuit represents the ground state of this model for any n and Δ . This gives us a good training ground to test the most general approach to a meta-VQE. We follow the same strategy as in other VQAs when applied to a general problem: we construct the quantum circuit with minimal blocks, called layers, that are concatenated and progressively introduce more variational parameters and generate entanglement. One may

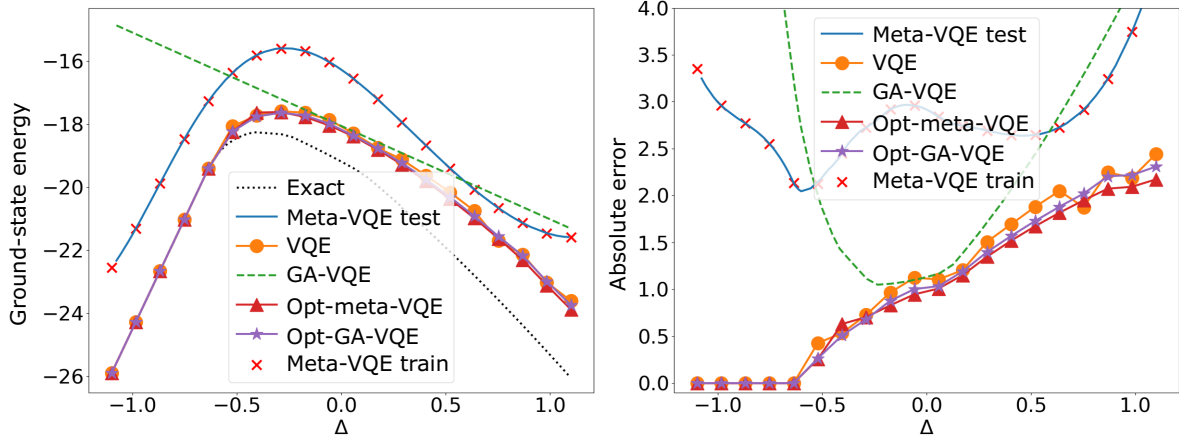


FIG. 2. Results for the ground-state energy (arbitrary units) in relation to the parameter Δ for an $n = 14$ XXZ spin chain with a transverse field $\lambda = 0.75$. The total number of layers considered is four, divided into encoding and processing for the meta-VQE and opt-meta-VQE. The meta-VQE learns the energy profile, whereas the standard VQE achieves better precision. However, using the results of the meta-VQE as a starting point for a VQE (opt-meta-VQE and opt-GA-VQE) improves the performance notably and avoids local minima, as shown by the plot of the absolute error with respect to the exact solution.

expect that the greater the number of layers that are considered, the closer to the ground state the algorithm will end up. The encoding part of the meta-VQE follows a similar strategy to data re-uploading for variational quantum classification [26]: we encode the parameter Δ into rotational single-qubit gates using a linear function of the form $w\Delta + \phi$, where w and ϕ are the variational parameters. Universality of this kind of encoding, when provided with enough layers with a single qubit, has been shown [26], so we expect to obtain a similar behavior when it is used on multiple qubits and entanglement is added between them. Each layer, in both the encoding and the processing part, contains first-neighbor controlled-NOT (CNOT) gates. This kind of entanglement ansatz has been proven to provide circuits with high expressibility [3].

An encoding layer l of the meta-VQE circuit for this model can be written in the following form:

$$S_l \equiv S(\Delta, \vec{\Phi}_l) = R(f(\Delta, \vec{\varphi}_{1l})) \otimes \cdots \otimes R(f(\Delta, \vec{\varphi}_{nl})) \otimes \text{CNOT} \otimes \cdots \otimes \text{CNOT}, \quad (4)$$

where $\vec{\Phi}_l = (\vec{\varphi}_{1l}, \dots, \vec{\varphi}_{nl})$ and

$$R(f(\Delta, \vec{\varphi}_{il})) = R_z(w_{il}^{(1)}\Delta + \phi_{il}^{(1)})R_y(w_{il}^{(2)}\Delta + \phi_{il}^{(2)}) \quad (5)$$

are single-qubit rotational gates with $\vec{\varphi}_{il} = (\vec{w}_{il}, \vec{\phi}_{il})$, $\vec{w}_{il} = (w_{il}^{(1)}, w_{il}^{(2)})$, and $\vec{\phi}_{il} = (\phi_{il}^{(1)}, \phi_{il}^{(2)})$. Then, each encoding layer contains $4n$ variational parameters. Notice that this linear encoding is similar to the classical neural-network encoding, where \vec{w}_{il} and $\vec{\phi}_{il}$ play the role of the weights and biases, respectively, and the rotational gate plays the role of the nonlinear activation function. The encoding gate is

then constructed with L_1 layers:

$$S(\Delta, \vec{\Phi}) = S_1 \otimes \cdots \otimes S_{L_1}, \quad (6)$$

where $\vec{\Phi} = (\vec{\Phi}_1, \dots, \vec{\Phi}_{L_1})$.

Each processing layer is constructed in the same way as in Eq. (4), now with rotations that do not depend on the metaparameters. Thus, each processing layer can be written as

$$U_l \equiv U(\vec{\Theta}_l) = R(\vec{\theta}_{1l}) \otimes \cdots \otimes R(\vec{\theta}_{nl}) \otimes \text{CNOT} \otimes \cdots \otimes \text{CNOT}, \quad (7)$$

where

$$R(\vec{\theta}_{il}) = R_z(\theta_{il}^{(1)})R_y(\theta_{il}^{(2)}). \quad (8)$$

Each layer of the processing unitary contains $2n$ variational parameters, the components of $\vec{\theta}_{il}$. Considering L_2 processing layers, the total unitary becomes

$$U(\vec{\Theta}) = U_1 \otimes \cdots \otimes U_{L_2}, \quad (9)$$

with $\vec{\Theta} = (\vec{\Theta}_1, \dots, \vec{\Theta}_{L_2})$.

