Deep variational quantum eigensolver for excited states and its application to quantum chemistry calculation of periodic materials

Kaoru Mizuta,[1,2,*] Mikiya Fujii,[3] Shigeki Fujii,[3] Kazuhide Ichikawa [3] Yutaka Imamura,[4] Yukihiro Okuno [5] and Yuya O. Nakagawa [1]

[1]*QunaSys Inc., Aqua Hakusan Building 9F, 1-13-7 Hakusan, Bunkyo, Tokyo 113-0001, Japan*
[2]*Department of Physics, Kyoto University, Kyoto 606-8502, Japan*
[3]*Technology division, Innovation Promotion Sector, Panasonic Corporation, 1006 Kadoma, Kadoma City, Osaka 571-8508, Japan*
[4]*Innovative Technology Laboratories, AGC Inc., Yokohama 230-0045, Japan*
[5]*Analysis Technology Center, FUJIFILM Corporation, Minamiashigara City, Kanagawa 250-0193, Japan*

A programmable quantum device that has a large number of qubits without fault-tolerance has emerged recently. Variational quantum eigensolver (VQE) is one of the most promising ways to utilize the computational power of such devices to solve problems in condensed matter physics and quantum chemistry. As the size of the current quantum devices is still not large for rivaling classical computers at solving practical problems, Fujii *et al.* proposed a method called "Deep VQE", which can provide the ground state of a given quantum system with the smaller number of qubits by combining the VQE and the technique of coarse graining [K. Fujii, K. Mitarai, W. Mizukami, and Y. O. Nakagawa, arXiv:2007.10917]. In this paper, we extend the original proposal of Deep VQE to obtain the excited states and apply it to quantum chemistry calculation of a periodic material, which is one of the most impactful applications of the VQE. We first propose a modified scheme to construct quantum states for coarse graining in Deep VQE to obtain the excited states. We also present a method to avoid a problem of meaningless eigenvalues in the original Deep VQE without restricting variational quantum states. Finally, we classically simulate our modified Deep VQE for quantum chemistry calculation of a periodic hydrogen chain as a typical periodic material. Our method reproduces the ground-state energy and the first-excited-state energy with the errors up to $O(1)\%$ despite the decrease in the number of qubits required for the calculation by two or four compared with the naive VQE. Our result will serve as a beacon for tackling quantum chemistry problems with classically-intractable sizes by smaller quantum devices in the near future.

I. INTRODUCTION

Noisy intermediate-scale quantum (NISQ) devices have a moderate number $[O(10) - O(100)]$ of qubits that we can control very precisely although they are not fault tolerant [1,2]. Variational quantum eigensolver (VQE), which computes the approximate ground state of quantum systems, is one of the most promising applications of the NISQ devices [3]. As the size and hardware-precision of the NISQ devices have been growing recently [4,5], extensive efforts are put into studies of the VQE, which leads to various extension of the VQE to its practical application to condensed matter physics and quantum chemistry. For instance, while the original VQE gives the approximate ground state, a bunch of methods to calculate low-energy excited states [6–11] were proposed. There are also VQE-based algorithms for calculating Green's functions [12,13], nonequilibrium dynamics [14–16], and its steady states [17] in dissipative systems, etc. As for applications to quantum chemistry [18,19], energies of electronic states of molecules were already computed experimentally by using the VQE [3,20–22]. Methods to obtain other important quantities such as the energy derivatives [23,24] and the nonadiabatic coupling [25] were proposed. Moreover, the possibility to explore periodic materials by the VQE was examined in Refs. [26–28].

While the above-mentioned methods based on the VQE will be utilized in the near future, the currently-available NISQ devices do not reach the stage to overwhelm computations by classical computers because the number of qubits and the precision of gate-operations on the qubits are still limited. To relax the requirement for the hardware of the NISQ devices, K. Fujii *et al.* recently proposed a method called deep variational quantum eigensolver (Deep VQE) [29]. Deep VQE combines the VQE and an idea of the divide-and-conquer method that is popular in quantum chemistry [30–33]. By performing the coarse graining of an original large problem based on the solutions of the VQE in smaller subsystems, Deep VQE allows us to obtain the approximate ground state of the original problem with the smaller number of qubits compared to the usual VQE.

*mizuta.kaoru.65u@st.kyoto-u.ac.jp

However, from both physical and technical aspects, it has been still nontrivial and essential whether Deep VQE can deal with low-energy properties, including excited states, which are of great interest both in physics and chemistry. The first problem for treating excited states is that coarse graining in the original Deep VQE aims to capture only the ground state by considering excitations from intersubsystem interactions but neglecting those from intrasubsystem interactions. We should carefully consider whether Deep VQE can yield low-energy excited states from the physical point of view, e.g., the type of excitations to be included in the algorithm. The second problem is a technical one that imposes strong limitations on variational quantum states employed in Deep VQE; that is, if we choose variational quantum states freely, wrong eigenvalues may be obtained because of the appearance of meaningless eigenvalues in the energy spectrum. This problem occurs not only for excited states but also for the ground state, and the use of special variational quantum states to avoid the meaningless eigenvalues is mentioned in the original Deep VQE proposal [29].

The first purpose of this paper is to elaborate the protocol of Deep VQE for correctly simulating low-energy eigenstates by solving the above-mentioned problems. Concretely, we propose a local basis set in which both excitations from intrasubsystem and intersubsystem interactions are taken into account in the procedure of the coarse graining. We show that our modified choice of the local basis set enables us to obtain more accurate low-energy eigenvalues than those obtained with the original Deep VQE by employing the perturbation theory and the quantum subspace expansion (QSE) method [6]. We also address the technical problem by giving an elaborate construction of a coarse grained Hamiltonian with additional penalty terms. The penalty terms make meaningless eigenvalues move away from the low-energy spectrum, enabling us to use any kind of variational quantum states. We provide a mathematically rigorous sufficient condition for the magnitude of the penalty terms to realize arbitrary options of variational quantum states. Finally, we numerically confirm those findings by classical simulations in a spin chain.

The second purpose is to examine the validity of Deep VQE for quantum chemistry problems, which has been raised as one of the motivations of the original Deep VQE [29]. In particular, we focus on low-energy eigenstates of periodic materials. Performing quantum chemistry calculations for periodic materials remains one of the ultimate goals of quantum chemistry and material science, but the huge computational cost hinders its realization. The use of quantum computers (possibly NISQ devices) [26–28,34] may circumvent the situation, but it still requires the large number of qubits to perform the calculation. In this regards, the coarse graining techniques such as Deep VQE are of great importance for periodic materials. We take a periodic hydrogen chain as the most straightforward example and classically simulate the performance of our modified Deep VQE by calculating its ground-state energy and its first-excited-state energy. We find a proper way of the coarse graining and obtain the low-energy eigenvalues with an error up to $O(1)\%$ despite reducing the number of qubits by up to four. Our results will serve as a beacon for simulating low-energy properties of a variety of materials with smaller quantum devices in the coming NISQ era.
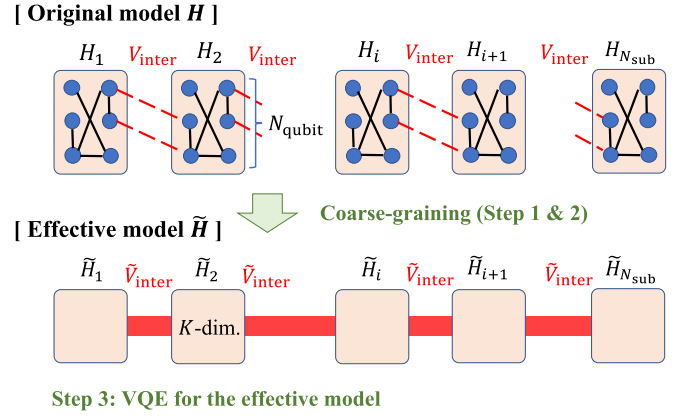

FIG. 1. Schematic picture of Deep VQE. The details are described in Sec. II A.

This paper is organized as follows. In Sec. II, we review Deep VQE and its pros and cons to describe our study's motivation in detail. Sections III and IV provide the main results of our research. In Sec. III, we formulate Deep VQE for low-energy excited states and show numerical results for simple spin systems. We discuss the validity of our proposed method with relating to other methods such as QSE. In Sec. IV, we apply our modified Deep VQE to a periodic hydrogen chain and show numerical results for the ground-state energy the first-excited-state energy. We conclude this paper in Sec. V.

II. PRELIMINARIES

In this section, we review the algorithm of the original Deep VQE proposed in Ref. [29] and discuss its pros and cons to motivate our study. Deep VQE combines the VQE [3] and the divide-and-conquer method [30,31], with which we can find the ground state with fewer qubits than those required when using the usual VQE.

A. Protocol of Deep VQE

We consider a N_{tot}-qubit quantum system and split it into N_{sub} subsystems, where each subsystem is composed of $N_{\mathrm{qubit}} = N_{\mathrm{tot}}/N_{\mathrm{sub}}$ qubits (see Fig. 1). Then, we assume that the Hamiltonian of the system is written in the following form:

$$H = \sum_{i=1}^{N_{\mathrm{sub}}} H_i + V_{\mathrm{inter}}, \quad V_{\mathrm{inter}} = \sum_{i \neq j}^{N_{\mathrm{sub}}} \sum_{\alpha} v_{ij}^{\alpha} V_i^{\alpha} \otimes W_j^{\alpha}, \quad (1)$$

where H_i represents an intrasubsystem Hamiltonian of the ith subsystem. The intersubsystem interactions are described by V_{inter}, in which V_i^{α} and W_j^{α} are operators acting on the ith and the jth subsystems, respectively. The index α enumerates such interaction terms between the subsystems i and j. The above equation only includes intersubsystem interactions over two subsystems, but the extension to the cases where intersubsystem interactions involves more than three subsystems is straightforward. Under this setup, Deep VQE is composed of the following three steps.

Step 1: Perform VQE for each subsystem.

We obtain a set of local ground states $\{|\psi_0\rangle_i\}_{i=1}^{N_{\mathrm{sub}}}$, where $|\psi_0\rangle_i$ represents the ground state of the ith subsystem Hamil-

tonian H_i, by the ordinary VQE. This can be completed by minimizing the cost function $\langle\psi(\vec{\theta})|H_i|\psi(\vec{\theta})\rangle$ with a variational quantum state $|\psi(\theta)\rangle = U(\vec{\theta})|0\rangle$ and setting $|\psi_0\rangle_i = |\psi(\vec{\theta}_i^*)\rangle$, where $\vec{\theta}_i^*$ is the optimal parameter set for ith subsystem. This step requires a N_{qubit}-qubit quantum computer.

Step 2: Construct an effective model in the restricted Hilbert space (coarse graining).

