Generalization of the Output of a Variational Quantum Eigensolver by Parameter Interpolation with a Low-depth Ansatz

Kosuke Mitarai, Tennin Yan, and Keisuke Fujii

Generalization of the output of variational quantum eigensolver by parameter interpolation with low-depth ansatz

Kosuke Mitarai,[1, 2, *] Tennin Yan,[2, †] and Keisuke Fujii[3, 4, ‡]

[1]*Graduate School of Engineering Science, Osaka University,*
1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan.
[2]*Qunasys Inc. High tech Hongo Building 1F, 5-25-18 Hongo, Bunkyo, Tokyo 113-0033, Japan*
[3]*Graduate School of Science, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8302, Japan.*
[4]*JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

The variational quantum eigensolver (VQE) is an attracting possible application of near-term quantum computers. Originally, the aim of the VQE is to find a ground state for a given specific Hamiltonian. It is achieved by minimizing the expectation value of the Hamiltonian with respect to an ansatz state by tuning parameters $\boldsymbol{\theta}$ on a quantum circuit which constructs the ansatz. Here we consider an extended problem of the VQE, namely, our objective in this work is to "generalize" the optimized output of the VQE just like machine learning. We aim to find ground states for a given set of Hamiltonians $\{H(\boldsymbol{x})\}$, where $\boldsymbol{x}$ is a parameter which specifies the quantum system under consideration, such as geometries of atoms of a molecule. Our approach is to train the circuit on the small number of $\boldsymbol{x}$'s. Specifically, we employ the interpolation of the optimal circuit parameter determined at different $\boldsymbol{x}$'s, assuming that the circuit parameter $\boldsymbol{\theta}$ has simple dependency on a hidden parameter $\boldsymbol{x}$ as $\boldsymbol{\theta}(\boldsymbol{x})$. We show by numerical simulations that, using an ansatz which we call the Hamiltonian-alternating ansatz, the optimal circuit parameters can be interpolated to give near-optimal ground states in between the trained $\boldsymbol{x}$'s. The proposed method can greatly reduce, on a rough estimation by a few orders of magnitude, the time required to obtain ground states for different Hamiltonians by the VQE. Once generalized, the ansatz circuit can predict the ground state without optimizing the circuit parameter $\boldsymbol{\theta}$ in a certain range of $\boldsymbol{x}$.

I. INTRODUCTION

Recent progress in the experimental realization of quantum computers is stimulating the research of their possible applications. Quantum computers which might realize in near-future are often called noisy intermediate-scale quantum (NISQ) devices [1]. Among possible applications of NISQ devices, variational quantum algorithms have recently attracted much attention [2–6]. In this variational approach, quantum computers have parameters which are optimized using classical algorithms with respect to a cost function, such as an expectation value of a Hamiltonian. To apply variational quantum algorithms to a problem, we divide the problem into two parts; one part which can be readily computed classically, and the other which is hard for classical computers but easy for quantum computers.

The most popular two of variational algorithms are the variational quantum eigensolver [2, 4, 7] (VQE) for quantum chemistry and materials science, and quantum approximate optimization algorithm [5, 8, 9] (QAOA) for combinatorial optimization problems. There are also many theoretical proposals which extend the hybrid framework to machine learning [10–17] along with the experimental demonstrations [6, 18]. We refer to these approach as quantum circuit learning (QCL) following Ref. [10]. In VQE or QAOA, a problem instance is en-

coded into an Hermitian operator $H(\boldsymbol{x})$, where $\boldsymbol{x}$ is an input parameter of the instance, such as the geometry of a molecule, external fields applied to a quantum system, or topology of a graph. Then we try to minimize an expectation value of $H(\boldsymbol{x})$ with respect to an ansatz state $|\psi(\boldsymbol{\theta})\rangle$, which is generated by a parameterized quantum circuit, by variating the parameter $\boldsymbol{\theta}$. The optimized parameter $\boldsymbol{\theta}^*$ gives a good approximate of the ground state of $H(\boldsymbol{x})$. On the other hand, the QCL tries to minimize the cost function L which depends on multiple inputs $\{\boldsymbol{x}_\alpha\}$ and multiple ansatz state $\{|\psi(\boldsymbol{x}_\alpha, \boldsymbol{\theta})\rangle\}$. Specifically, in supervised learning, the cost function measures the difference between the teacher data $\{y_j\}$ and the corresponding output with respect to the set of input $\{\boldsymbol{x}_\alpha\}$. We can, therefore, construct a trained quantum circuit that gives outputs close to the teacher, by minimizing the cost function [10]. The main difference between VQE/QAOA and QCL is that whether the output is "generalized". In the machine learning approach, the output is generalized in the sense that a trained circuit can predict the correct answers even when it is provided with a wide range of unknown inputs. In contrast, VQE and QAOA are specialized in solving one problem with a specific input. However, in many-body physics and chemistry, there is a high demand to know the changes of important properties of the system, such as the ground state energy or the order parameter, with respect to the system parameter $\boldsymbol{x}$, such as the distance between the atoms or the magnitude of the applied magnetic field.

In this work, we extend VQE to give us "generalized" outputs in the sense described above. Specifically, we aim to find a trained circuit which can output approximate

ground states in a certain range of the input $\boldsymbol{x}$. Classical machine learning approach using the Boltzmann machine has succeeded in learning the ground states of quantum many-body systems [19], however, we expect that the quantum ansatz is more suited for the objective. We first propose the ansatz which, we claim from numerical simulation, are able to represent ground states with the small number of parameters in Sec. II. The ansatz includes the adiabatic state preparation (ASP) in the limit of the infinite depth, which always outputs the ground state of Hamiltonian $H(\boldsymbol{x})$ with the sufficiently large annealing time and it is irrespective of the parameter $\boldsymbol{x}$. However, the depth-limited ansatz does not allow us to achieve the robustness of the ASP with respect to the change in the parameter $\boldsymbol{x}$, as shown in Appendix by numerical simulations. Instead, the interpolation between the optimal parameters $\{\boldsymbol{\theta}^*(\boldsymbol{x}_\alpha)\}$ of the ansatz at different inputs $\{\boldsymbol{x}_\alpha\}$ can be utilized to achieve the objective of generalization. Numerical simulations using Hamiltonians of hydrogen molecule, H_3 molecule, and water molecule in Sec. III support this claim. Especially, for hydrogen and H_3 molecule, we obtained the precision on the order of 10^{-5} Hartree in between the training points.