Figure 3 shows the circuit ansatz described above for $n = 4$ qubits and two encoding and processing layers.

We run the meta-VQE for the XXZ model with a transverse field $\lambda = 0.75$. The training set of Δ points is composed of 20 equispaced points between $\Delta = -1.1$ and $\Delta = 1.1$. The performance of the meta-VQE is then tested through evaluation on 100 equispaced testing points between the same Δ values (and using the parameter values found after the training).

To check whether the encoding strategy of the meta-VQE entails an advantage, we run a meta-VQE with

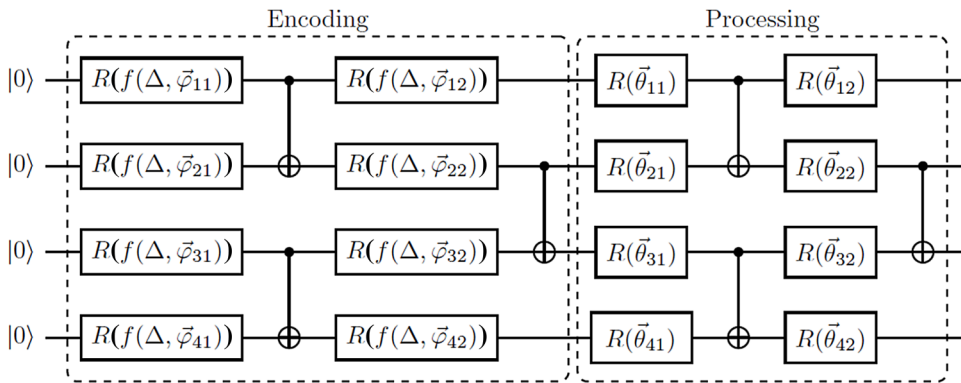


FIG. 3. Circuit ansatz corresponding to four qubits and two encoding and processing layers. This ansatz is used for the XXZ meta-VQE. Each $R(\vec{x})$ gate corresponds to $R_z(x^{(1)})R_y(x^{(2)})$. The function used for the encoding layer is $f(\Delta, \vec{\varphi}) = w\Delta + \phi$.

no encoding layers, i.e., all layers have the form of U_l in Eq. (7) (a globally averaged VQE, *GA-VQE*). The number of optimization parameters is then lower than for the original meta-VQE, so we keep the same total number of layers to maintain the circuit depth.

To check the possible advantage of the meta-VQE learning strategy, we compare the results with those from a standard VQE. Again, we keep the total number of layers at $L_1 + L_2$, and all layers are of the form of U_l in Eq. (7). We do not consider a single-point VQE with an encoding part, because the encoding function $w\Delta + \phi$ can be expressed with a single angle, θ , reducing the circuit to a standard VQE. Encoding has an effect only in the simultaneous optimization of different points. In total, the meta-VQE and opt-meta-VQE need $n(4L_1 + 2L_2)$ variational parameters, and GA-VQE and the standard VQE need $n(2L_1 + 2L_2)$.

Finally, we check whether we can use the results of a meta-VQE and GA-VQE, that is, with and without the encoding strategy, respectively, as starting points for a standard VQE (*opt-meta-VQE* and *opt-GA-VQE*). We run the simulation of a VQE again, using as initial parameters the results of the meta-VQE and of the meta-VQE without encoding (GA-VQE). In both cases, the circuit depth is the same but the optimization parameters are different. The latter initializes the quantum circuit to the same state, regardless of the Hamiltonian parameter, while the former provides a parameter-dependent initialization. In contrast to the standard VQE with encoding, in this case the encoding helps us to start at a specific point: if the initialization is random, the encoding can be reduced to a single rotational parameter, but if it is not, it can be used to help the algorithm to start at a particular point.

Figure 2 shows the results of these four simulations for an XXZ spin chain of 14 qubits, with two encoding and two processing layers. In all cases, the circuit is not good enough to find the ground-state energy for all values of Δ . The reason is simple: the circuit ansatz does not generate enough entanglement, and the rotational gates are not expressible enough. This is something that we should expect, since we are considering just a few layers and, in turn, their design is not physically inspired. As we mentioned before, we chose this model as an example of a

worst-case scenario for a VQE-type simulation, where we have no clue regarding the circuit ansatz and the initial-state preparation. The results are much better for $\Delta \lesssim 0.6$ because the ground state there corresponds to all spins being aligned with the external magnetic field (a ferromagnetic phase), i.e., in the computational basis, the state $|11 \dots 1\rangle$. This basis element is easy to find for the algorithms, due to the full basis superposition induced by the R_y gates.

It seems that a meta-VQE can find the general energy shape but not provide an accurate value, in contrast to the standard VQE. GA-VQE can reproduce only a linear profile, not learning the behavior of the energy function. However, opt-meta-VQE proves valuable, achieving better results than the standard VQE with random initialization. This result is better shown in Fig. 2 right, where we plot the absolute error of all algorithms with respect to the exact solution. Whereas the standard VQE gets stuck in local minima for some values of Δ , opt-meta-VQE is able to deal with those false minima and achieve better precision.