We use a divide-and-conquer method here. For each subsystem, we choose a set of local excitation operators $\{P_k^{(i)}\}_{k=1}^K$ with $P_1^{(i)} = I$ (identity operator). Let $\tilde{\mathcal{H}}_i$ denote the K-dimensional local Hilbert space spanned by $\{P_k^{(i)}|\psi_0\rangle_i\}_{k=1}^K$. We focus on the restricted Hilbert space $\tilde{\mathcal{H}} = \otimes_i^{N_{\text{sub}}}\tilde{\mathcal{H}}_i$, and construct the effective Hamiltonian after coarse graining by

$$\tilde{H} = H|_{\tilde{\mathcal{H}}} = \sum_{i=1}^{N_{\text{sub}}}\tilde{H}_i + \sum_{i\neq j}^{N_{\text{sub}}}\sum_\alpha v_{ij}^\alpha \tilde{V}_i^\alpha \otimes \tilde{W}_j^\alpha, \qquad (2)$$

which is reminiscent of the internally contracted multireference-configuration interaction method in quantum chemistry [35,36]. Here, we define $\tilde{H}_i = H_i|_{\tilde{\mathcal{H}}_i}$, $\tilde{V}_i^\alpha = V_i^\alpha|_{\tilde{\mathcal{H}}_i}$, and $\tilde{W}_j^\alpha = W_j^\alpha|_{\tilde{\mathcal{H}}_j}$. We should compute $K \times K$ matrix representations of $\tilde{H}_i$, $\tilde{V}_i^\alpha$ and $\tilde{W}_j^\alpha$ by using the result of Step 1 and the N_{qubit}-qubit quantum computer. Recalling that the local basis $\{P_k^{(i)}|\psi_0\rangle_i\}_{k=1}^K$ is not orthonormal, we first construct an orthonormal basis $\{|\phi_k\rangle_i\}_{k=1}^K$ by the Gram-Schmidt method. Concretely, we compute a $K \times K$ matrix $G^{(i)}$ representing the inner product,

$$G_{kk'}^{(i)} = {}_i\langle\psi_0|P_k^{(i)\dagger}P_l^{(i)}|\psi_0\rangle_i, \quad k, k' = 1, 2, \ldots, K, \qquad (3)$$

by N_{qubit}-qubit quantum devices. With classical computation using the matrix $G^{(i)}$, we can obtain a $K \times K$ matrix $S^{(i)}$ so that the set of states, given by

$$|\phi_k\rangle_i = \sum_{k'=1}^K S_{k'k}^{(i)}\big(P_{k'}^{(i)}|\psi_0\rangle_i\big), \quad k = 1, 2, \ldots, K, \qquad (4)$$

can provide an orthonormal basis of $\tilde{\mathcal{H}}_i$ (see Appendix A for detail). We then obtain the matrix elements

$$\begin{aligned}(\tilde{A})_{kl} &= {}_i\langle\phi_k|A|\phi_l\rangle_i \\ &= \sum_{k',l'} S_{k'k}^{(i)*}S_{l'l}^{(i)}{}_i\langle\psi_0|P_{k'}^{(i),\dagger}AP_{l'}^{(i)}|\psi_0\rangle_i\end{aligned} \qquad (5)$$

by using the N_{qubit}-qubit quantum device of to evaluate the expectation values of $P_{k'}^{(i),\dagger}AP_{l'}^{(i)}$ for $A = H_i$, V_i^α or W_i^α.

In this step, there are several options for the local excitation operators $\{P_k^{(i)}\}$, or equivalently the local basis set $\{|\phi_k\rangle_i\}$. In Ref. [29], a series of local operators that occur in the intersubsystem interactions and nontrivially act on the boundary sites of subsystems are adopted as $\{P_k^{(i)}\}$ for evaluating the ground state of spin systems.

Step 3: Perform VQE for the effective model.

We finally compute the ground state of the effective Hamiltonian $\tilde{H}$ by using VQE, which is expected to approximately reproduce that of the original Hamiltonian H. To perform VQE for $\tilde{H}$ on quantum computers composed of qubits, we should choose an integer $N_{\text{eff}} = \lceil\log_2 K\rceil$ to embed $\tilde{H}$ into a Hilbert space of N_{eff} qubits. We add auxiliary dimensions to

each local Hilbert space $\tilde{\mathcal{H}}_i$ and replace the effective Hamiltonian by

$$\tilde{H}_{\text{eff}} = \sum_{i=1}^{N_{\text{sub}}}\tilde{H}_{i,\text{eff}} + \sum_{i\neq j}^{N_{\text{sub}}}\sum_\alpha v_{ij}^\alpha \tilde{V}_{i,\text{eff}}^\alpha \otimes \tilde{W}_{j,\text{eff}}^\alpha, \qquad (6)$$

$$\tilde{H}_{i,\text{eff}} = \tilde{H}_i \oplus 0_{2^{N_{\text{eff}}}-K}, \qquad (7)$$

$$\tilde{V}_{i,\text{eff}}^\alpha = \tilde{V}_i^\alpha \oplus 0_{2^{N_{\text{eff}}}-K}, \quad \tilde{W}_{j,\text{eff}}^\alpha = \tilde{W}_j^\alpha \oplus 0_{2^{N_{\text{eff}}}-K}, \qquad (8)$$

where 0_M represents a $M \times M$ zero matrix. The replaced effective Hamiltonian $\tilde{H}_{\text{eff}}$ is defined on $N_{\text{sub}} \times N_{\text{eff}}$ qubits, and we can perform VQE for it. By carefully choosing a variational quantum circuit $V(\vec{\theta})$ so that $V(\vec{\theta})|0\rangle$ has no components out of the $K^{N_{\text{sub}}}$-dimensional subspace corresponding to $\tilde{\mathcal{H}}$, we obtain the ground-state energy by minimizing the cost function $\langle 0|V^\dagger(\vec{\theta})\tilde{H}_{\text{eff}}V(\vec{\theta})|0\rangle$. This Step 3 requires quantum computation on $(N_{\text{sub}} \times N_{\text{eff}}) = N_{\text{sub}}\lceil\log_2 K\rceil$ qubits.

Steps 1, 2, and 3 are crucial procedures of Deep VQE. The number of qubits required for the whole algorithm is given by

$$N_{\text{req}} = \max(N_{\text{qubit}}, N_{\text{sub}} \times N_{\text{eff}}). \qquad (9)$$

When K is not taken as so large $[\sim O(e^{N_{\text{qubit}}})]$, N_{req} is small compared to the required number of qubits for simulating the original system, $N_{\text{tot}} = N_{\text{sub}}N_{\text{qubit}}$. In the original proposal [29], the repetition of Steps 1–3 are also proposed to further reduce the number of qubits.

B. Motivation of our study

The pros of Deep VQE is that we can calculate the ground state with the smaller number of qubits. The key to reducing the required qubits is the choice of local excitation operators $\{P_k^{(i)}\}_{k=1}^K$ in Step 2, in which we discard some information of the whole Hilbert space. As is usual with the divide-and-conquer methods, it is argued in Ref. [29] that the resulting restricted Hilbert space can describe the ground state of the original Hamiltonian well as long as it is low-entangled and we can choose the proper local basis $\{P_k^{(i)}\}_{k=1}^K$.

We motivate our study by raising several points for the original Deep VQE proposal. The first point is whether it is possible to describe the low-energy eigenstates of the original Hamiltonian in the restricted Hilbert space, or simply stating, the applicability of Deep VQE to the excited states. It is nontrivial whether the same local basis for the ground state works also for the excited states. The second point lies in Step 3. When we replace the effective Hamiltonian $\tilde{H}$ by $\tilde{H}_{\text{eff}}$ with inserting auxiliary dimensions, some meaningless (or artificial) eigenvalues appear in the spectrum of $\tilde{H}_{\text{eff}}$. As we explicitly show in Sec. III B, this can alter both the ground-state energy and the low-excited-state energies of $\tilde{H}_{\text{eff}}$ from those of $\tilde{H}$. This is the reason why we restrict the variational quantum states in Step 3 following Ref. [29]. In practice, construction of such variational quantum states is not so straightforward and it is desirable to develop a way to use arbitrary variational quantum states for searching low-energy eigenstates of $\tilde{H}_{\text{eff}}$. The third point relates to the applicability of Deep VQE to fermionic systems such as quantum chemistry problems. The performance of Deep VQE in fermionic systems has not been explored so far, where quantum states mapped to qubit

systems can be more entangled in general because of the nonlocal terms in the mapped Hamiltonian.

This paper aims to construct the modified Deep VQE protocol for low-energy eigenstates by solving the above problems and to examine its practical use for chemistry problems, particularly the simulation of periodic materials. The simulation of periodic materials in *ab initio* level, which always lies at the center of condensed matter physics, requires many qubits to predict accurate results in the thermodynamic limit. Thus, periodic materials are expected to be a good target of the modified Deep VQE, in that we would like to obtain both their ground state and the low-energy-excited states with the smaller number of qubits.

III. MAIN RESULT: DEEP VQE FOR LOW-ENERGY PHYSICS

This section provides the first half of the main results. As discussed in the previous section, the original Deep VQE has some nontrivial problems in calculating low-energy eigenstates. In Secs. III A and III B, we propose the modified Deep VQE protocol for excited states that solves these problems. In Sec. III D, we provide a numerical example to confirm the validity of our modified Deep VQE.

A. Choice of local basis for obtaining accurate excited-states

Here, we provide the modified way of choosing the local basis in Step 2 of Deep VQE so that low energy eigenstates of the original Hamiltonian can be well represented in the restricted Hilbert space.

Before proposing our modified choice of the local basis to obtain accurate excited states, let us recall the local basis of the original Deep VQE and argue that it may fail for the excited states. The local basis $\{P_k^{(i)}\}_k$ in the original Deep VQE is chosen as a set of Pauli operators contained in the intersubsystem interactions (an explicit example can be found in Sec. III D). Assuming that H_i has a unique ground state $|\psi_0\rangle_i$ and that the VQE in Step 1 of Deep VQE is accurate enough, it follows that $|\Psi_0\rangle = \bigotimes_i |\psi_0\rangle_i$ is the ground state of $H_{\text{intra}} = \sum_i H_i$. Regarding H_{intra} and V_{inter} as an unperturbed Hamiltonian and a perturbation, respectively, Deep VQE always gives the ground-state energy more accurate than that of the first-order perturbation theory, $\langle\Psi_0|H|\Psi_0\rangle$. This is because the restricted Hilbert space $\tilde{\mathcal{H}}$ includes $|\Psi_0\rangle$ by construction. Although neither the original Deep VQE overwhelms higher-order perturbation theories nor vice versa, they commonly capture local excitations from the unperturbed ground state $|\Psi_0\rangle$ invoked by $\{P_k^{(i)}\}$ (or V_{inter}), which thereby validate the accuracy of the original Deep VQE for the ground state. On the other hand, the restricted Hilbert space of the original Deep VQE cannot reproduce the low-energy-excited-states eigenvalues even at the first-order perturbation level. The *n*th excited states $|\Psi_n\rangle$ of the unperturbed Hamiltonian H_{intra} is not generated by applying only the excitation operators stemming from the intersubsystem interactions to $|\Psi_0\rangle$. Hence, the choice of the local basis of the original Deep VQE will fail to produce accurate excited-state energies (see also the numerical results in Sec. III D).