This work can be regarded as one of the first approach to apply a quantum computer, especially a NISQ device, to learn a quantum system, by finding hidden patterns in ground states. It can significantly reduce the time required to find ground states in a wide range of parameters, which can be used for, for example, drawing a potential energy surface of a chemical compound. The application of the proposed method is not limited to NISQ devices. In fault-tolerant quantum computation, which is a long-term goal for the experiments, the approximate ground state prepared by our method can then be used as an input to the phase estimation algorithm for a more accurate and efficient energy determination.

II. METHODS

A. Variational quantum eigensolver

Here we briefly review the algorithm called VQE [2]. The VQE aims to find a ground state of a Hamiltonian $H(\boldsymbol{x})$ using the variational approach, where $\boldsymbol{x}$ are parameters that specify a quantum system such as a molecule. For an electronic problem with two-body interactions, Hamiltonians have the following form,

$$H(\boldsymbol{x}) = \sum_{ij} h_{ij}(\boldsymbol{x}) c_i^\dagger c_j + \sum_{ijkl} h_{ijkl}(\boldsymbol{x}) c_i^\dagger c_j^\dagger c_k c_l, \quad (1)$$

where $c_i^\dagger$ and c_i are fermionic creation and annihilation operators, and $h_{ij}(\boldsymbol{x})$ and $h_{ijkl}(\boldsymbol{x})$ are coefficients. When applying VQE to molecular problems, we usually start with Hamiltonians that are described in Hartree-Fock (HF) basis [20]. We then map the fermionic Hamiltonian Eq. (1) to a qubit Hamiltonian using a conversion

such as Jordan-Wigner or Bravyi-Kitaev transformations [21, 22]. After the transformation, Hamiltonians acting on an n-qubit system have the following form,

$$H(\boldsymbol{x}) = \sum_{P \in \mathcal{P}_{diag}} h_P^{(HF)}(\boldsymbol{x}) P + \sum_{Q \in \mathcal{P}_{nondiag}} h_Q^{(cor)}(\boldsymbol{x}) Q, \quad (2)$$

where $\mathcal{P}_{diag} \subset \{I, Z\}^{\otimes n}$ is a set of Pauli products which only contains Pauli Z's and are diagonal in the computational basis, and $\mathcal{P}_{nondiag} \subset \{I, X, Y, Z\}^{\otimes n} \setminus \mathcal{P}_{diag}$ is a set of Pauli products which are not diagonal in the computational basis. $h_P^{(HF)}(\boldsymbol{x})$ and $h_Q^{(cor)}(\boldsymbol{x})$ are real-valued coefficients. The ground state energy calculated by the HF approximation only depends on the first sum of Eq. (2). The second sum of Eq. (2) determines the correlation energy, which is discarded in the HF approximation. For later convenience, we define the HF Hamiltonian as $H_{HF} = \sum_{P \in \mathcal{P}_{diag}} h_P^{(HF)}(\boldsymbol{x}) P$, and the correlation Hamiltonian as $H_{cor} = \sum_{Q \in \mathcal{P}_{nondiag}} h_Q^{(cor)}(\boldsymbol{x}) Q$.

To find a ground state, we construct a specific ansatz $|\psi(\boldsymbol{\theta})\rangle$ that depends on the variational parameter $\boldsymbol{\theta}$. It is usually created by applying a parameterized unitary gate $U(\boldsymbol{\theta})$ to an initialized state $|0\rangle^{\otimes n}$. There are several theoretical proposals addressing the form of $U(\boldsymbol{\theta})$ [23–25].

B. Variational ansatz

We extend the VQE so that an optimized quantum circuit gives us a generalized output with respect to the parameter $\boldsymbol{x}$, which specifies Hamiltonians $H(\boldsymbol{x})$, that is, we aim to construct a quantum circuit that outputs ground states of $\{H(\boldsymbol{x})\}$. When we want to find a ground state and its energy of each Hamiltonian $H(\boldsymbol{x})$, the ansatz, in general, can be an arbitrary one that has enough power to represent ground states of each $\{H(\boldsymbol{x})\}$ by optimizing the parameter $\boldsymbol{\theta}$ of the ansatz at each $\boldsymbol{x}$ independently. However, for the objective considered here, the ansatz must be constructed in a form that includes the parameter $\boldsymbol{x}$ and can represent the ground states.

Our idea is to use an ansatz which is inspired by the ASP [8]. The ASP finds a ground state of a Hamiltonian $H(\boldsymbol{x})$ by implementing a time-dependent Hamiltonian,

$$H_{ann}(t) = A(t) H_0 + B(t) H(\boldsymbol{x}), \quad (3)$$

where $A(T) = B(0) = 0$, $A(0) = B(T) = 1$ and H_0 is a Hamiltonian of which ground state is trivial, for a duration T. The output of the ASP is always the ground state of $H(\boldsymbol{x})$ when T is sufficiently large. The ASP with the large T can be regarded as an ansatz circuit parameterized by $H(\boldsymbol{x})$ itself that outputs the ground state of $H(\boldsymbol{x})$ without any optimization, that is, it does not require any training.