Finally, we compare the scaling of the results for different numbers of qubits and, therefore, an increasing number of optimization parameters. Figure 4 shows the relative error with respect to the exact ground-state energy for the four algorithms considered for $n = 8, 10, 12, 14$. The errors decrease slightly with the number of qubits for opt-meta-VQE and opt-GA-VQE. However, the meta-VQE shows fluctuations depending on the phase of the system and the number of qubits. This behavior may be a consequence of the hardness of the optimization; we are trying to learn the ground-state energy of a critical system that undergoes a quantum phase transition with a single global cost function. The ground state will change drastically when the critical point is crossed, and the encoding may be dependent on the phase. As mentioned before, we chose this model as an example of a hard case.

C. Molecular Hamiltonian: H_4

To explore the role of the encoding function, we employ an example of an electronic structure problem in quantum chemistry. As before, the objective is to

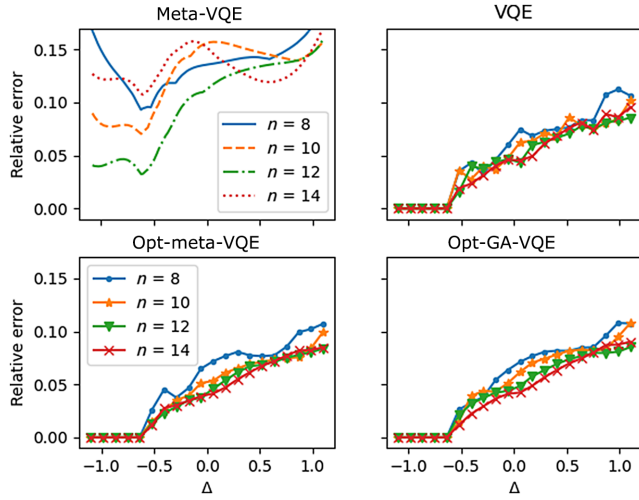


FIG. 4. Relative errors with respect to the exact ground state of the 1D XXZ model with $\lambda = 0.75$ for the four algorithms considered in this work in relation to the Δ parameter and the number of qubits n . The errors for the VQE, opt-meta-VQE, and opt-GA-VQE decrease slightly with the number of qubits. For the meta-VQE, the error scaling with n depends on the phase of the system, a trait that can be explained by the hardness of learning a critical system with a single global cost function.

approximate ground states of qubit Hamiltonians, generated in this example by transforming second-quantized fermionic operators using various transforms [8,35]. We use the k -UpCCGSD approach of Ref. [17] in a separated form, equivalent to a single Trotter step of the full approach, as a VQE ansatz. This denotes a unitary coupled-cluster ansatz with generalized electronic single and double excitations, where the double excitations are limited to electron pairs within the same spatial orbital. This ansatz is repeated k times, similarly to the layered construction used before. When applied to a reference state, unitary coupled-cluster approaches are number conserving, which is a necessary condition for electronic eigenstates. In this work, we use the Hartree-Fock state (the best product state obtained from a classical mean-field optimization) as a reference. In contrast to the spin Hamiltonians above, the circuit construction is physically motivated, illustrating a different use case of the meta-VQE approach. In the next paragraphs, we give a brief description of the construction of the parameterized quantum circuit following the implementations presented in Refs. [30,36].

The overall circuit consists of a collection of primitive unitary coupled-cluster operations that describe collective excitations of n electrons from spin orbitals $\mathbf{p} = (p_0, p_1, \dots, p_n)$ into spin orbitals $\mathbf{q} = (q_0, q_1, \dots, q_n)$,

$$U(\theta_{\mathbf{pq}}) = e^{-i(\theta_{\mathbf{pq}}/2)G_{\mathbf{pq}}}. \quad (10)$$

The generators $G_{\mathbf{pq}}$ are constructed from pairs of fermionic creation and annihilation operators and mapped to linear

combinations of Pauli strings by various transformations, with the most prominent being the Jordan-Wigner transformation, which we use in this paper. In the case of k -UpCCGSD, the types of excitation are restricted to single and double excitations, where the corresponding spin orbitals p_0, p_1, q_0, q_1 are restricted to the same spatial orbital. If we index the spatial orbitals with P and label the spin-up (spin-down) orbitals with even (odd) numbers, a pair-restricted excitation generator in the Jordan-Wigner representation can be written as

$$G_{(2P,2P+1),(2Q,2Q+1)} \xrightarrow{\text{JW}} i \left(\sigma_{2P}^+ \sigma_{2Q}^- \sigma_{2P+1}^+ \sigma_{2Q+1}^- - \text{H.c.} \right), \quad (11)$$

where $\sigma^\pm = (\sigma^x \pm i\sigma^y)/2$. The single excitations behave similarly but introduce additional terms σ^z into intermediate qubits,

$$G_{2P,2Q} \xrightarrow{\text{JW}} i \left[\sigma_{2P}^+ \left(\prod_{k=2P+1}^{2Q-1} \sigma_k^z \right) \sigma_{2Q}^- - \text{H.c.} \right], \quad (12)$$

and analogously for $G_{2P+1,2Q+1}$. A single layer of the factorized k -UpCCGSD circuit is then constructed from all possible restricted double excitations and the corresponding single excitations excluding spin flips:

$$U_{\text{pCCGSD}}(\boldsymbol{\theta}) = \prod_{P < Q} \left[e^{-i(\theta_{(2P,2Q),(2P+1,2Q+1)}/2)G_{(2P,2Q),(2P+1,2Q+1)}} \times e^{-i(\theta_{2P,2Q}/2)G_{2P,2Q}} \times e^{-i(\theta_{2P+1,2Q+1}/2)G_{2P+1,2Q+1}} \right], \quad (13)$$

with $\boldsymbol{\theta} = (\theta_{0213}, \theta_{02}, \theta_{13}, \dots, \theta_{2P,2Q,2P+1,2Q+1}, \dots)$, following the implementation in Ref. [30]. The full k -UpCCGSD circuit consists of k layers, with individual parameter sets acting on an initial state, commonly chosen as the Hartree-Fock state

$$U(\boldsymbol{\theta}) = \prod_k U_{\text{pCCGSD}}(\boldsymbol{\theta}^{(k)}) U_{\text{HF}}. \quad (14)$$