We propose an operator set $\{P_k^{(i)}\}_{k=1}^K$ for properly representing low-energy excited states as the modified Deep VQE protocol. The main idea is to take into account not only excitations by intersubsystem interactions V_{inter} but also those by intrasubsystem interactions H_{intra}. Namely, the simplest choice of $\{P_k^{(i)}\}_{k=1}^K$ is

$$P_k^{(i)} \in \mathcal{W} := \{I\} \cup \left(\bigcup_{j=1}^{N_{\text{qubit}}} \{X_j, Y_j, Z_j\}\right), \tag{10}$$

composed of the identity operator and Pauli operators on each sites of the subsystem *i*. The local Hilbert space dimension $K = 3N_{\text{qubit}} + 1$ is still much smaller than the original one $2^{N_{\text{qubit}}}$, meaning the decrease of qubits required for simulation.

We can show the validity of the modified local basis generated by $\mathcal{W}$ for Deep VQE of the excited states under certain assumptions. Concretely, we can show that our modified Deep VQE yields more accurate low-excited-state energies than the first-order perturbation theory of V_{inter}. We assume that the low-lying excited states of the *i*th subsystem, $|\psi_n\rangle_i$ ($n = 1, 2, \ldots$), can be well described by the restricted Hilbert space for the subsystem, $\{P_k^{(i)} |\psi_0\rangle_i \,|\, P_k^{(i)} \in \mathcal{W}\}_{k=1}^K$. This assumption is justified when the excited states of the subsystem are not so entangled and approximately generated only from the first-order excitation of the terms included in H_i. This assumption can also be rephrased that the QSE method [6] works well in the *i*th subsystem, where a similar restricted Hilbert space is constructed and the Hamiltonian within that space is solved classically. To see how the assumption leads to the validity of our choice of the local basis, let us consider the first-excited states of the original Hamiltonian composed of N_{sub} identical subsystems. The degenerated first-excited states of the unperturbed Hamiltonian H_{intra},

$$\left|\Psi_1^{(i)}\right\rangle = |\psi_1\rangle_i \otimes \left(\bigotimes_{j:\, j\neq i}^{N_{\text{sub}}} |\psi_0\rangle_j\right), \tag{11}$$

are included in the subspace $\tilde{\mathcal{H}}$ because of the assumption. The perturbation theory for V_{inter} tells us that the subspace $\mathcal{S}_1 = \text{span}(\{|\Psi_1^{(i)}\rangle \,|\, i = 1, \ldots, N_{\text{sub}}\})$ can reproduce the approximate first-excited-state energy as

$$E_1(\text{local}) = \min\left\{\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \,\Big|\, |\Psi\rangle \in \mathcal{S}_1\right\}. \tag{12}$$

Because of the inclusion $\mathcal{S}_1 \subset \tilde{\mathcal{H}}$ by construction, we can conclude that the modified Deep VQE always provides more accurate low-excited energies than that of the first-order perturbation.

The higher-order perturbation theories involve an infinite series of the unperturbed eigenstates $|\Psi_k\rangle$, thereby making it difficult to obtain rigorous relation with the Deep VQE. However, we can intuitively give some correspondence as well as the ground states. For instance, in the second-order perturbation theory without degeneracy, the unperturbed eigenstates $|\Psi_k\rangle$ with the large value of $|\langle\Psi_k|V_{\text{inter}}|\Psi_n\rangle / (\langle\Psi_n|H_{\text{intra}}|\Psi_n\rangle - \langle\Psi_k|H_{\text{intra}}|\Psi_k\rangle)|$ have principal contributions to the *n*th low-energy eigenvalues. Such states $|\Psi_k\rangle$, having the low-energy under H_{intra} and the connection via local excitations from V_{inter}, are approximately included in the restricted Hilbert

space of our method. Thus, our method is expected to provide approximate low-energy eigenvalues as well as the higher-order perturbation theories. In fact, in the later numerical simulations, we will see that the modified Deep VQE comparably outperforms the first-order perturbation result, supporting this expectation.

We note another choice of the local basis set to evaluate the excited states. One possible choice is to include the states like $\{P_k^{(i)} |\psi_n\rangle_i\}$, where $|\psi_n\rangle_i$ is the subsystem excited eigenstate. Since the restricted Hilbert space after the coarse graining directly includes $|\Psi_1^{(i)}\rangle$ [Eq. (11)], we can expect that this choice also gives approximate excited states. The caveat of choice is that we should employ a proper algorithms to calculate the excited states $|\psi_n\rangle_i$ in the subsystems to construct evaluate the matrix elements of $\tilde{H}$. We discuss the details of this choice of the local basis in Appendix B.

To summarize, in the modified Deep VQE for calculating low-energy eigenstates, we prepare local excitation operators in Step 2 as $\mathcal{W}$ [Eq. (10)]. This corresponds to taking into account excitations by both intrasubsystem and intersubsystem terms.

B. Construction of effective qubit model with penalty

Next, we provide the modified way of constructing the effective qubit model in Step 3 of Deep VQE so that we can adopt any variational quantum circuit.

Let us first see the problem of meaningless eigenvalues in the original Deep VQE by taking simple examples. In Step 3 of Deep VQE, we insert additional dimensions to the effective Hamiltonian $\tilde{H}$ obtained in Step 2 and get $\tilde{H}_{\rm eff}$ defined on qubits. This introduces some meaningless eigenvalues and even changes the ground-state energy and the low-energy eigenvalues. For instance, we consider a two-qubit Hamiltonian

$$\tilde{H}_a = Z \otimes I + I \otimes P^{z+} + P^{x+} \otimes P^{x+}, \tag{13}$$

where $P^{x+} = (I + X)/2$ [$P^{z+} = (I + Z)/2$] is a projection to an up-spin state in x direction [in z direction]. We add a dimension by one in the same way as Eqs. (6)–(8):

$$\tilde{H}_{a,{\rm eff}} = \tilde{Z} \otimes \tilde{I} + \tilde{I} \otimes \tilde{P}^{z+} + \tilde{P}^{x+} \otimes \tilde{P}^{x+} \tag{14}$$

with $\tilde{Z} = Z \oplus (0)$, $\tilde{P}^{\alpha+} = P^{\alpha+} \oplus (0)$, and $\tilde{I} = I \oplus (1)$ (note that the dimension itself is not essential here). While the energy spectrum of $\tilde{H}_a$ is $E \in [-0.836, 0.201, 1.245, 2390]$, that of $\tilde{H}_{a,{\rm eff}}$ is $[-1, -0.836, 0, 0, 0.201, 1, 1, 1.245, 2.390]$, changing the ground state. As well, another example is $\tilde{H}_b = 0.2Z \otimes I + 0.7I \otimes Z + 0.3X \otimes X$. It originally has the first-excited-state energy $E_1 = -0.583$, but that of $\tilde{H}_{b,{\rm eff}}$ is -0.7.

These meaningless eigenvalues come from inserting the identity by $\tilde{I} = I \oplus (1)$, which gives meaningless energies to states in the auxiliary dimensions. We note that inserting the identity by $\tilde{I}' = I \oplus 0_M$ generates no meaningless eigenvalues other than zero. However, the operator $\tilde{I}'$ is decomposed into a number of terms composed of many Pauli Z operators. This makes $\tilde{H}_{\rm eff}$ have a lot of nonlocal Pauli operators among the subsystems, so the extension by $\tilde{I} = I \oplus 0_M$ is not practical for performing the VQE for the effective Hamiltonian in Step 3. In the original Deep VQE, to avoid the change of the ground-state energy (and the low-energy eigenvalues), the variational

quantum circuits should be chosen so that all the components in the additional Hilbert space become zero.

We propose an alternative way to construct the effective qubit model so that $\tilde{H}$ and $\tilde{H}_{\rm eff}$ have the same low-energy spectrum, thereby enabling the use of any variational quantum circuit. Concretely, we specify the terms in the auxiliary dimensions as follows:

$$\tilde{H}_{\rm eff} = \sum_{i=1}^{N_{\rm sub}} \tilde{H}_{i,{\rm eff}} + \sum_{i \neq j}^{N_{\rm sub}} \sum_\alpha v_{ij}^\alpha \tilde{V}_{i,{\rm eff}}^\alpha \otimes \tilde{W}_{j,{\rm eff}}^\alpha, \tag{15}$$

$$\tilde{H}_{i,{\rm eff}} = \tilde{H}_i \oplus \lambda_i I_{2^{N_{\rm eff}} - K}, \tag{16}$$

$$\tilde{V}_{i,{\rm eff}}^\alpha = \tilde{V}_i^\alpha \oplus 0_{2^{N_{\rm eff}} - K}, \quad \tilde{W}_{j,{\rm eff}}^\alpha = \tilde{W}_j^\alpha \oplus 0_{2^{N_{\rm eff}} - K}. \tag{17}$$

The difference from Eq. (6) is the insertion of $\lambda_i I_{2^{N_{\rm eff}} - K}$ in $\tilde{H}_{i,{\rm eff}}$ with $\lambda_i > 0$, corresponding to a penalty term to components in auxiliary dimensions. For $\tilde{H}$ and $\tilde{H}_{\rm eff}$ having the same low-energy spectrum, we should choose a proper set $\{\lambda_i\}_{i=1}^{N_{\rm sub}}$. We have derived the following proposition, which gives a mathematically rigorous sufficient condition on the choice of λ_i.

Proposition. Let $E_n(H)$ denote the nth smallest eigenvalue of a Hamiltonian H with the dimension d ($n = 0, 1, 2, \ldots, d - 1$). When we choose $\{\lambda_i\}_i$ in Eq. (16) to satisfy

$$\lambda_i > e_{\tilde{H}}(i) + E_n(\tilde{H}) - E_0(\tilde{H}), \tag{18}$$

the spectrum of $\tilde{H}$ coincides that of $\tilde{H}_{\rm eff}$, defined by Eqs. (15)-(17), up to the nth smallest eigenvalue:

$$E_m(\tilde{H}) = E_m(\tilde{H}_{\rm eff}) \quad {\rm for} \quad m = 0, 1, \ldots, n. \tag{19}$$

Here, $e_{\tilde{H}}$ is the value called extensiveness of the ith subsystem under the Hamiltonian $\tilde{H}$, representing the maximal energy of the ith subsystem (see Appendix C for the rigorous definition). When the Hamiltonian $\tilde{H}$ is given by Eq. (2), the extensiveness is given by

$$e_{\tilde{H}}(i) = ||\tilde{H}_i||_{\rm op} + \sum_{j,k:\{j,k\}\ni i} \sum_\alpha |v_{jk}^\alpha| \cdot ||\tilde{V}_j^\alpha||_{\rm op} \cdot ||\tilde{W}_k^\alpha||_{\rm op}. \tag{20}$$

We give a proof for this proposition in Appendix C.