The implementation of the ASP is unlikely on the NISQ devices, while it can be a nice starting point for the construction of the ansatz for the generalization considered here. We seek to resolve the objective considered here with what we call an Hamiltonian-alternating
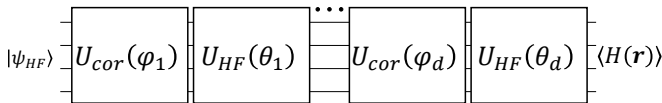
FIG. 1. Hamiltonian-alternating ansatz for generalized VQE.

ansatz,

$$U(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \prod_{k=1}^{d} U_{HF}(\theta_k) U_{cor}(\varphi_k). \qquad (4)$$

The quantum circuit corresponding to this is shown in Fig. 1. Hereafter we refer to the integer d as the depth of the ansatz circuit. $U_{HF}(\theta_k)$ is a unitary operator generated by the HF Hamiltonian $H_{HF}(\boldsymbol{x})$; $U_{HF}(\theta_k) = e^{-i\theta_k H_{HF}(\boldsymbol{x})}$, and $U_{cor}(\varphi_k)$ is a unitary operator generated by the non-diagonal part of the qubit Hamiltonian $H_{cor}(\boldsymbol{x})$; $U_{cor}(\varphi_k) = e^{-i\varphi_k H_{cor}(\boldsymbol{x})}$. We use the HF ground state $|\psi_{HF}\rangle$ as the input to the circuit. This ansatz includes the ASP in the limit of $d \to \infty$ and appropriate $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ are chosen. The optimization of the circuit parameters $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ at a certain d finds the shorter circuit than the ASP, although the optimized circuit with optimal $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ can lose the robustness with respect to the change of $\boldsymbol{x}$ of the ASP. The ansatzes having similar forms can be found in Ref. [8, 23]. Especially, in the first proposal of QAOA [8], they pointed out the connection between the ansatz and quantum annealing.

We need the Trotterization of $U_{cor}(\varphi_k)$ for accurate execution of it. However, in this work, we simplify it to one step Trotter expansion considering the limited circuit depth of the NISQ devices;

$$\tilde{U}_{cor}(\varphi_k) = \prod_{Q \in \mathcal{P}_{nondiag}} \exp(-i\varphi_k h_Q^{(cor)} Q), \qquad (5)$$

where the product is taken in an arbitrary order.

C. Generalizing Variational Quantum Eigensolver

To "generalize" the outputs of an optimized circuit, we first tried simultaneous minimization of $H(\boldsymbol{x})$ at different $\boldsymbol{x}_\alpha$'s. The subscript α denotes an index of points on which the quantum circuit is trained. In this approach, we set the cost function, which we aim to minimize by variating $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$, as $\sum_\alpha \langle \psi(\boldsymbol{x}_\alpha, \boldsymbol{\theta}, \boldsymbol{\varphi})| H(\boldsymbol{x}_\alpha) |\psi(\boldsymbol{x}_\alpha, \boldsymbol{\theta}, \boldsymbol{\varphi})\rangle$. Note that the ansatz state $|\psi(\boldsymbol{x}_\alpha, \boldsymbol{\theta}, \boldsymbol{\varphi})\rangle$ includes the parameter $\boldsymbol{x}$ since the ansatz is constructed from the Hamiltonian $H(\boldsymbol{x})$. Unfortunately, we found that this approach does not work well at small d's, as shown in Appendix. With the low-depth ansatz, the optimization could not find the optimal $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ that is robust to the change in the input parameter $\boldsymbol{x}$. The reason for this may be that the compression of the ASP to the low-depth circuit depends on the parameter $\boldsymbol{x}$, and the depth of the circuit was traded-off with the robustness to the change in $\boldsymbol{x}$. The ASP is

included in the limit of $d \to \infty$ and in that case the ansatz must work, however, the depth of such a circuit is prohibitively large for the NISQ devices.

To avoid the above issue and find a low-depth ansatz that returns a ground state irrespective to the change of the external parameter $\boldsymbol{x}$, we propose another approach to "generalize" with low-depth circuits. We seek a solution by assuming that the optimal $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ themselves also depend on the input parameter $\boldsymbol{x}$; $\boldsymbol{\theta} \to \boldsymbol{\theta}(\boldsymbol{x})$ and $\boldsymbol{\varphi} \to \boldsymbol{\varphi}(\boldsymbol{x})$, since we have found in the preliminary simulations that the optimal $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ tended to have simple trends. To find these functions, we interpolate the optimal parameters determined at different $\boldsymbol{x}$'s. The procedure is as follows. First we optimize the parameter $\boldsymbol{\theta}, \boldsymbol{\varphi}$ independently at different $\boldsymbol{x}_\alpha$'s. The resulting optimal parameters are denoted as $\boldsymbol{\theta}^*(\boldsymbol{x}_\alpha)$ and $\boldsymbol{\varphi}^*(\boldsymbol{x}_\alpha)$. Then we interpolate the points between $\boldsymbol{\theta}^*(\boldsymbol{x}_\alpha)$ and $\boldsymbol{\varphi}^*(\boldsymbol{x}_\alpha)$. The interpolation gives us the approximately optimal parameter function $\boldsymbol{\theta}^*(\boldsymbol{x})$ and $\boldsymbol{\varphi}^*(\boldsymbol{x})$ in between the training points $\{\boldsymbol{x}_\alpha\}$. The ansatz in this approach can be written as:

$$U(\boldsymbol{\theta}(\boldsymbol{x}), \boldsymbol{\varphi}(\boldsymbol{x})) = \prod_{k=1}^{d} U_{HF}(\theta_k(\boldsymbol{x})) \tilde{U}_{cor}(\varphi_k(\boldsymbol{x})). \qquad (6)$$

We note that the reduction in the number of parameters might be crucial for this approach.

III. NUMERICAL SIMULATION

For all simulations that we present in this section, the molecular Hamiltonian is calculated by OpenFermion [26], OpenFermion-Psi4 [26], and Psi4 [27]. We used the STO-3G minimal basis set and the Jordan-Wigner transformation for all calculations. Therefore the number of qubits in the simulated quantum circuit in this section were up to 14 for the water molecule in Sec. III D.

At training points, the parameters were optimized using BFGS method [28] provided in SciPy library. Quantum circuits were simulated with a variational quantum circuit simulator Qulacs [29]. The gradients can be calculated using the method described in Ref. [10, 30]. The starting points for the optimization were chosen randomly near the origin. More specifically, we sampled from the uniform distribution on $[0, 10^{-2}]$ for initial values for each parameter $\{\theta_k\}, \{\varphi_k\}$. The optimization procedure is repeated 10 times starting from different initial parameters, and we picked the best one before the interpolation. After the optimal parameters are found, we used the quadratic interpolation.