The unitary U_{HF} prepares the Hartree-Fock state, which is the fermionic state with the first $N_{\text{electrons}}$ molecular spin orbitals occupied. In the Jordan-Wigner encoding, this translates to a simple computational basis state

$$U_{\text{HF}} \xrightarrow{\text{JW}} \prod_{l=0}^{N_{\text{electrons}}} \sigma_l^x. \quad (15)$$

In the standard meta-VQE approach, each UpCCGSD angle is encoded in the same way as in the previous

sections using a metaparameter d and a linear function,

$$\theta_{\text{pq}} = w_{\text{pq}}d + \phi_{\text{pq}}, \quad (16)$$

with independent variational parameters w_{pq} and ϕ_{pq} . Additionally, we introduce a nonlinear encoding (*nl-meta-VQE*) using a single Gaussian to encode each UpCCGSD angle as

$$\theta_{\text{pq}} = \alpha_{\text{pq}}e^{\beta_{\text{pq}}(\gamma_{\text{pq}}-d)} + \delta_{\text{pq}}, \quad (17)$$

with individual variational parameters α_{pq} , β_{pq} , γ_{pq} , and δ_{pq} . In comparison with the previous sections and Fig. 1, this approach contains a static (nonparameterized) initial part U_{HF} and an encoded part U_{pCCGSD} . Note that, although it could be envisioned in future approaches, there is no processing part in the examples explored here.

In Fig. 5, we show the results of the meta-VQE with 2-UpCCGSD applied to a dimer of two hydrogen molecules in a rectangular arrangement (see for example Ref. [17]), using the distance between the two molecules as a metaparameter. At 1.23 Å, this results in a square structure and degeneracies within the orbitals due to a change in the point-group symmetry, making this point challenging for standard single-reference approaches in classical quantum chemistry.

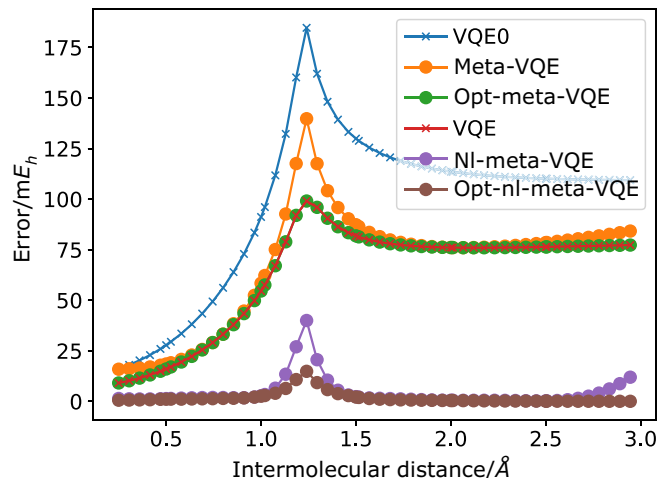


FIG. 5. Relative errors for the rectangular H_4 molecule in eight spin orbitals (STO-3G basis set) in relation to the intermolecular bond distance d . The 2-UpCCGSD model is used for all VQEs. The VQE lines denote the standard optimized 2-UpCCGSD model starting from the Hartree-Fock configuration (VQE0). Linear encoding (meta-VQE) replaces the UpCCGSD angles with an encoding of the form $\theta = \alpha + d\beta$, while the nonlinear encoding (nl-meta-VQE) uses a floating Gaussian with $\theta = \alpha e^{\beta(\gamma-d)} + \delta$. The opt-meta-VQE line denotes the canonical VQE initialized with the angles obtained from the meta-VQE. All errors are with respect to the exact diagonalization of the Hamiltonian within this basis set. Training points are chosen from 0.5 to 2.5 Å with 0.5 Å distance between them.

A single layer of the UpCCGSD ansatz is not able to achieve the accuracies below the millihartree threshold, in both absolute and relative energies, that are required for accurate chemical predictions. Using more than one layer of UpCCGSD gives enough additional freedom to the circuit, but the usual optimization strategy that initializes all angles to zero (i.e., starting from the reference state) fails to converge here and leads to no improvement from the additional layers. The meta-VQE is able to capture the basic form of the VQE results at the individual points and reaches similar accuracy to the canonical VQE at the individual points if the system is not close to the critical point.

In the original work [17], repeated random initialization was used to obtain convergence to the best possible result. Here, we initialize all angles to zero for the meta-VQE as well as for the regular VQE, using the BFGS optimizer in SCIPY with settings that worked well for previously investigated molecular systems [37].

We extend the meta-VQE to a nonlinear encoding as given in Eq. (17), abbreviated as *nl-meta-VQE*, using a single Gaussian to encode each UpCCGSD angle and the intermolecular distance of the two hydrogen molecules as a metaparameter d . The parameters are initialized such that the initial values produced are the same as for the linear meta-VQE and the canonical VQE ($\alpha, \delta = 0, \beta, \gamma = 1$). The nonlinear encoding results in significantly improved convergence, which can be further improved by individually optimizing the individual points with a regular VQE. Floating Gaussians are fairly good general function approximators, and in this example using a single one is sufficient for the corresponding interval of interest. In addition, floating Gaussians are able to capture the asymptotic behavior of molecular distances over the constant offset δ . It is expected that with increasing intermolecular distance d the dependence of the UpCCGSD parameters should asymptotically approach a constant value, since the system will resemble two isolated molecules. A representation with floating Gaussians automatically incorporates this property. In the same way, this holds true for bond distances—see our online material [38] for a detailed example based on a single hydrogen molecule. More general nonlinear encodings are currently under investigation.