We discuss the implication of this proposition. When we calculate the ground state by our modified Deep VQE with Eqs. (15)–(17) and choose λ_i larger than $e_{\tilde{H}}(i)$, Eq. (19) indicates that we can find the ground state of $\tilde{H}$ by searching that of $\tilde{H}_{\rm eff}$ in the whole Hilbert space of $2^{N_{\rm eff}}$ dimension. In other words, we can exploit any variational quantum circuits for VQE on the extended Hamiltonian $\tilde{H}_{\rm eff}$. Similarly, when we calculate the excited states of $\tilde{H}_{\rm eff}$ up to nth level and choose λ_i larger than $e_{\tilde{H}}(i) + E_n(\tilde{H}) - E_0(\tilde{H})$, the spectra of $\tilde{H}_{\rm eff}$ and $\tilde{H}$ coincide and we can use an arbitrary variational quantum circuit.

The value of $e_{\tilde{H}}(i)$ can be calculated by classical computers as a sum of $O(N_{\rm sub} \cdot K^3)$ terms since the minimal or maximal eigenvalue of $K \times K$ matrices is required for evaluating the operator norm. On the other hand, the energy gap $E_n(\tilde{H}) - E_0(\tilde{H})$ is not known *a priori* before performing the VQE for the excited states of $\tilde{H}_{\rm eff}$. In practice, we can obtain rough and typical energy scale $\Delta_n^* \sim E_n(\tilde{H}) - E_0(\tilde{H})$

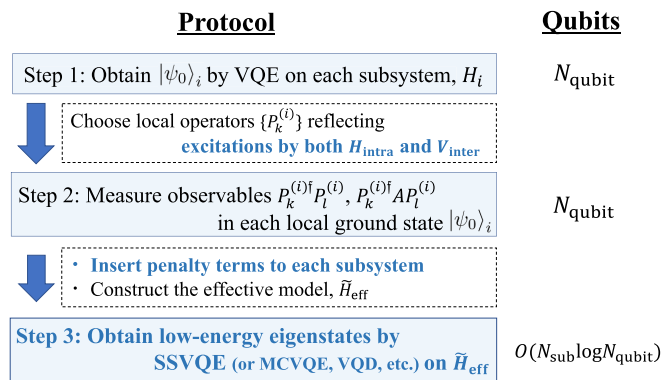| Protocol	Qubits
Step 1: Obtain $|\psi_0\rangle_i$ by VQE on each subsystem, H_i	N_{qubit}
⬇ Choose local operators $\{P_k^{(i)}\}$ reflecting **excitations by both H_{intra} and V_{inter}**	
Step 2: Measure observables $P_k^{(i)\dagger}P_l^{(i)}$, $P_k^{(i)\dagger}AP_l^{(i)}$ in each local ground state $|\psi_0\rangle_i$	N_{qubit}
⬇ • **Insert penalty terms to each subsystem** • Construct the effective model, $\tilde{H}_{eff}$	
Step 3: Obtain low-energy eigenstates by SSVQE (or MCVQE, VQD, etc.) on $\tilde{H}_{eff}$	$O(N_{sub}\log N_{qubit})$

FIG. 2. Protocol of the Deep VQE modified for low-lying eigenstates. The blue bold parts are the differences compared to the original Deep VQE [29], which deals with the ground state. "Qubits" gives the typical size of NISQ devices required for each step.

with other computationally-light methods such as a perturbation theory and a mean-field theory. With replacing the gap $E_n(\tilde{H}) - E_0(\tilde{H})$ in Eq. (18) by a certain value larger than the estimated gap Δ^*, we can safely search the low-energy excited states by VQE on $\tilde{H}_{eff}$ with any variational quantum circuit.

In short, by constructing the effective qubit model $\tilde{H}_{eff}$ as Eqs. (15)–(17) and setting λ_i to satisfy Eq. (18), we can employ any variational quantum state to run algorithms in the literature [6–11] to find excited states of $\tilde{H}_{eff}$, which are ensured to be identical to those of $\tilde{H}$. We also note that introducing the energy shift λ_i hardly affects the number of measurements. The energy shift terms $0_K \oplus \lambda_i I_{2^{N_{eff}}-K}$, composed of ZZ-type Pauli operators, have at-most $2^{N_{eff}}$ Pauli terms. This is much less than $O(4^{N_{eff}})$, required for the effective Hamiltonian $\tilde{H}_i \oplus 0_{2^{N_{eff}}-K}$.

C. Protocol and number of measurements

Here, we summarize the protocol of the modified Deep VQE for low-lying eigenstates and the number of measurements in each step. Figure 2 shows a protocol on NISQ devices required for each step.

Let us consider a simple Hamiltonian H involving at-most two-body interactions. In Step 1, the VQEs on subsystems require measurements of $O(N_{sub}N_{qubit}^2)$ terms in H_{intra}. In Step 2, we pick up $\{P_k^{(i)}\}$, which reflects the local excitations from H_{intra} and V_{inter}. For a simple choice of $\{P_k^{(i)}\}$ given by Eq. (10) with $K \sim N_{qubit}$, which captures linear excitations, we should measure $P_k^{(i)\dagger}P_l^{(i)}$ and $P_k^{(i)\dagger}AP_l^{(i)}$ in $|\psi_0\rangle_i$ [see Eqs. (3) and (5)]. They involve at-most $O(N_{sub}^2N_{qubit}^3)$ Pauli terms. In Step 3, we analyze the effective Hamiltonian $\tilde{H}_{eff}$ by elaborated VQE for low-lying eigenstates, such as the subspace-search VQE (SSVQE) [7], the multistate-contracted VQE (MCVQE) [8], and the variational quantum deflation (VQD) [9]. Since each of $\tilde{H}_i$, $\tilde{V}_i^\alpha$, $\tilde{W}_i^\alpha$ is decomposed into at-most $4^{N_{eff}} \sim K^2$ Pauli terms, we should measure $O(N_{sub}^2N_{qubit}^4)$ terms in $\tilde{H}_{eff}$. We note that introducing the penalty term λ_i by Eq. (16) does not severely increase the cost of measurements. While the number of measurements increases compared to the conventional VQE for the whole system, given by $O(N^2) = O(N_{sub}^2N_{qubit}^2)$,

the size of quantum devices decreases to $O(N_{sub}\log N_{qubit})$ in the modified Deep VQE.

In the numerical simulations below, we employ SSVQE in Step 3 due to its simplicity, and we give detailed description of its algorithm in Appendix D.

D. Example: Spin chain

Based on the modified Deep VQE protocol in the previous subsections, we numerically examine its validity by exemplifying a simple spin model. We consider a one-dimensional antiferromagnetic (AFM) Heisenberg model of N_{tot}-site,

$$H = \sum_{i=1}^{N_{tot}-1}(X_iX_{i+1} + Y_iY_{i+1} + Z_iZ_{i+1}), \quad (21)$$

under the open boundary condition. We split the system into N_{sub} subsystems, each of which is composed of neighboring N_{qubit} qubits ($N_{tot} = N_{sub} \times N_{qubit}$). Then, the intersubsystem interactions involve qubits only at the boundaries of the subsystems.

We introduce two different sets of local excitation operators to compare the modified Deep VQE with the original Deep VQE. The first one is

$$\mathcal{W}_1 = \{I\} \cup \{X_i, Y_i, Z_i | \, i \text{ at the boundary of subsystem}\}. \quad (22)$$

We note that we have not written the index of the subsystem, but we define $\mathcal{W}_1$ for each subsystem. An excitation operator $P_k \in \mathcal{W}_1$ only acts on the boundaries of subsystems and is relevant to the intersubsystem interactions, so the choice of local basis based on $\mathcal{W}_1$ correspond to the original Deep VQE. The dimension of the local Hilbert space K is 4 (for subsystems at the edges) or 7 (for subsystems in the bulk). The second choice of local excitation operators is

$$\mathcal{W}_2 = \{I\} \cup \{X_i, Y_i, Z_i | \, i \in \Lambda'\}, \quad (23)$$

where Λ' is a set of qubits in a subsystem except for the right edge. This choice also captures any single-spin excitation $(\alpha X_i + \beta Y_i + \gamma Z_i)|\psi_0\rangle$ $(\alpha, \beta, \gamma \in \mathbb{C})$, emerging from a generic local extensive intrasubsystem Hamiltonian H_i as well as those of V_{inter}, thereby corresponding to the choice in our modified Deep VQE. We omit qubits at the right edge because of the SU(2) symmetry in the AFM model; for the ground state $|\psi_0\rangle$ of each subsystems, it holds $(\sum_{j\in\text{subsystem}} X_j)|\psi_0\rangle = (\sum_{j\in\text{subsystem}} Y_j)|\psi_0\rangle = (\sum_{j\in\text{subsystem}} Z_j)|\psi_0\rangle = 0$. One of the states in $\{X_j|\psi_0\rangle | \, j \text{ in subsystem}\}$ is not linearly independent from the others, and the same relations for Y and Z also hold. We do not include the Pauli operators (X, Y, Z) at the right edge in $\mathcal{W}_2$ due to this fact. The local Hilbert space dimension K is $3N_{qubit} - 2$ in the case of $\mathcal{W}_2$.

We note on the insertion of auxiliary dimensions in this model analyzed in Sec. III B. We can roughly estimate the extensiveness $e_{\tilde{H}}(i)$ for each subsystem. In the case of $N_{sub} = 2$ and $N_{qubit} = 4$, we numerically confirm that $||\tilde{H}_i||_{op}$ coincides with the absolute value of the ground-state energy for each subsystem. The operator norm of the intersubsystem interactions is at most $2(||X_i||_{op}||X_{i+1}||_{op} + ||Y_i||_{op}||Y_{i+1}||_{op} + ||Z_i||_{op}||Z_{i+1}||_{op}) = 6$, so we obtain $e_{\tilde{H}}(i) \leqslant$

TABLE I. Numerical results for low-energy eigenvalues E_n ($n = 0, 1$) of a one-dimensional AFM Heisenberg model (21). The excitation operator set $\{P_k^{(i)}\}$ is taken as $\mathcal{W}_1$ (by $\mathcal{W}_2$) for the original Deep VQE (for our modified Deep VQE). The values E_n(DeepVQE) and E_n(ED) represent the results of Deep VQE simulation and those of exact diagonalization of the original Hamiltonian H. The values E_n(Effective) are obtained by substituting ED for the VQE in Step 3 of Deep VQE, which give the best performance of Deep VQE in theory. The left values and the right values of E_n(Effective) and E_n(DeepVQE) are the obtained energies and their relative errors from E_n(ED), respectively. The truncation rate TR is given by Eq. (24). We also show the number of qubits required to run Deep VQE by N_{req}.