A. Hydrogen molecule

First we consider the hydrogen molecule. In this simulation, the parameter $\boldsymbol{x}$ of the Hamiltonian is the distance r between two hydrogen atoms. We
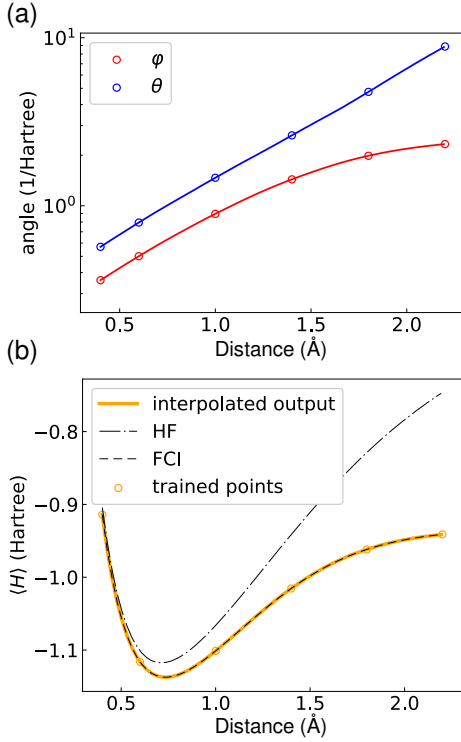
FIG. 2. The result of numerical simulation for the hydrogen molecule with circuit depth $d = 1$. (a) Optimal parameters plotted against the interatomic distance r. Circles: the best results among the optimizations started from 10 different initial values. Solid lines: drawn from quadratic interpolation between the markers. (b) Markers: optimized output at the training points. Solid line: output from the quantum circuit using the interpolated parameters. The HF and FCI energies are plotted for reference.

search for the optimal parameter $\boldsymbol{\theta}^*$, $\boldsymbol{\phi}^*$ at $\{r_\alpha\} = \{0.4, 0.6, 1.0, 1.4.1.8, 2.2\}$ Å. We found that the ansatz in Fig. 1 with $d = 1$ can achieve the full configuration interaction (FCI) energy almost to the machine precision, and therefore the all of the result in this section is for $d = 1$.

Fig. 2 (a) shows the dependency of optimal paramters. One can clearly see the trend in the optimal θ and φ at the training points $\{r_\alpha\}$ drawn by the markers. The solid lines in the figure are drawn from the interpolation. Using these interpoalted parameters, in Fig. 2 (b) we plot the output $\langle H(r, \boldsymbol{\theta}^*(r), \boldsymbol{\varphi}^*(r)) \rangle$ from the optimized quantum circuit. The output well matches the exact solution calculated by the FCI method. The error between the interpolated output and the FCI energy does not exceed 6×10^{-5} Hartree.

B. Linearly aligned three hydrogen chain

Here we consider a somewhat artificial problem of finding ground state energies of the linearly aligned H$_3$ molecule, to see if the proposed method works in bigger
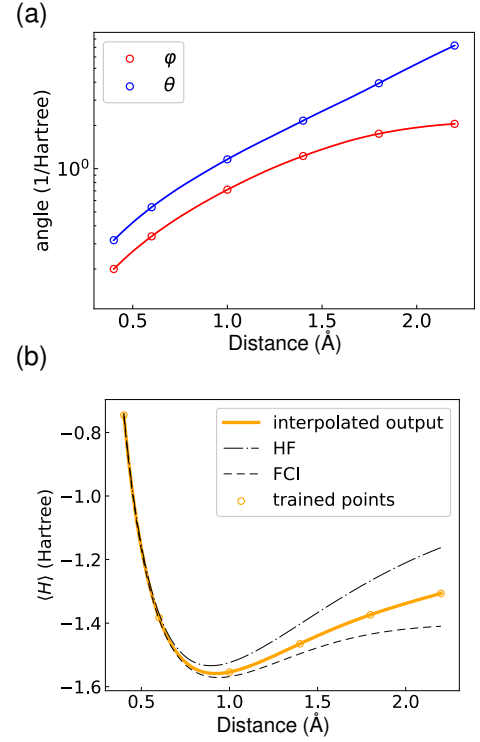

FIG. 3. The result of numerical simulation for the linearly aligned H_3 with circuit depth $d = 1$. The notations are similar to that of Fig. 2.

quantum systems. When we use the minimal basis set for this molecule, which is the case studied here, the system that we aim to solve can be regarded as a three-site Fermi-Hubbard model with periodic boundary conditions at half-filling. The coordinates of three hydrogen atoms are set to $-r$, 0, and r, respectively, with r being the parameter $\boldsymbol{x}$ which determines a Hamiltonian. Training points are set to $\{r_\alpha\} = \{0.4, 0.6, 1.0, 1.4, 1.8, 2.2\}$ Å.

The result for depth $d = 1$ and $d = 2$ are shown in Figs. 3 and 4, respectively. From Fig. 3, it is clear that the circuit with $d = 1$ is not capable of describing the ground state of the system since the output is not achieving the FCI energy. On the other hand, increasing the circuit depth to $d = 2$, the optimized output reaches the FCI energy as shown in Fig. 4, indicating the circuit can generate the approximate ground state. It is notable that the number of parameter at $d = 2$ is only 4 while the dimension of Hilbert space under the search is $\begin{pmatrix} 6 \\ 3 \end{pmatrix} = 20$. The interpolation of the parameter also works nicely in this case, probably benefitting from the small number of parameters used in the ansatz. The maximum error between the FCI energy and the optimized output is 2×10^{-3} Hartree at the training point $r = 2.2$Å
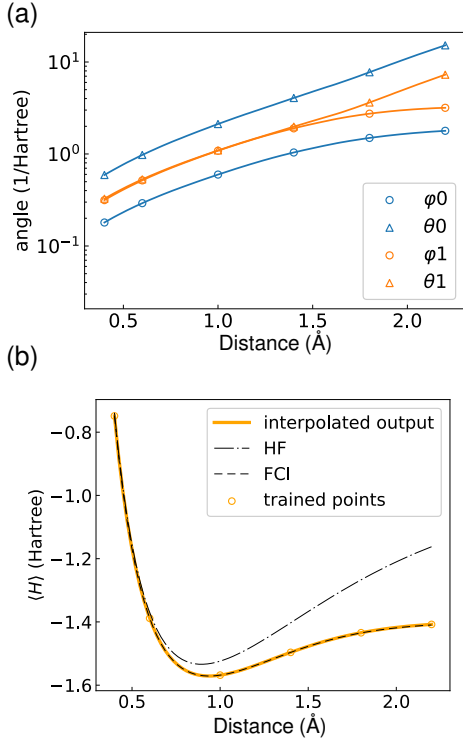
FIG. 4. The result of numerical simulation for the linearly aligned H_3 with circuit depth $d = 2$. The notations are similar to that of Fig. 2.