In Fig. 5, a standard VQE takes on average 3345 evaluations of the objective function, while opt-nl-meta-VQE takes 204. The training of nl-meta-VQE takes 12 760 evaluations of the objective, including all five training points. The cost of obtaining a given number of test points in terms of individual expectation-value estimations can then be estimated as $12\,760n_{\text{training points}} + 204n_{\text{test points}}$ for opt-nl-meta-VQE and as $3345n_{\text{test points}}$ for the standard VQE. Since we use five training points, opt-nl-meta-VQE becomes cheaper when more than 20 test points are required. Note that those numbers should not be interpreted as rigorous benchmarks. They reflect the state of our

implementation in TEQUILA at the time of writing of this manuscript. We expect, for example, further improvement by incorporating analytical gradients for unitary coupled-cluster methods [36] in combination with an improved automatic differentiation scheme that avoids multiple evaluations of the same objective.

D. Transmon simulation

As a final example, we apply the meta-VQE to a state-of-the-art application of quantum computation: the quantum computer-aided design (QCAD) of superconducting digital quantum computers [29]. We use this example to show how the meta-VQE can be applied straightforwardly to recent work.

We consider the truncated four-qubit Gray-encoded [39] Hamiltonian for a flux-tunable transmon presented in Ref. [29]. Following the original reference, as a circuit ansatz for both the encoding and the processing layer of the meta-VQE, we consider single-qubit rotations R_x and R_z applied to each qubit and parameterized XX gates applied to all six pairs of qubits. As a further simplification, the parameters of the XX gates are common to all layers, leading to the same entangling gate after each layer of single-qubit rotations. Finally, we add a layer of R_x and R_z gates at the end of the processing layer. We encode linearly the Hamiltonian parameter f (flux) in only the rotational gates of the encoding layers, applying the same encoding strategy as for the XXZ Hamiltonian.

In this example, we compare the meta-VQE and opt-meta-VQE with the standard VQE and improved versions

of the VQE. In particular, we define *VQE-smart* as a standard VQE that uses as initial points the optimized points of the previous Hamiltonian-parameter optimization. This approach assumes that the ground state does not change drastically when moving from close points in the Hamiltonian parameter space. We also consider the effect of the encoding layers in this kind of VQE. We define *VQE-enc* as a VQE that has the same parameterized quantum circuit as in the meta-VQE and in opt-meta-VQE but uses as the starting point of each minimization the optimized variables of the previous point, as in VQE-smart. The main difference between VQE-enc and opt-meta-VQE is the choice of the initialization parameters: the result of the previous optimization at each step and the result of the meta-VQE, respectively. This methodology can be further improved by running, for instance, layerwise optimization strategies [40].

The results for one encoding and one processing layer are shown in Fig. 6. The meta-VQE learns the ground-state energy with high accuracy using only 10 training points. The performance of opt-meta-VQE is better than that of the VQE, VQE-smart, and VQE-enc. As one might expect, the VQE with random initialization gets stuck in local minima for some of the test points. On the other hand, VQE-smart and VQE-enc show more consistent results. The encoding strategy in VQE-enc does not prove beneficial for increasing the precision. The increase in the number of parameters to be optimized in comparison with VQE-smart can produce the emergence of local minima or barren plateaus [19], i.e., the gradients are close to zero, making it difficult to explore the parameter space. However, opt-meta-VQE

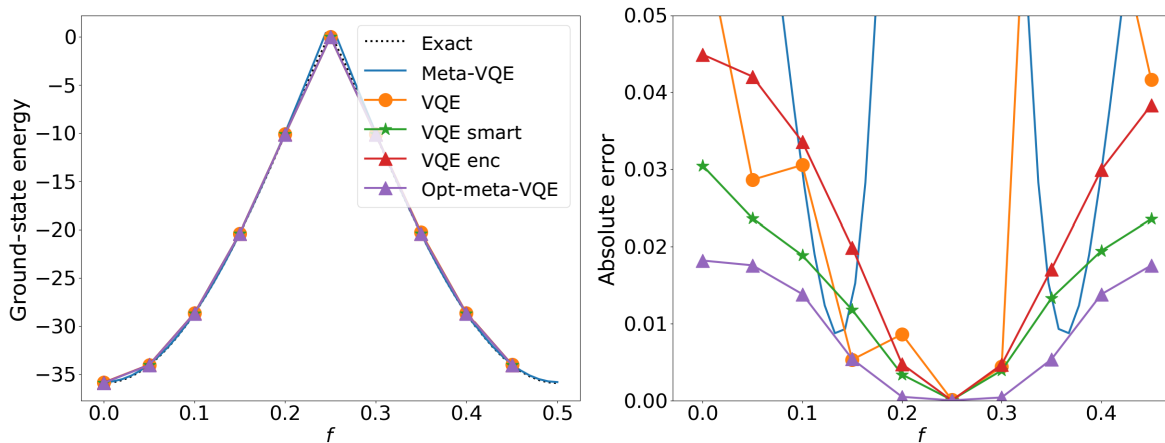


FIG. 6. Left: ground-state energy as a function of the normalized external magnetic flux f for a QCAD [29] simulation of a single transmon. Right: absolute error with respect to the exact diagonalization of the Hamiltonian. We compare the meta-VQE and opt-meta-VQE with the VQE and improved versions of it, where we use the result of the previous optimization point to initialize the parameters of the next one (VQE-smart) and the VQE with encoding layers (VQE-enc). Both VQE-smart and VQE-enc show more consistent results in comparison with the VQE with random initialization, where the algorithm gets trapped in local minima for some points. The introduction of encoding layers in VQE-enc seems not to be beneficial in comparison with the standard VQE circuit (VQE-smart). This could be a consequence of the increase in the number of parameters, where the optimization can face the barren-plateau problem [19]. On the other hand, opt-meta-VQE shows the best performance, indicating that the use of the meta-VQE result as the initialization point better captures the parameter-space region of the ground state.