	$\{P_k^{(i)}\}_k$	E_0(Effective)	E_0(Deep VQE)	E_0(ED)	E_1(Effective)	E_1(Deep VQE)	E_1(ED)	TR	N_{req}
$N_{\text{sub}} = 2$	$\mathcal{W}_1$	$-13.445/0.4\%$	$-13.445/0.4\%$	-13.500	$-11.169/6.4\%$	$-11.169/6.4\%$	-11.929	6.3×10^{-2}	4
$N_{\text{qubit}} = 4$	$\mathcal{W}_2$	$-13.497/0.02\%$	$-13.488/0.09\%$		$-11.882/0.4\%$	$-11.863/0.6\%$		3.9×10^{-1}	8
$N_{\text{sub}} = 3$	$\mathcal{W}_1$	$-20.413/0.8\%$	$-20.413/0.8\%$	-20.568	$-18.665/4.0\%$	$-18.665/4.0\%$	-19.445	2.7×10^{-2}	7
$N_{\text{qubit}} = 4$	$\mathcal{W}_2$	$-20.513/0.3\%$	$-20.486/0.4\%$		$-19.265/0.9\%$	$-19.199/1.3\%$		2.4×10^{-1}	12
$N_{\text{sub}} = 2$	$\mathcal{W}_1$	$-20.480/0.4\%$	$-20.480/0.4\%$	-20.568	$-18.286/6.0\%$	$-18.286/6.0\%$	-19.445	3.9×10^{-3}	4
$N_{\text{qubit}} = 6$	$\mathcal{W}_2$	$-20.560/0.04\%$	$-20.551/0.08\%$		$-19.343/0.5\%$	$-19.324/0.6\%$		6.3×10^{-2}	8
$N_{\text{sub}} = 2$	$\mathcal{W}_1$	$-27.535/0.4\%$	$-27.535/0.4\%$	-27.647	$-25.374/5.2\%$	$-25.374/5.2\%$	-26.770	2.4×10^{-4}	4
$N_{\text{qubit}} = 8$	$\mathcal{W}_2$	$-27.634/0.05\%$	$-27.612/0.1\%$		$-26.620/0.6\%$	$-26.578/0.7\%$		7.4×10^{-3}	10

$6.464 + 6 = 12.464$. Therefore, it is sufficient to insert auxiliary dimensions by $J_i = \lambda_i I_{2^{N_{\text{eff}}} - K}$ with $\lambda_i > 12.464$ (for ground states) or $\lambda_i > 12.464 + \Delta_n^*$ (for excited states) with the estimated gap $\Delta_n^* \sim E_n(\tilde{H}) - E_0(\tilde{H})$. Nevertheless, we directly confirm that both the ground-state energy and the first-excited-state energy of $\tilde{H}$ are the same as those of $\tilde{H}_{\text{eff}}$ by exact diagonalization in this simulation, so we simply set $\lambda_i = 0$ in the following.

We run a classical simulation of Deep VQE for the AFM Heisenberg model (21). We exploit a hardware-efficient type ansatz [20,37] as a variational quantum circuit. We use this ansatz in the VQE for finding the ground states of the subsystem Hamiltonians (Step 1 of Deep VQE) and SSVQE for finding the ground and first-excited state of the effective Hamiltonian $\tilde{H}_{\text{eff}}$ (Step 3 in Deep VQE). The simulation of quantum circuits is performed by using the libraries Qulacs [38,39] and OpenFermion [40]. More details on numerical simulations are described in Appendix D.

We show the numerical results for the AFM Heisenberg chain with various partitioning of the total system in Table I. We calculate the ground-state energy E_0 and the first-excited-state energy E_1 in three ways and compare them. The first one, "Effective" is obtained by exact diagonalization of the effective Hamiltonian $\tilde{H}$. This value tells us provides the best performance of Deep VQE in theory. The second one, "Deep VQE," is obtained by simulations of Deep VQE. The difference between "Effective" and "Deep VQE" indicates errors originating from VQEs in Step 1 and 3 of Deep VQE. The last one, "ED," is the exact energy of of the original Hamiltonian H calculated by the exact diagonalization. The difference between "Effective" and "ED" tells the validity of the choice of the local basis in Step 2 of Deep VQE. In the table, we also introduce the truncation rate TR defined by

$$\text{TR} = \dim(\tilde{\mathcal{H}})/\dim(\mathcal{H}) = (K/2^{N_{\text{qubit}}})^{N_{\text{sub}}}, \quad (24)$$

as a figure of merit of the reduction of the Hilbert space. The accurate results for E_n(Effective) ($n = 0, 1$) with small TR mean that the coarse graining is performed with a proper choice of a local basis set.

As shown in Table I, in terms of the ground-state energy E_0, both the original Deep VQE ($\mathcal{W}_1$) and the modified Deep VQE ($\mathcal{W}_2$) exhibit good performance with relative errors to

"ED" values (eigenvalues of the original Hamiltonian) up to 0.8%, as expected. On the other hand, in terms of the first-excited-state energy E_1, the modified Deep VQE ($\mathcal{W}_2$) shows comparably good performance with relative errors from 0.6% to 1.3%, compared to the original Deep VQE ($\mathcal{W}_1$) results having relative errors from 4.0% to 6.4%. Considering that the value of "Effective" in the original Deep VQE ($\mathcal{W}_1$) deviates from that of "ED", the modified choice of local basis is essential for obtaining low-energy excited states accurately, i.e., the larger error in the original Deep VQE ($\mathcal{W}_1$) is not due to the imperfection of the optimization of the VQE in Deep VQE algorithm. We also evaluate the number of measurements. For $N_{\text{sub}} = 2$ and $N_{\text{qubit}} = 8$ with $\mathcal{W}_2$, each subsystem term $\tilde{H}_i \oplus 0_{2^{N_{\text{eff}}} - K}$ is decomposed into 452 Pauli terms. While we do not introduce the energy shift λ_i for the simulation, we can complete the protocol with the same number of measurements 452 for $\tilde{H}_i \oplus \lambda_i I_{2^{N_{\text{eff}}} - K}$.

To further understand the higher accuracy for the excited states achieved with our local basis choice $\mathcal{W}_2$, we confirm the validity of QSE for each subsystem as discussed in Sec. III A. QSE for each subsystem with the local excitations $\mathcal{W}_1$ ($\mathcal{W}_2$) and the reference state $|\psi_0\rangle_i$ is equivalent to calculating the eigenvalues of $\tilde{H}_i$ obtained from $\mathcal{W}_1$ ($\mathcal{W}_2$) and $|\psi_0\rangle_i$. In the case of $N_{\text{qubit}} = 6$, the exact first-excite-state energy for each subsystem is $E_1^{\text{sub}} = -8.008$ with three-fold degeneracy, which we calculate by the original H_i. On the other hand, by diagonalizing $\tilde{H}_i$, we obtain the QSE results $E_1^{\text{sub}} = -6.415$ with nearly three-fold degeneracy (for $\mathcal{W}_1$) and $E_1^{\text{sub}} = -8.000$ with three-fold degeneracy (for $\mathcal{W}_2$). This means the failure of QSE with $\mathcal{W}_1$ and the success of QSE with $\mathcal{W}_2$.

As discussed in Sec. III A, the success of QSE in expressing the local first-excited states ensures the accurate first-excited-state energy by the consistency with the perturbation theory. The above result on the invalidity and validity of QSE for $\mathcal{W}_1$ and $\mathcal{W}_2$ explains why our modified local excitations $\mathcal{W}_2$ gives much better approximation for the excited states than the original one $\mathcal{W}_1$. This example also supports the scenario of the modified Deep VQE in Sec. III A— the modified local excitations reflecting the intrasubsystem Hamiltonians result in the accurate low-lying excited states due to the success of QSE for each subsystem. Based on

this finding, although we simply use all the Pauli operators X, Y, Z, we can find a better choice which efficiently describes each subsystem by employing some methods reflecting the type of intrasubsystem Hamiltonians H_i such as Ref. [41].

E. Discussion for improving the accuracy

We finally discuss how to improve the accuracy. The key is to capture higher-order excitations caused by intrasubsystem and intersubsystem interactions within each subsystem. For example, the product of Pauli operators like $A_i B_j$ ($A, B = X, Y, Z$) for different qubits i, j within the same subsystem can be included in the local operator set in Step 2 of Deep VQE, while our simulation only considers linear excitations of such Pauli operators. This is similar to considering higher-order excitations in QSE. The drawback for including the higher-order excitations is the increase of the effective system size. When we consider up to the nth order product of Pauli operators like $P_{j_1} \cdots P_{j_n}$, the dimension of the local Hilbert space becomes $K \sim (N_{\text{qubit}})^n$. Although we have $(N_{\text{qubit}})^n \ll 2^{N_{\text{qubit}}}$ (the dimension of the original local Hilbert space) for small n, the order of excitations n has to be determined by comparing the benefit in the accuracy with the computational cost.

IV. MAIN RESULT: APPLICATION TO CHEMISTRY PROBLEMS

In this section, we provide the second half of our main results. We apply our modified VQE to the simplest example of a quantum chemistry calculation for periodic materials: electronic states of a periodic hydrogen chain. We show numerical simulation results and examine the validity of our modified Deep VQE on the low-energy eigenstates of periodic materials. Our numerical results also can be seen as the first application of Deep VQE to fermionic systems that map to nonlocal qubit Hamiltonians.

A. Model and Method

Let us describe our model and how to perform Deep VQE. We consider a one-dimensional chain under periodic boundary condition whose unit cell is composed of two hydrogen atoms [see Fig. 3(a)]. The distance between the two atoms is d and the length of the unit cell is $2d$. After performing the crystal Hartree-Fock calculation with STO-3G basis set [28], we obtain the Hamiltonian in the second-quantized form as

$$H = \sum_k \sum_{pq} t_k^{pq} c_{kp}^\dagger c_{kq}$$
$$+ \sum_{k_1 k_2 k_3 k_4}{}' \sum_{pqrs} v_{k_1 k_2 k_3 k_4}^{pqrs} c_{k_1 p}^\dagger c_{k_2 q}^\dagger c_{k_3 r} c_{k_4 s}, \quad (25)$$

where $c_{kp}^\dagger$ (c_{kp}) is a creation annihilation operators of electrons with crystalline momentum k and the spin orbital p. The crystalline momentum k is uniformly sampled from the first Brillouin zone $[0, 2\pi]$, which is renormalized by the unit-cell length $2d$. The index $p \in [1, 2, 3, 4]$ represents the spin-orbital within a unit cell, which comes from two spin-orbitals in the STO-3G basis for two hydrogen atoms. The orbitals $p = 1, 2$ ($p = 3, 4$) are occupied (unoccupied)
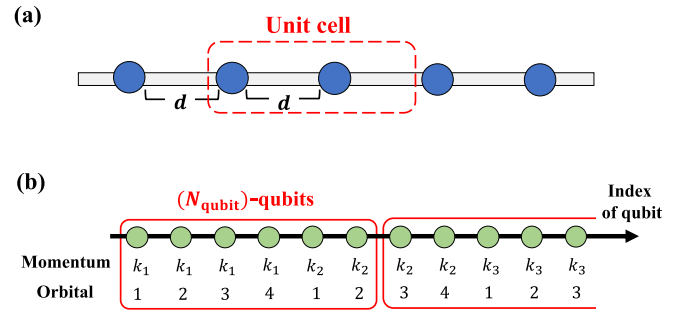

FIG. 3. (a) A one-dimensional hydrogen chain. A unit cell includes two hydrogen atoms, and the distance between the atoms is d. (b) The way of splitting the hydrogen chain into subsystems in the momentum space. Neighboring N_{qubit} qubits, corresponding to electrons with similar momentum, belong to the same subsystem. Here, the orbitals 1 and 2 (3 and 4) represent the occupied (unoccupied) crystalline orbitals under the Hartree-Fock approximation.

orbitals in the Hartree-Fock state. The symbol Σ' represents the summation over the crystalline momentum under the crystalline momentum conservation, satisfying

$$k_1 + k_2 - k_3 - k_4 \in 2\pi \mathbb{Z}. \quad (26)$$

The sets of coefficients $\{t_k^{pq}\}$ and $\{v_{k_1 k_2 k_3 k_4}^{pqrs}\}$ represent one-body and two-body electron integrals between different crystalline Hartree-Fock orbitals (k, p), determined by classical computers.