FIG. 5. The result of numerical simulation for triangle H_3^+ with circuit depth $d = 2$. The notations are similar to that of Fig. 2.

C. Triangle H_3^+ ion

We performed simulation on the triangle H_3^+ ion. The hydrogen atoms are placed at $(0,0)$, $(r,0)$, and $(r/2, \sqrt{3}r/2)$, with r being the parameter which determines the Hamiltonian. Here we use $\{r_\alpha\} = \{0.5, 1.0, 1.5, 2.0, 2.5\}$Å We found, as same as in the case of the previous section, the output well converges to FCI energy at $d = 2$.

Fig. 5 shows the simulated result of $d = 2$. The optimal parameters do not appear to have the simple trend as in the previous two examples, nevertheless, the interpolation approach works well in this case too. The error between the FCI energy and the output does not exceed 2×10^{-5} Hartree.

D. Water molecule

Here we simulate the water molecule H_2O. In this simulation, we fix the bond length r between H and O to a constant ($r = 0.96$Å) and set the angle β of H-O-H bonding as the parameter of the Hamiltonian; the oxygen atom was placed at $(0,0)$ and the two hydrogen atoms were placed at $(r,0)$ and $(r\cos\beta, r\sin\beta)$. The optimization was performed at $\{\beta_\alpha\} = \{54, 72, 90, 108, 126, 154\}$ deg.
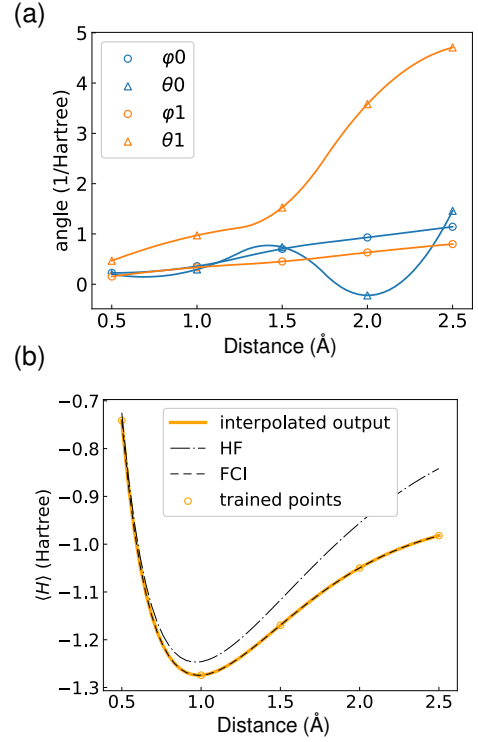
The result for $d = 2$ are shown in Fig. 6. In this

case, the optimal parameters give the better approximation than the HF approximation but do not reach the FCI energy. It is due to the limited representation power at $d = 2$. From the energy plot in Fig. 6 (b) we conclude that the interpolation approach can be utilized even at the level of 14 qubits. However, we found that, at $d = 3$ and $d = 4$, the parameter tends to be trapped at a local minima due to a complicated landscape of the energy expectation value. In those cases, the parameter found by the optimization process did not have the simple trends as in the figures, therefore the interpolation fails to predict the ground state energy in between the training points. More sophisticated optimization algorithms, such as the imaginary time evolution [31], or the usage of the optimal parameter at one training point as the initial parameter for another training point, can be a solution to this problem by finding the global minimum of the energy.

IV. DISCUSSION

We first discuss the applicability of the presented method. The dependence of $\boldsymbol{\theta}^*$ on the system parameter $\boldsymbol{x}$ might not be as smooth as the examples which we have shown in the previous section. For example, we do not expect the method to work in the case where the system undergoes the phase transition. In this case,
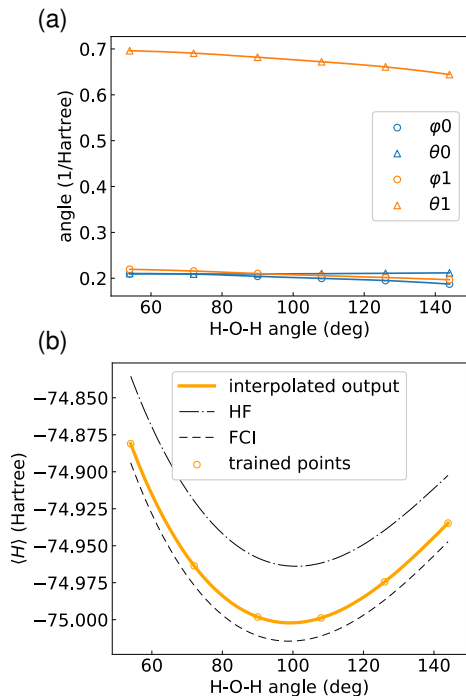
(a)

(b)

FIG. 6. The result of numerical simulation for H_2O with circuit depth $d = 2$. The notations are similar to that of Fig. 2.

the dependence of $\boldsymbol{\theta}^*$ might even be discontinuous at the transition point. There can be more instances where $\boldsymbol{\theta}^*$ does not have as simple dependency on $\boldsymbol{x}$ as the presented examples. In those cases, one can either increase the number of training points to track such complicated dependency, or change the ansatz circuit that has more simple $\boldsymbol{\theta}^*$. The detailed strategy should be studied in a future research. However, we note that the ansatz used in the examples gave us nice results.