shows the most precise results. This fact implies that a good initialization is crucial for finding the best minimum. In this case, using the initial parameters provided by the meta-VQE outperforms the initialization strategy of using the previously optimized parameters as a starting point, maybe because the assumption that the ground state does not change significantly from one point to another is not accurate enough.

Besides the better performance of opt-meta-VQE, notice that for the VQE and the VQE with encoding, to obtain any point of the ground-state energy, one needs to first optimize a close point and then use the result to optimize the target point. With opt-meta-VQE this is not necessary, and any point can be obtained by using the result of the meta-VQE. The number of optimization variables involved is 30 for the VQE and 38 for the rest (the meta-VQE, opt-meta-VQE, and the VQE with encoding).

III. DISCUSSION

We present the meta-VQE, an algorithm to compute the ground-state energy of a parameterized Hamiltonian. In contrast to the standard VQE, the meta-VQE learns the Hamiltonian ground-state structure and delivers an initial circuit parameterization that can be used to compute any ground-state energy. To do so, the meta-VQE circuit is divided into two parts, the encoding and processing unitaries. First, this circuit is trained with a set of Hamiltonian parameters, which are encoded in the gates of the encoding unitary. By designing a cost function with the expected values of all these Hamiltonian training points, the algorithm extracts the optimal values of the variational parameters. Then, we can compute the energy for other Hamiltonian values by just running the meta-VQE circuit with the parameters obtained in the minimization. In addition, we can also use the result of a meta-VQE training process as a starting point for a standard VQE algorithm, referred to as opt-meta-VQE, instead of using random initialization. The meta-VQE captures global correlation with a few training points, alleviating the need for refined optimization of the individual points in a later step. With this approach, we do not intend primarily to find stationary points on the potential-energy surface but instead to capture the dependence of the VQE model parameters with metaparameters that appear in the Hamiltonian in specific subspaces.

We demonstrate the efficacy of the algorithm numerically by simulating a spin-chain Hamiltonian (the XXZ model), the electronic Hamiltonian of H_4 , and a single-transmon Hamiltonian, representing different scenarios for future applications. In the first example, we run a meta-VQE for an $n = 14$ XXZ spin chain, with two encoding and processing layers and 10 training points. We compare the results with those of a standard VQE, opt-GA-VQE, and opt-meta-VQE. The results show that the meta-VQE

learns the ground-state energy shape but its accuracy is lower than that of a standard VQE. However, in many cases, opt-meta-VQE is more precise than the standard VQE. For the molecular case, the H_4 system illustrates an application where physically inspired circuit construction can be applied. The meta-VQE is again able to capture the basic features of the energy profile. We use this particular example to illustrate a possible extension of the meta-VQE to nonlinear encodings where a simple first approach could globally improve the convergence beyond that of the standard VQE, reaching the full potential of the underlying circuit. In the third example, the meta-VQE learns the ground-state energy with high accuracy, and the performance of opt-meta-VQE is better than that of the standard VQE and of the VQE with encoding.

We use the XXZ spin chain as a test because of its generality. There is no known circuit ansatz to compute the ground state of this model, so the meta-VQE structure is a general one. This example represents a worst-case scenario. In the molecular case, we illustrate a scenario where circuit construction based on physical principles is possible. Finally, the single-transmon example represents a state-of-the-art application and shows how the meta-VQE can be easily adapted to increase the accuracy of both well-known and state-of-the-art algorithms.

The fact that opt-meta-VQE is more precise and is able to avoid some local minima can be interpreted as a new way to try to avoid the barren-plateau problem [19] using global information. For the transmon-qubit simulation example, opt-meta-VQE is significantly more accurate than other initialization strategies. The meta-VQE approach can be added to other solutions to this problem by, for instance, using local cost functions [41,42], adding correlations between the circuit parameters [43], using a different initialization strategy [44], or exploiting noise [40,45]. All these other approaches can be adapted both to the training of the meta-VQE and to opt-meta-VQE.

We can also adapt the meta-VQE to other improved VQE models [46–50]. The encoding strategy also has possibilities for improvement. Classical and quantum machine-learning techniques can be used to encode the Hamiltonian parameters more efficiently. In the end, the aim of this work is to show the general strategy of the meta-VQE when used in the most general way and for the most general VQE circuit. We illustrate the potential of nonlinear encodings with the H_4 example, show how better accuracy can be obtained with the transmon-simulation example, and anticipate further more generalized improvements in the future. One example is the usage of multiple metaparameters, such as multiple bond lengths and angles in the molecular case or the transverse field λ in the XXZ example. In such cases, we expect the nonlinear encoding of Eq. (17), in the form $\theta_{\mathbf{pq}} = f(d_0, d_1, \dots)$, to have significant advantages compared with the linear encoding. Following the data re-uploading strategy of

Ref. [26], we can also encode each metaparameter in one single-qubit rotational gate [e.g., Δ in the R_z gate and λ in the R_y gate in Eq. (5)]. Apart from multidimensional floating Gaussians, we can also envision more sophisticated approaches, such as neural networks, to approximating and training the nonlinear function f . All simulations shown in this work are exact, i.e., they do not contain any noise model. We expect the meta-VQE to suffer similar effects of noise to other VQAs, effects that can be partially compensated by the use of error-mitigation techniques [51].