To perform Deep VQE, we transform the fermionic Hamiltonian (25) into the one on qubits by the Jordan-Wigner transformation [42]. Each qubit is still labeled by (k, p) of a corresponding electron. Through this transformation, some nonlocal terms appear in the Hamiltonian, e.g., a Pauli string $X_1 Z_2 X_3 X_5 Z_6 X_7$ coming from $c_3^\dagger c_7^\dagger c_1 c_5$. This nonlocality potentially harms the validity of Deep VQE because the eigenstates can be entangled nonlocally among the system, so our simulation also examines how the nonlocality affects the accuracy of Deep VQE. When we sample N_k points as the crystalline momentum, $4N_k$ qubits are required to represent the original Hamiltonian (25).

Next, we depict how to perform each step of Deep VQE for the periodic hydrogen chain including the way to divide it into subsystems. We define subsystems so that each orbital (qubit) in the subsystem has similar the crystalline momentum k. To be precise, we label the qubits as $(k, p) = (k_1, 1), (k_1, 2), (k_1, 3), (k_1, 4), (k_2, 1), (k_2, 2), \ldots$ with the order $0 = k_1 < k_2 < \ldots$ [see Fig. 3(b)] and group neighboring N_{qubit} qubits with the smaller index step by step. We stress that we define the subsystems in the momentum space, not the real space.

We consider two set of local excitation operators for Step 2 of Deep VQE as

$$\mathcal{W}_s = \{I\} \cup \{c_j', c_j'^\dagger \mid j \in \Lambda_e\}, \quad (27)$$

$$\mathcal{W}_d = \mathcal{W}_s \cup \{c_j'^\dagger c_k' \mid j, k \in \Lambda_e\}, \quad (28)$$

where Λ_e is the set of (k, p) in each subsystem (we have not explicitly written the index for the subsystem in $\mathcal{W}_s$ and $\mathcal{W}_d$). The qubit operator c_i' ($c_i'^\dagger$) is obtained by a fermion

annihilation (creation) operator truncated within each subsystem. Namely, we define c'_i $(c'^{\dagger}_i)$ by restricting the transformed Pauli operators of c_i $(c^{\dagger}_i)$ to act only on the target subsystem. For instance, in the case of $N_{\text{qubit}} = 4$, we define $c'_6 = Z_5(X_6 - iY_6)/2$ while the genuine qubit-operator representation of c_6 obtained by the Jordan-Wigner transformation is $c_6 = Z_1 Z_2 Z_3 Z_4 Z_5 (X_6 - iY_6)/2$. Although this truncation means neglecting the anticommutation relations between c'_i and $c'^{\dagger}_i$ in different subsystems, we expect that they are sufficient to describe locally excited states in each subsystem. Under this definition, $\mathcal{W}_s$ is a set of single-particle excitations giving a local Hilbert space dimension $K = |\mathcal{W}_s| = 2N_{\text{qubit}} + 1$. On the other hands, $\mathcal{W}_d$ is a set of single- and double- particle excitations giving a local Hilbert space dimension $K = (N_{\text{qubit}} + 1)^2$.

In Step 3 of Deep VQE, we should specify the way of inserting auxiliary dimensions following Eqs. (15)–(17). We again employ $\lambda_i = 0$ because we directly confirm that the low-energy eigenvalues of $\tilde{H}_{\text{eff}}$ and those of $\tilde{H}$ coincide by the exact diagonalization.

B. Numerical results

In numerical simulations of Deep VQE for the periodic hydrogen chain, three k points are sampled (the total number of qubits is $3 \times 4 = 12$), and we split the system by $N_{\text{sub}} = 2$ and $N_{\text{qubit}} = 6$. Here, we employ a hardware-efficient type ansatz [20,37] of depth 20 for VQE in Step 1 of Deep VQE, and that of depth 80 for SSVQE in Step 3 to simply examine the performance of our protocol. To compute larger systems with smaller depth of quantum circuits, it will be better to employ VQD or MCVQE instead of SSVQE (see Fig. 2) and other ansatz avoiding the barren plateau problem [43] instead of hardware-efficient ansatz. All circuit simulations are performed by using the libraries Qulacs [38,39] and Open-Fermion [40], and the crystalline Hartree-Fock calculation is done by PySCF package [44,45] (see Appendix D for the detail).

To assess the accuracy of Deep VQE result, we introduce two kinds of values, dubbed "Local" and "Effective: ED." The former one is defined by

$$E_0(\text{local}) = \langle \Psi_0 | H | \Psi_0 \rangle, \tag{29}$$

where $|\Psi_0\rangle = \bigotimes_{i=1}^{N_{\text{sub}}} |\psi_0\rangle_i$ is the ground states of the intrasubsystem Hamiltonian $H_{\text{intra}} = \sum_{i=1}^{N_{\text{sub}}} H_i$. The energy $E_0(\text{local})$ represents the result of the first-order perturbation theory with the unperturbed Hamiltonian H_{intra} and the perturbation $V_{\text{inter}} = H - H_{\text{intra}}$. The latter one "Effective: ED" is calculated by substituting VQE in Step 1 of Deep VQE by exact diagonalization and performing the exact diagonalization again to solve $\tilde{H}_{\text{eff}}$ in Step 3 of Deep VQE. The value of "Effective: ED" gives the best performance of Deep VQE in theory since we solve all the Hamiltonians in Deep VQE exactly. (Note that its definition is different from that of "Effective" in Sec. III in that the latter exploits VQE in Step 1 and ED in Step 3.)

First, we choose the single-particle excitations $\mathcal{W}_s$ as a local basis of Step 2 of Deep VQE and set $d = 1.4$ Bohr. The results are shown in Table II. We find a nice agreement between the "Deep VQE" and the "ED" within relative error

TABLE II. Numerical results for the ground-state and the first-excited-state energies of the periodic hydrogen chain under $d = 1.4$ Bohr. The left values in the cells represent energies delivered in Hartree. The right values are relative errors from the exact results $E_n(\text{ED})$.

$E_0(\text{Local})$	$E_0(\text{Effective: ED})$	$E_0(\text{Deep VQE})$	$E_0(\text{ED})$
$-0.743/31\%$	$-1.067/1.4\%$	$-1.067/1.4\%$	-1.082
	$E_1(\text{Effective: ED})$	$E_1(\text{Deep VQE})$	$E_1(\text{ED})$
	$-0.743/4.1\%$	$-0.743/4.1\%$	-0.775

up to a few percent, especially compared to the value of the perturbation theory "$E_0(\text{local})$". It should be noted that this accuracy characterized by the relative errors up to 4.1% is achieved despite truncating the vast majority of the whole Hilbert space, TR = 0.0413, and reducing the number of qubits by four ($12 \rightarrow 8$). We also find that the values of "Effective: ED" and those of "Deep VQE" are almost identical, which implies that the error of the "Deep VQE" results from the exact ones solely comes from the truncation of the Hilbert space in Step 2 of Deep VQE.

Next, let us investigate how the validity of Deep VQE depends on the atom-atom distance d and the local basis. From now on, we consider the values of $E_n(\text{Effective: ED})$ $(n = 0, 1)$ instead of the Deep VQE results $E_n(\text{Deep VQE})$ due to the computational cost of classical simulations. As discussed in the above, the values of $E_n(\text{Effective : ED})$ give the best performance of the Deep VQE protocol. In some cases, we observe the two-fold degeneracy for the ground states of each subsystem, and we randomly choose the local ground states $|\psi_0\rangle_i$ in such situations. Figure 4 shows the low-energy eigenvalues (upper panels) and the relative error (lower panels) of $E_n(\text{Effectice: ED})$ from the exact results $E_n(\text{ED})$ for (a) the ground state and (b) the first-excited states. For the ground-state energy, Fig. 4(a) shows that the Deep VQE results with the local operators $\mathcal{W}_s$ and $\mathcal{W}_d$ both can provide much more accurate values than those of the first-order perturbation theory, $E_0(\text{local})$, within a wide range of the distance d. When we only consider the single-particle excitations by $\mathcal{W}_s$, the obtained ground-state energy is accurate under small distance d while the result approaches $E_0(\text{local})$ with increasing d. By taking into account double-particle excitations with $\mathcal{W}_d$, we can improve the accuracy, especially in the large-d regime with making the relative error approximately half. The results for the first-excited states have a similar tendency. We can see that the first-excited-state energy error is suppressed up to 10%. The improvement of the accuracy by considering double-particle excitations becomes larger as the distance d increases, with giving the errors approximately half as large as those with single-particle excitations.