The potential impact of this work can be estimated as follows. In Ref. [4], the experimental VQE of H_2, LiH_2, BeH_2 and the Heisenberg model took $\approx$ 200-300 iterations to converge. They employed the simultaneous perturbation stochastic approximation method for the optimization. They took 10^3 samples in the optimization procedure, resulting 10^5 to 10^6 samplings in total at one r, the interatomic distance. Thus, for example, to plot an energy landscape using 10^2 points of r, the number of samples can go up to 10^8. Since, from numerical simulations, our approach only needs around 10 optimizations at different r's for the same task, the requirement can be dramatically lowered to the tenth of the original approach. Note that the direct interpolation between the calculated energies $\{\langle H(\boldsymbol{x}_\alpha)\rangle\}$ might be able to do the same task, but our approach can prepare useful quantum outputs, that is, the approximate ground states in a certain range of $\boldsymbol{x}$, not only the classical output like the energy. The approximate ground state can provide us other observables such as electron density. Also, we have observed that the direct interpolation approach tends to

give a poorer approximation than the interpolation of the parameter utilized in this work.

To conclude, we proposed an idea to "generalize" the output of VQE. The simple interpolation of the circuit parameters, not of the calculated energy, enables us to resolve the objective. We gave the numerical evidence for the effectiveness of the approach. If one needs more accuracy, the interpolated values of parameters can be used as the starting point of the optimization for VQE. Our proposal can greatly reduce the time required for obtaining the ground states for a set of Hamiltonian $\{H(\boldsymbol{x})\}$ characterized by some parameter $\boldsymbol{x}$, and therefore advances the practicability of VQE.

Appendix: Simultaneous optimization

In this Appendix we show the result with simultaneous optimization approach which is mentioned in Section II C. Unfortunately, we find that this approach does not work well at small d's that we simulated. The compression of the ASP to the low-depth circuit depends on the parameter $\boldsymbol{x}$, and the depth of the circuit might be traded-off with the robustness to the change in $\boldsymbol{x}$. This has led us to explore the approach described in the main text. For numerical simulations, same procedures and packages as the main text is utilized for following numerical simulations.

1. H_2

In this simulation, the parameter $\boldsymbol{x}$ of the Hamiltonian is the distance r between two hydrogen atoms.

We compare the result with our ansatz (Fig. 1) and that with so-called hardware efficient (HE) ansatz [4] (Fig. 7). Parameters are optimized to minimize the cost Eq. (II C) with BFGS method, which is a gradient-based algorithm, using python library SciPy [32].

Fig. 8 (a) and (b) show an example of the output from the optimized circuit with our ansatz and the hardware efficient ansatz with the same circuit depth $d = 5$, respectively. We performed the optimization using molecular Hamiltonians at r = 0.4, 0.6, 1.0, and 1.4 Å. In spite of the fact that the quantum circuit of the HE ansatz does not depend on Hamiltonians and always outputs the same quantum state, the energy obtained from the HE ansatz achieves quite close to FCI energy in a wide range. However, due to the lack of the information about the Hamiltonians to be minimized in the circuit, it performs

poorly when compared with the Hamiltonian-alternating ansatz. This is apparent especially at the training points of $r = 0.4$ and 1.4 Å. With the Hamiltonian-alternating ansatz, one can see that the trained points are actually optimized at the same time, and also that the optimized circuit gives us the energies that are close to the ones obtained from FCI calculations, even in the ranges between the trained points.

Fig. 8 (c) shows the depth dependency of our algorithm. We have conducted optimizations with the Hamiltonian-alternating ansatz with depths $d = 1$ to 9. Optimization was performed for 10 times for each depth starting from different initial parameters that are chosen randomly, and we picked an optimized circuit that gave us the lowest cost. One can see that with increasing depth of the ansatz, the trained points get close to FCI energies, as expected. On the other hand, increasing depth sometimes lowers the performance in between the training points.

2. Linearly aligned H_3

The coordinates of three hydrogen atoms are set to $-r$, 0, and r, respectively, with r being the parameter $\boldsymbol{x}$ which determines a Hamiltonian. We trained quantum circuits on the four values of the parameter $r = 0.6, 1.0,$ 1.4, and 1.8 Å, using the same method as the previous section.

The optimized output from the HE ansatz with $d = 5$ is shown in Fig. 9 for comparison. As same as in Fig. 8, the circuit parameters seems to be tuned as to minimize the Hamiltonian at around the middle of the trained points. It is as expected because intuitively the state which minimizes the Hamiltonian at the middle point would minimize the cost Eq. (II C). However, the output at $r = 0.6$ fails to achieve even to the HF energy.

Fig. 9 (b) shows the depth dependency of the optimized output from the AI ansatz. The results are generated with the same method as the previous section. The output at the trained points gets close to FCI energy as we increase the depth of the circuit, however, we can observe the complex behavior in deeper circuits, which is not apparent in H_2 problem. We suspect that it is due to the more complicated structure of Hamiltonians of this problem. The complex behavior is similar to the overfitting problem which is encountered in standard machine learnings. In general, overfitting happens when the model parameter, in this case, $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$, are too many in number with respect to the trained points. Several approaches can be taken to address the problem. The immediate one is to increase the number of the training points, which has a severe disadvantage in the time required for the optimization. The other approach without increasing the optimization time is stopping the optimization before the exact minimum is found.