Finally, we would like to emphasize the gain in efficiency that a meta-VQE introduces. When dealing with parameterized Hamiltonians, the only way to find the configuration with the lowest ground-state energy is to scan over the parameters and run a VQE for each point, i.e., a minimization procedure for each of these points. Gradient-based minimization, as often performed for molecules, can be applied if analytical derivatives of the Hamiltonian are available, but those depend on good initial guesses of the Hamiltonian parameters. Furthermore, such approaches currently are not able to exploit global information from previous points. For multiparameter Hamiltonian studies, this reduction in the number of points to be scanned to find the interesting energy regions is even more relevant. With a meta-VQE, one can scan over the same set of parameters with a single minimization, obtain the energy profile, and identify the areas of interest, e.g., those with minimum (for equilibrium structures) or maximum (for transition-state searches) energy. In combination with more advanced algorithms, such as excited-state optimization, a meta-VQE could be used to represent the ground state in a first coarse search for points of interest, such as points with a low energy gap. This is, for example, the case for the transmon system of Ref. [29] that we illustrate here as one potential use case. Then, if more accuracy is required, one can run an individual VQE minimization or opt-meta-VQE in these identified regions. With this approach, one can save precious computational power by avoiding minimizations in uninteresting regions of the Hamiltonian phase space. In addition, meta-VQEs can be envisioned as a valuable tool for the development of more sophisticated initialization schemes in the spirit of the molecular example investigated in this work.

All data used to generate the plots and extract the conclusions of this article can be found in a public repository [38]. The main code used to obtain these data can also be found in the same repository in a tutorial form.

ACKNOWLEDGMENTS

We are thankful for comments and suggestions from our Matter Lab colleagues, in particular Thi Ha Kyaw and Sukin Sim. This work was supported by the U.S. Department of Energy under Award No. DE-AC02-05CH11231 (LBNL—505736) and Award No. DE-SC0019374. We

acknowledge generous support from Google, Inc. in the form of a Google Focused Award. A.A.-G. also acknowledges support from the Canada Industrial Research Chairs Program and the Canada 150 Research Chairs Program. We are grateful for the generous support of Anders G. Frøseth at the Vannevar Bush Faculty Fellowship Program.

-
- [1] J. Preskill, Quantum Computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 - [2] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, *et al.*, Noisy Intermediate-Scale Quantum algorithms, *ArXiv:2101.08448* (2021).
 - [3] 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).
 - [4] 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).
 - [5] 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).
 - [6] J. R. McClean and A. Aspuru-Guzik, Compact wavefunctions from compressed imaginary time evolution, *RSC Adv.* **5**, 102277 (2015).
 - [7] 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).
 - [8] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, quantum chemistry in the age of quantum computing, *Chem. Rev.* **119**, 10856 (2019).
 - [9] Y. S. Yordanov, D. R. M. Arvidsson-Shukur, and C. H. W. Barnes, Efficient quantum circuits for quantum computational chemistry, *ArXiv:2005.14475* (2020).
 - [10] B. Kraus, Compressed Quantum Simulation of the Ising Model, Efficient quantum circuits for quantum computational chemistry, *Phys. Rev. Lett.* **107**, 250503 (2011).
 - [11] M. Hebenstreit, D. Alsina, J. I. Latorre, and B. Kraus, Compressed quantum computation using a remote five-qubit quantum computer, *Phys. Rev. A* **95**, 052339 (2017).
 - [12] P. Scholl and R. Orus, Kitaev honeycomb tensor networks: Exact unitary circuits and applications, *Phys. Rev. B* **95**, 045112 (2017).
 - [13] F. Verstraete, J. I. Cirac, and J. I. Latorre, Quantum circuits for strongly correlated quantum systems, *Phys. Rev. A* **79**, 032316 (2009).
 - [14] A. Cervera-Lierta, Exact ising model simulation on a quantum computer, *Quantum* **2**, 114 (2018).
 - [15] A. Montanaro and S. Stanisic, Compressed variational quantum eigensolver for the Fermi-Hubbard model, *ArXiv:2006.01179* (2020).
 - [16] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, *ArXiv:1411.4028* (2014).