We remark on the required number of qubits for simulation. For both cases with the single- and double-particle excitations, we observe some states in $\{P_k^{(i)} |\psi_0\rangle_i\}_{k=1}^K$ are not linearly independent. For the former case $\mathcal{W}_s$, the numerically-obtained dimension K of the restricted Hilbert space ranges from 11 to 13, giving the Hilbert space truncation rate TR [Eq. (24)] ranging from 0.030 to 0.035. (Note that TR obtained here is different from that of the Deep VQE, 0.0413, due to the
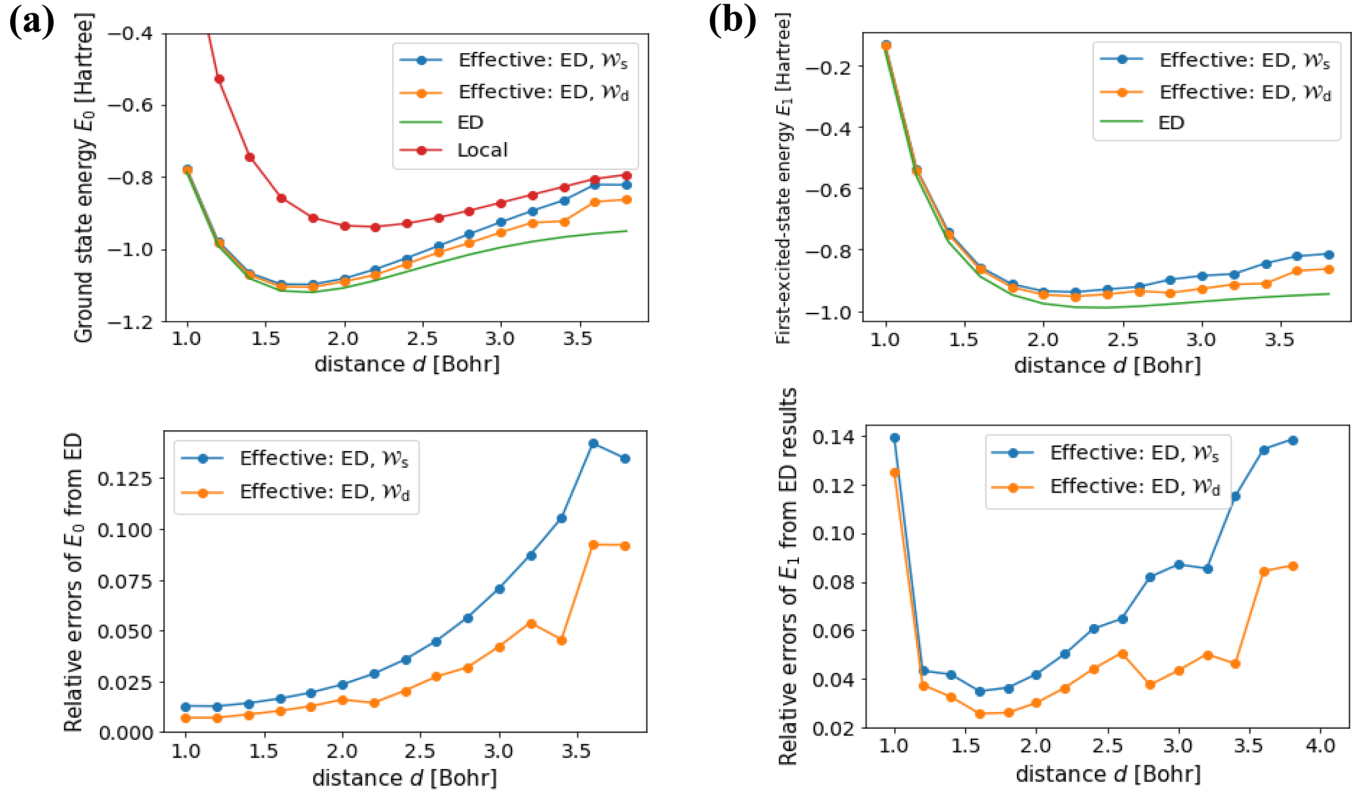
FIG. 4. Numerical results for the hydrogen chain with picking up three k-points. The values of "ED" and "Local" show the exact and first-order-perturbation (29) results, respectively. "Effective: ED" results are obtained by substituting VQEs with ED in Steps 1 and 3 of Deep VQE, giving the best performance of Deep VQE in theory. (a) The ground-state energy (upper panel) and the relative errors from the ED values (lower panel). (b) The first-excited-state energy (upper panel) and the errors from the ED values (lower panel). For both the ground- and first-excited-state energies, the relative errors from the exact results with double-particle excitations are suppressed approximately half compared to those with single-particle excitations.

degenerate ground states in the local Hilbert space.) The Deep VQE calculation for the original model [Eq. (25)] of 12 qubits can be executed with a 8-qubit quantum device in this case. On the other hand, for the latter case $\mathcal{W}_d$, we find the numerically-obtained dimension K of the restricted Hilbert space and the truncation rate is $24 \leqslant K \leqslant 34$ and $0.15 \leqslant \mathrm{TR} \leqslant 0.22$, respectively. The number of the required qubits for the Deep VQE protocol is 11 for the atom-atom distance $d = 3.4$ Bohr and 10 for the other choices of d.

Finally, we discuss two possible directions to improve the accuracy. The first way is to increase the number of the reference states for the local basis, e.g., taking the local basis like $\{P_k^{(i)} |\psi_0\rangle_i\} \cup \{P_k^{(i)} |\psi_1\rangle_i\} \cup \cdots$. It is reasonable to include low-energy eigenstates of the subsystems other than the ground state for constructing the local basis to capture the low-energy physics in the original system. We discuss such a protocol in Appendix B. The second possible way is to consider higher-order excitations for the construction of local basis set. We can show that the Deep VQE with higher-order local excitations gives the better result than the combination of QSE and Deep VQE for the whole system with lower-order excitations (see Appendix E for the detail). As well as Sec. III E, the number of qubits for simulation is given by

$$N_{\mathrm{req}} \sim \min(N_{\mathrm{qubit}}, nN_{\mathrm{sub}}\lceil \log_2 N_{\mathrm{qubit}} \rceil) \ll N_{\mathrm{qubit}} \times N_{\mathrm{sub}} \quad (30)$$

when we consider up to the nth order excitation. Therefore, as the excitations up to double-particle $\mathcal{W}_d$ overwhelms those up to single-particle $\mathcal{W}_s$ in terms of accuracy, the local basis with higher order excitations will give more accurate low-lying excited states with keeping the reduction of qubits.

V. DISCUSSION AND CONCLUSIONS

In this paper, we have proposed the improved way for performing Deep VQE, in which we can properly obtain low-energy eigenstates with arbitrary preferable variational quantum circuits on a smaller number of qubits. We have composed a set of local excitation operators used for the coarse graining in Deep VQE by focusing on the excitations caused by intrasubsystem interactions. The perturbation theory and QSE ensure the validity for low-energy eigenstates obtained by such Deep VQE protocols. We have also provided an alternative way of constructing the effective Hamiltonian defined on qubits, in which we introduce penalty terms to the auxiliary dimensions. We have derived a rigorous bound on the penalty that makes arbitrary variational quantum states available. After reformulating Deep VQE for excited states, we have applied it to periodic materials, namely, a periodic hydrogen chain. We have shown that the low-energy eigenvalues are well reproduced with relative errors up to $O(1)\%$ in a small atom-atom distance regime by splitting the system into

subsystems based on crystalline momentum and considering the local basis introduced by single-particle and double-particle excitations. Our results enlarge the possibility of simulating large systems by a small-sized quantum computers even for excited states of quantum systems including, quantum chemistry.

We provide some comments on how our method outperforms conventional classical methods. Our method searches low-energy eigenstates from the restricted Hilbert space spanned by

$$\bigotimes_{i=1}^{N_{\text{sub}}} \left(P_{k_i}^{(i)} |\psi_0\rangle_i \right), \quad k_i = 1, 2, \ldots, K. \tag{31}$$

Due to the coarse graining, this can accurately capture strong intrasubsystem correlations and weak intersubsystem correlations. In Eq. (31), each pair of subsystems is equivalent to one another. Thus, our method can deal with 1D systems with long-range interactions, which is usually difficult for classical methods based on matrix product states (MPS) [46–48]. This will benefit chemistry problems where the Coulomb interactions play a central role. In addition, the applicability of long-ranged models indicates the validity of our method for higher dimensional systems unlike MPS-based methods. Our method will overwhelm the cluster mean-field theory (MFT), which is a classical method for high-dimensional systems combined with the coarse graining, since it can deal with entanglement between subsystems.

We leave some future directions for this study. First, while our results provide approximate ground-state and first-excited-state energies for periodic materials with relative errors up to a few percents, the ultimate goal is to predict them with the chemical accuracy (1.6×10^{-3} Hartree) by Deep VQE. As discussed in Secs. III E and IV B, the way to improve the accuracy would be to take higher-order excitations into account for the local basis. Although the models examined in this paper are too small to get benefit from considering higher-order excitations because those excitations exhaust the original (unrestricted) Hilbert space and there is no decrease in the number of qubits, large systems targeted by the NISQ devices with hundreds or thousands of qubits can gain the improvement of the accuracy and the decrease of the number of qubits by considering higher-order excitations. Second, it is intriguing to study the performance of our modified VQE when we repeat the coarse graining of Deep VQE many times, as proposed in Ref. [29]. The Deep VQE discussed in this paper performs the coarse graining once in Step 2. As the number of repetition of the coarse graining increases, the Deep VQE can simulate much larger systems, but instead, its effective Hamiltonian gradually becomes nonlocal and discards the information of the original systems. It should be an important problem whether the modified Deep VQE with repeating the coarse graining many times well reproduces low-energy eigenstates of huge systems. Third, from the practical point of view, the feasibility of our modified VQE in the real NISQ hardware can be further investigated. For example, analyzing the effect of noise in the NISQ devices on the coarse graining of Deep VQE is helpful to find how much we alleviate the noise by error mitigation techniques [49–51]. Another possible obstacle is the problem so-called

barren plateau [43], in which the optimization of variational quantum circuits gets difficult for large and deep quantum circuits. Our results in this paper nevertheless provides the solid support that our modified VQE works in noiseless situations.

APPENDIX A: GRAM-SCHMIDT METHOD

We describe the classical algorithm for obtaining the orthonormal basis $\{|\phi_k\rangle_i\}_{k=1}^K$ in Step 2 of the original and modified Deep VQE based on the Gram-Schmidt method [29]. To be precise, we derive the matrix elements $S_{kk'}^{(i)}$ satisfying $|\phi_k\rangle_i = \sum_{k'=1}^K S_{k'k}^{(i)}(P_{k'}^{(i)}|\psi_0\rangle_i)$ from the overlap $G_{kk'}^{(i)} = \langle\psi_0|P_k^{(i)\dagger}P_{k'}^{(i)}|\psi_0\rangle_i$. By the standard formulation of the Gram-Schmidt method, one can choose the first normalized basis by $|\phi_1\rangle_i = P_1^{(i)}|\psi_0\rangle_i/(G_{11}^{(i)})^{1/2}$, giving the matrix elements

$$S_{11}^{(i)} = \left(G_{11}^{(i)}\right)^{-1/2}, \quad S_{1l'}^{(i)} = 0, \quad (l' = 2, 3, \ldots, K). \tag{A1}$$

The second basis is given by

$$|\phi_2\rangle_i = \frac{P_2^{(i)}|\psi_0\rangle_i - \langle\phi_1|P_2^{(i)}|\psi_0\rangle_i|\phi_1\rangle_i}{\left\| P_2^{(i)}|\psi_0\rangle_i - \langle\phi_1|P_2^{(i)}|\psi_0\rangle_i|\phi_1\rangle_i \right\|}, \tag{A2}$$

which becomes a linear combination of $P_1^{(i)}|\psi_0\rangle_i$ and $P_2^{(i)}|\psi_0\rangle_i$ by Eq. (A1). The denominator $D > 0$ is expressed in $S^{(i)}$ and $G^{(i)}$ as

$$D^2 = \left\| P_2^{(i)}|\psi_0\rangle_i - \langle\phi_1|P_2^{(i)}|\psi_0\rangle_i|\phi_1\rangle_i \right\|^2$$
$$= G_{22}^{(i)} - 2\left|S_{11}^{(i)}G_{12}^{(i)}\right|^2 + \left|S_{11}^{(i)}\right|^4\left|G_{12}^{(i)}\right|^2 G_{11}^{(i)}, \tag{A3}$$

which results in the elements $S_{21}^{(i)} = -|S_{11}^{(i)}|^2 G_{12}^{(i)}/D$, $S_{22}^{(i)} = 1/D$, and $S_{2l'}^{(i)} = 0$ for $l' = 3, 4, \ldots, K$.