Considering these, we conducted experiments with a threshold on gradients at each iteration; optimization is stopped when the norm of the gradient goes below the threshold. Fig. 9 (c) shows the result with such threshold. It is apparent that the complex behavior is suppressed with this approach. However, this approach reduces the accuracy of the optimized output at the trained points. It seems that to recover the accuracy that is lost by this approach, it is necessary to increase the depth of the circuit. Therefore we conclude the simultaneous optimization approach does not fit as an application of NISQ devices.
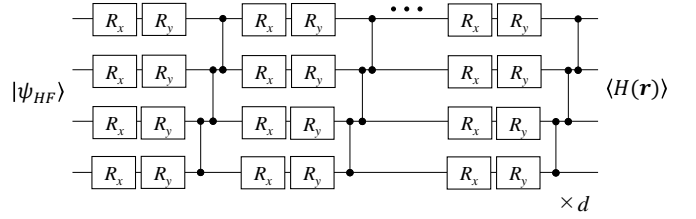

FIG. 7. Hardware efficient circuit that is used for comparison. R_x and R_y are rotation gates around x and y axis respectively. Rotation angles of each rotation gate are independent parameters to be optimized.
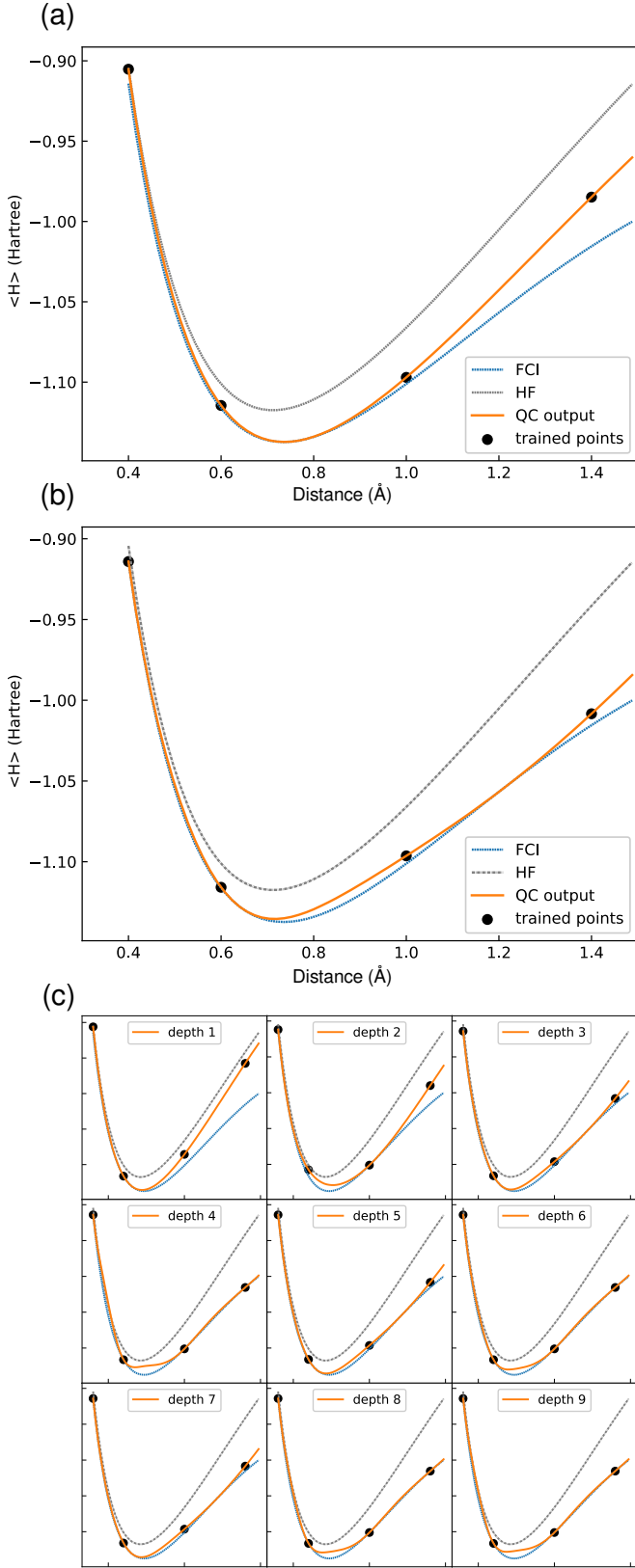
FIG. 8. (a), (b) Example of optimized outputs from the hardware efficient circuit (Fig. 7) and the Hamiltonian-alternating circuit (Fig. 1), respectively. Depth d was set to 5, and the circuits are trained with the H_2 molecular Hamiltonians of interatomic distances $r = 0.4, 0.6, 1.0,$ and 1.4 Å. The result from the Hartree-Fock (HF) calculations and the full configuration interaction (FCI), which is the exact ground state energy with the chosen basis set, are drawn for references. (c) Depth dependence of optimization result with the Hamiltonian-alternating ansatz.
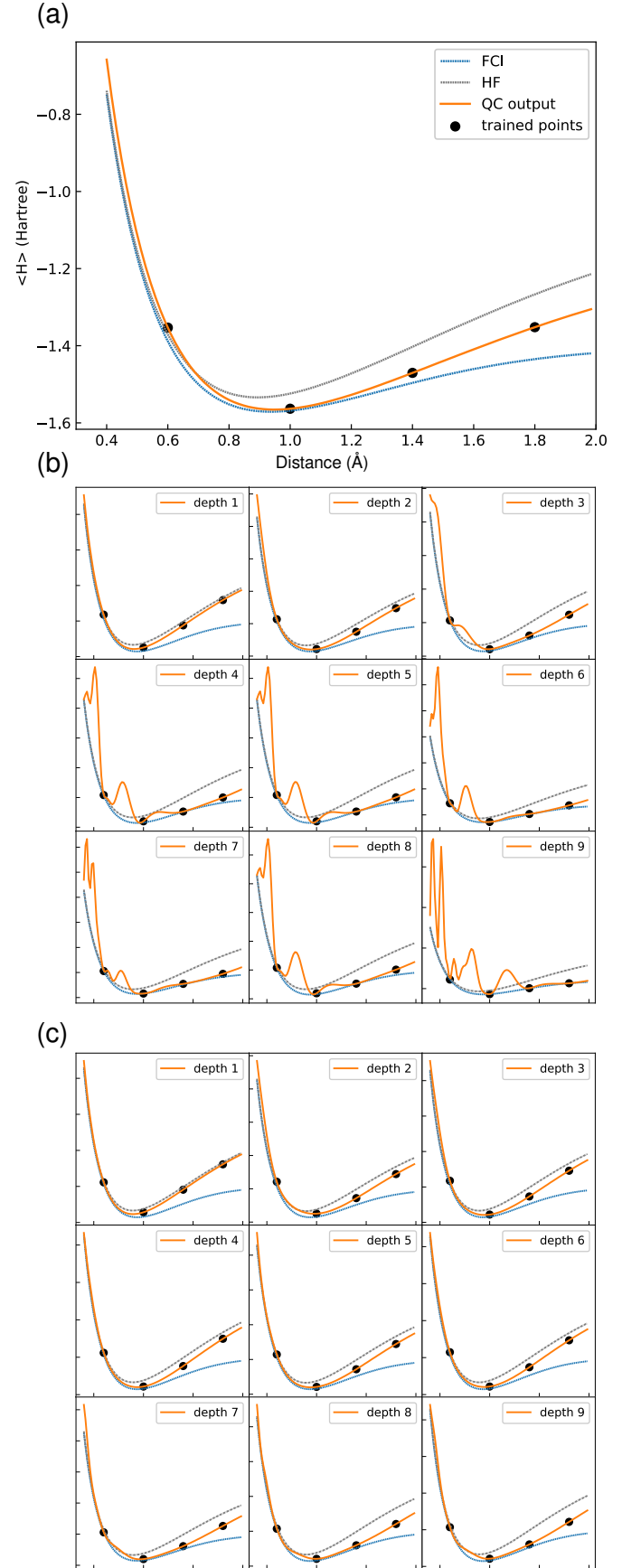