- [17] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, Generalized unitary coupled cluster wave functions for quantum computation, *J. Chem. Theory Comput.* **15**, 311 (2018).
- [18] K. Nakaji and N. Yamamoto, Expressibility of the alternating layered ansatz for quantum computation, [ArXiv:2005.12537](https://arxiv.org/abs/2005.12537) (2020).
- [19] 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).
- [20] K. Mitarai, T. Yan, and K. Fujii, Generalization of the Output of a Variational Quantum Eigensolver by Parameter Interpolation with a Low-Depth Ansatz, *Phys. Rev. Appl.* **11**, 044087 (2019).
- [21] Y. Cao, G. Giacomo Guerreschi, and A. Aspuru-Guzik, Quantum neuron: an elementary building block for machine learning on quantum computers, [ArXiv:1711.11240](https://arxiv.org/abs/1711.11240) (2017).
- [22] J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum autoencoders for efficient compression of quantum data, *Quantum Sci. Technol.* **2**, 045001 (2017).
- [23] M. Schuld and F. Petruccione, Quantum ensembles of quantum classifiers, *Sci. Rep.* **8**, 2772 (2018).
- [24] 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](https://arxiv.org/abs/1907.05415) (2019).
- [25] M. Schuld, A. Bocharov, K. Svore, and N. Wiebe, Circuit-centric quantum classifiers, *Phys. Rev. A* **101**, 032308 (2020).
- [26] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, Data re-uploading for a universal quantum classifier, *Quantum* **4**, 226 (2020).
- [27] A. Abbas, M. Schuld, and F. Petruccione, On quantum ensembles of quantum classifiers, [ArXiv:2001.10833](https://arxiv.org/abs/2001.10833) (2020).
- [28] D.-B. Zhang and T. Yin, Collective optimization for variational quantum eigensolvers, *Phys. Rev. A* **101**, 032311 (2020).
- [29] T. H. Kyaw, T. Menke, S. Sim, N. P. D. Sawaya, W. D. Oliver, G. Giacomo Guerreschi, and A. Aspuru-Guzik, Quantum computer-aided design: digital quantum simulation of quantum processors, [ArXiv:2006.03070](https://arxiv.org/abs/2006.03070) (2020).
- [30] J. S. Kottmann, S. Alperin-Lea, T. Tamayo-Mendoza, A. Cervera-Lierta, C. Lavigne, T.-C. Yen, V. Verteletskiy, P. Schleich, A. Anand, M. Degroote, *et al.*, TEQUILA: a platform for rapid development of quantum algorithms, *Quantum Sci. Technol.* **6**, 024009 (2021).
- [31] Y. Suzuki, Y. Kawase, Y. Masumura, Y. Hiraga, M. Nakadai, J. Chen, K. M. Nakanishi, K. Mitarai, R. Imai, S. Tamiya, *et al.*, Qulacs: a fast and versatile quantum circuit simulator for research purpose, [ArXiv:2011.13524](https://arxiv.org/abs/2011.13524) (2020).
- [32] A. Langari, Phase diagram of the antiferromagnetic XXZ, *Phys. Rev. B* **58**, 14467 (1998).
- [33] J. S. Van Dyke, G. S. Barron, N. J. Mayhall, E. Barnes, and S. E. Economou, Preparing bethe ansatz eigenstates on a quantum computer, [ArXiv:2103.13388](https://arxiv.org/abs/2103.13388) (2021).
- [34] J. I. Cirac and G. Sierra, Infinite matrix product states, conformal field theory, and the haldane-shastry model, *Phys. Rev. B* **81**, 104431 (2010).
- [35] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [36] J. S. Kottmann, A. Anand, and A. Aspuru-Guzik, A feasible approach for automatically differentiable unitary coupled-cluster on quantum computers, *Chem. Sci.* **12**, 3497 (2021).
- [37] J. S. Kottmann, P. Schleich, T. Tamayo-Mendoza, and A. Aspuru-Guzik, Reducing qubit requirements while maintaining numerical precision for the variational quantum eigensolver: a basis-set-free approach, [ArXiv:2008.02819](https://arxiv.org/abs/2008.02819) (2020).
- [38] A. Cervera-Lierta, J. S. Kottmann, and A. Aspuru-Guzik, Data for: “Meta-variational quantum eigensolver: Learning energy profiles of parameterized Hamiltonians for quantum simulation,” <https://github.com/aspuru-guzik-group/Meta-VQE> (2021).
- [39] N. P. D. Sawaya, T. Menke, T. H. Kyaw, S. Johri, A. Aspuru-Guzik, and G. G. Guerreschi, IBM Q experience as a versatile experimental testbed for simulating open quantum systems, *NPJ Quantum Inf.* **6**, 1 (2020).
- [40] A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib, Layerwise learning for quantum neural networks, [ArXiv:2006.14904](https://arxiv.org/abs/2006.14904) (2020).
- [41] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow quantum neural networks, [ArXiv:2001.00550](https://arxiv.org/abs/2001.00550) [quant-ph] (2020).
- [42] S. P. Premaratne and A. Y. Matsuura, Engineering the cost function of a variational quantum algorithm for implementation on near-term devices, [ArXiv:2006.03747](https://arxiv.org/abs/2006.03747) (2020).
- [43] T. Volkoff and P. J. Coles, Large gradients via correlation in random parameterized quantum circuits, [ArXiv:2005.12200](https://arxiv.org/abs/2005.12200) (2020).
- [44] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, *Quantum* **3**, 214 (2019).
- [45] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, [ArXiv:2007.14384](https://arxiv.org/abs/2007.14384) (2020).
- [46] 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, and I. Tavernelli, Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wave-function expansions, *Phys. Rev. A* **98**, 022322 (2018).
- [47] A. Cowtan, W. Simmons, and R. Duncan, A generic compilation strategy for the unitary coupled cluster ansatz, [ArXiv:2007.10515](https://arxiv.org/abs/2007.10515) (2020).
- [48] K. Fujii, K. Mitarai, W. Mizukami, and Y. O. Nakagawa, Deep variational quantum eigensolver: a divide-and-conquer method for solving a larger problem with smaller size quantum computers, [ArXiv:2007.10917](https://arxiv.org/abs/2007.10917) (2020).
- [49] W. J. Huggins, J. Lee, U. Baek, B. O’Gorman, and K. B. Whaley, A non-orthogonal variational quantum eigensolver, *New J. Phys.* **22**, 073009 (2020).
- [50] D. Chivilikhin, A. Samarin, V. Ulyantsev, I. Iorsh, A. R. Oganov, and O. Kyriienko, MoG-VQE: Multiobjective genetic variational quantum eigensolver, [ArXiv:2007.04424](https://arxiv.org/abs/2007.04424) (2020).
- [51] S. Endo, S. C. Benjamin, and Y. Li, Practical quantum error mitigation for near-future applications, *Phys. Rev. X* **8**, 031027 (2018).