FIG. 9. Simulation results for linearly aligned H_3 molecular Hamiltonian. (a) Optimized output from the hardware efficient circuit for comparison. Depth d is set to 5. (b) Depth dependency of the optimization result with the Hamiltonian-alternating ansatz. (c) The optimization result with a threshold in gradient.

3. Discussion

We suspect the reason why this simultaneous optimization approach does not work is due to the fact that, for fixed (optimized) parameter $\boldsymbol{\theta}^*$ and $\boldsymbol{\varphi}^*$, the change in ansatz state $|\psi(\boldsymbol{x}, \boldsymbol{\theta}^*, \boldsymbol{\varphi}^*)\rangle$ with respect to the change in the parameter $\boldsymbol{x}$ is not same as in the exact ground state $|\psi_g(\boldsymbol{x})\rangle$. The ansatz $|\psi(\boldsymbol{x}, \boldsymbol{\theta}^*, \boldsymbol{\varphi}^*)\rangle$ that outputs approximate ground states with respect to different $\boldsymbol{x}$'s must have the approximately same gradient with respect to $\boldsymbol{x}$, at least around trained points $\{\boldsymbol{x}_\alpha\}$, as the exact ground state $|\psi_g(\boldsymbol{x})\rangle$. This can be a guiding instruction on constructing quantum circuits for VQE. Interpolation can be regarded as a way of realizing this concept.

[1] J. Preskill, Quantum **2**, 79 (2018).
[2] A. Peruzzo, J. McClean, P. Shadbolt, M. Yung, X. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5** (2014).
[3] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X **6**, 031007 (2016).
[4] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).
[5] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, P. Karalekas, C. B. Osborn, A. Papageorge, E. C. Peterson, G. Prawiroatmodjo, N. Rubin, C. A. Ryan, D. Scarabelli, M. Scheer, E. A. Sete, P. Sivarajah, R. S. Smith, A. Staley, N. Tezak, W. J. Zeng, A. Hudson, B. R. Johnson, M. Reagor, M. P. da Silva, and C. Rigetti, (2017), arXiv:1712.05771.
[6] V. Havlicek, A. D. Córcoles, K. Temme, A. W. Harrow, J. M. Chow, and J. M. Gambetta, Nature **567**, 209 (2018).
[7] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, Phys. Rev. X **6**, 031045 (2016).
[8] E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028.
[9] E. Farhi and A. W. Harrow, arXiv:1602.07674.
[10] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Phys. Rev. A **98**, 032309 (2018).
[11] W. Huggins, P. Patel, K. B. Whaley, and E. M. Stoudenmire, Quantum Sci. Technol. **4** (2018).
[12] E. Farhi and H. Neven, (2018), arXiv:1802.06002.
[13] M. Schuld and N. Killoran, Phys. Rev. Lett. **122**, 040504 (2019).
[14] M. Schuld, A. Bocharov, K. Svore, and N. Wiebe, (2018), arXiv:1804.00633.
[15] M. Benedetti, D. Garcia-Pintos, Y. Nam, and A. Perdomo-Ortiz, (2018), arXiv:1801.07686.
[16] M. Benedetti, E. Grant, L. Wossnig, and S. Severini, (2018), arXiv:1806.00463.
[17] K. Fujii and K. Nakajima, Phys. Rev. Appl. **8**, 024030 (2017).
[18] M. Negoro, K. Mitarai, K. Fujii, K. Nakajima, and M. Kitagawa, (2018), arXiv:1806.10910.
[19] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[20] S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin, and X. Yuan, (2018), arXiv:1808.10402.
[21] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, Annu. Rev. Phys. Chem. **62**, 185 (2011).
[22] J. T. Seeley, M. J. Richard, and P. J. Love, J. Chem. Phys. **137**, 224109 (2012).
[23] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A **92**, 042303 (2015).
[24] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys. **18**, 023023 (2016).
[25] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, (2018), arXiv:1801.01053.
[26] J. R. McClean, I. D. Kivlichan, K. J. Sung, D. S. Steiger, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, C. Huang, J. Izaac, Z. Jiang, X. Liu, M. Neeley, T. O'Brien, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, M. Steudtner, Q. Sun, W. Sun, F. Zhang, and R. Babbush, (2017), arXiv:1710.07629.
[27] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).
[28] J. Nocedal and S. Wright, *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering (Springer, New York, 2006).
[29] "Qulacs," https://github.com/qulacs/qulacs (2018).
[30] G. G. Guerreschi and M. Smelyanskiy, (2017), arXiv:1701.01450.
[31] S. McArdle, S. Endo, T. Jones, Y. Li, S. Benjamin, and X. Yuan, (2018), arXiv:1804.03023.
[32] E. Jones, T. Oliphant, P. Peterson, et al., "SciPy: Open source scientific tools for Python," http://www.scipy.org/ (2001–).