In a similar way, when considering the kth basis given by

$$|\phi_k\rangle_i = \frac{P_k^{(i)}|\psi_0\rangle_i - \sum_{k'=1}^{k-1}\langle\phi_{k'}|P_k^{(i)}|\psi_0\rangle_i|\phi_{k'}\rangle_i}{\left\| P_k^{(i)}|\psi_0\rangle_i - \sum_{k'=1}^{k-1}\langle\phi_{k'}|P_k^{(i)}|\psi_0\rangle_i|\phi_{k'}\rangle_i \right\|}, \tag{A4}$$

we already have the matrix elements $S_{ll'}^{(i)}$ for $l = 1, 2, \ldots, k-1$ and $l' = 1, 2, \ldots, K$. By expanding $|\phi_k'\rangle_i$ by $\{P_1^{(i)}|\psi_0\rangle_i, \ldots, P_{k-1}^{(i)}|\psi_0\rangle_i\}$ and expressing the denominator in $S^{(i)}$ and $G^{(i)}$, we can obtain $S_{kl'}$ for $l' = 1, 2, \ldots, K$. By repeating this up to $k = K$, we obtain all the matrix elements of $S^{(i)}$.

Since the above protocol repeats processing matrices with $O(K^2)$ size K times, this step requires $O(N_{\text{sub}}K^3)$ time in classical computation. We do not apply any approximation here, and no errors other than computational errors arise.

These classical errors ε_{cl} deviate the low-energy eigenvalues by at-most $O(N_{sub}\varepsilon_{cl}K^2)$ from Eqs. (5) and (6). The errors ε_{cl} usually have sufficiently small values, and their effects on the low-energy eigenvalues are negligible. However, if we have vanishing or nearly-vanishing denominators in Eq. (A4), the classical errors ε_{cl} can become non-negligible. In that case, some of $\{P_k^{(i)}|\psi_0\rangle_i\}$ are linearly dependent, and we should exclude the corresponding states. In our simulation, we observe such linear dependence for the periodic hydrogen chain in Sec. IV. For simplicity, we set $S_{kl'}^{(i)} = 0$ instead of excluding the state $P_k^{(i)}|\psi_0\rangle_i$.

APPENDIX B: MULTISTATE DEEP VQE FOR EXCITED STATES

In this section, we discuss another choice of the local basis in Step 2 of Deep VQE for calculating low energy excited states. As discussed in Sec. III A of the main text, it is important to choose local excitation operators $\{P_k^{(i)}\}$ so that they reproduce excitations by intrasubsystem terms in addition to those by intersubsystem terms. While we choose Pauli operators on the whole subsystem as such excitation operators in the main text, we can also pick up the following local basis set instead:

$$\bigcup_{m=1}^{M} \{P_k^{(i)}|\psi_m\rangle_i\}_{k=1}^{K}, \tag{B1}$$

with the original choice of $\{P_k^{(i)}\}$ (local operators relevant to the intersubsystem interactions). The low-energy excited states of each subsystem $|\psi_m\rangle_i$ can be obtained by using VQE-based algorithms for excited states such as SSVQE [7] instead of the usual VQE in Step 1, with a N_{qubit}-qubit device. The dimension of the local Hilbert space is $M \times K$, typically smaller than $2^{N_{qubit}}$.

Reference [29] reported that the Deep VQE based on this choice with $M > 1$, $K = 1$ fails to capture the ground state of the whole system, since it neglects the local excitations coming from intersubsystem interactions. However, if we consider Deep VQE with $M > 1$ and $K > 1$, it is expected to well describe both the ground state and the low-energy eigenstates since it captures excitations from both intersubsystem and intrasubsystem terms, just as the modified Deep VQE in the main text. One of the possible problems of this version of Deep VQE compared to the one in the main text lies in calculating matrix elements of the effective Hamiltonian in Step 2: we should compute off-diagonal elements $_i\langle\psi_m|P_k^{(i),\dagger}P_l^{(i)}|\psi_n\rangle_i$ and $_i\langle\psi_m|P_k^{(i),\dagger}AP_l^{(i)}|\psi_n\rangle_i$ $(A = H_i, V_i^\alpha, W_i^\alpha)$ for $m \neq n$. We should choose a proper method for calculating the excited states of the subsystem in Step 1 of Deep VQE that enables us to evaluate those off-diagonal elements easily. The algorithms SSVQE and MCVQE satisfy this requirement while VQD dose not, demanding auxiliary qubits and/or the increase of the circuit depth to evaluate the matrix elements.

Thus, the choice of the local basis described in the main text is more preferable for calculating low-energy eigenstates in that we can use any kind of VQE-based algorithms in Step 1 of Deep VQE. On the other hand, the multistate local basis like Eq. (B1) can be suitable compared to the one in the main text when the intrasubsystem Hamiltonian is nonlocal enough

to break the validity of QSE in each subsystem (see discussion in Sec. III A). We also note that, when the gap above the subsystem ground state $|\psi_0\rangle_i$ is small enough, we may have to consider the multistate local basis to capture the low-energy states for the whole system.

APPENDIX C: PROOF OF PROPOSITION IN SEC. III B

We derive the proposition in Sec. III B. First, we clarify the setup for the construction of the effective qubit model and the notation. We consider an effective Hamiltonian

$$\tilde{H} = H|_{\tilde{\mathcal{H}}} = \sum_{i=1}^{N_{sub}} \tilde{H}_i + \sum_{i\neq j}^{N_{sub}} \sum_\alpha v_{ij}^\alpha \tilde{V}_i^\alpha \otimes \tilde{W}_j^\alpha, \tag{C1}$$

which is defined on the $K^{N_{sub}}$-dimensional Hilbert space $\tilde{\mathcal{H}}$ as a result of choosing the local basis (see Step 2 in Sec. II). We can regard the Hamiltonian $\tilde{H}$ as that on a lattice $\Lambda = \{1, \ldots, i, \ldots, N_{sub}\}$ where each subsystem $i \in \Lambda$ has K levels. We define N_{eff} by the smallest integer that exceeds $\log_2 K$, and we introduce auxiliary dimensions by

$$\tilde{H}_{eff} = \sum_{i=1}^{N_{sub}} \tilde{H}_{i,eff} + \sum_{i\neq j}^{N_{sub}} \sum_\alpha v_{ij}^\alpha \tilde{V}_{i,eff}^\alpha \otimes \tilde{W}_{j,eff}^\alpha, \tag{C2}$$

$$\tilde{H}_{i,eff} = \tilde{H}_i \oplus J_i, \tag{C3}$$

$$\tilde{V}_{i,eff}^\alpha = \tilde{V}_i^\alpha \oplus 0_M, \quad \tilde{W}_{j,eff}^\alpha = \tilde{W}_j^\alpha \oplus 0_M, \tag{C4}$$

with $M = 2^{N_{eff}} - K$. As discussed in Sec. II, this insertion of auxiliary dimension generates additional meaningless eigenvalues giving strong limitation on variational quantum circuits. We derive the way to construct $M \times M$ matrices $\{J_i\}_{i=1}^{N_{sub}}$ so that the Hamiltonians $\tilde{H}$ and $\tilde{H}_{eff}$ can have the same low-energy spectrum.

We introduce some notations for matrices. For a finite-dimensional hermitian matrix A, we describe its eigenvalues by $E_0(A), E_1(A), \ldots,$ with $E_n(A) \leqslant E_{n+1}(A)$ and the set of the eigenvalues by $\mathrm{Spec}(A) = \{E_n(A)\}_{n=0}^{\dim(A)-1}$. When H is a Hamiltonian defined on a lattice of subsystems $\Lambda = \{1, \ldots, i, \ldots, N_{sub}\}$, we can always decompose it as

$$H = \sum_{X \subset \Lambda} h_{X,\Delta}, \tag{C5}$$

where $h_{X,\Delta}$ is a Hermitian operator nontrivially acting just on a domain $X \subset \Lambda$. For instance, two-body interactions are represented by a series of $h_{X,\Delta}$ with $|X| = 2$. We note that the way of the decomposition is not unique, and hence we designate its choice by the subscript Δ. Then, we define $H(D)$ for a certain domain $D \subset \Lambda$ by

$$H_\Delta(D) = \sum_{X:X\cap D\neq\phi} h_{X,\Delta}, \tag{C6}$$

which represents the part of H nontrivially acting on some sites in D. When we choose a domain D as the ith subsystem, it becomes

$$H_\Delta(D = \{i\}) = \sum_{X:X\ni i} h_{X,\Delta}. \tag{C7}$$

Then, we define the extensiveness of Hamiltonian H for the ith subsystem by

$$e_{H,\Delta}(i) \equiv \sum_{X:X \ni i} ||h_{X,\Delta}||_{\mathrm{op}}, \qquad (C8)$$

in which $||\ \ ||_{\mathrm{op}}$ denotes the operator norm. The extensiveness $e_{H,\Delta}(i)$ represents the maximal energy on the ith subsystem under the Hamiltonian H. Note that $e_{H,\Delta}(i)$ depends on the decomposition of the Hamiltonian Δ. When we designate the decomposition by Eq. (C1), the extensiveness becomes

$$e_{\tilde{H}}(i) = ||\tilde{H}_i||_{\mathrm{op}} + \sum_{\alpha} \sum_{j,k:\{j,k\} \ni i} |v_{jk}^{\alpha}| \cdot ||\tilde{V}_j^{\alpha}||_{\mathrm{op}} \cdot ||\tilde{W}_k^{\alpha}||_{\mathrm{op}}$$

$$(C9)$$

However, the choice is not essential in the discussion below, and hence we omit the subscript Δ like h_X, $H(D)$, and $e_H(i)$ in the following.

Before proving the proposition in the main text, we show the relation between eigenvalues of $\tilde{H}$ and those of $\tilde{H}_{\mathrm{eff}}$. The result is summarized as follows.

Proposition 1. We introduce auxiliary dimension by Eqs. (C2)–(C4). For a domain $D \subset \Lambda$, we define a $(K^{N_{\mathrm{sub}}-|D|} \times M^{|D|})$-dimensional Hilbert space $\tilde{\mathcal{H}}_D$, where subsystems out of D and those in D have K and M degrees of freedom, respectively. Then, the set of eigenvalues of $\tilde{H}_{\mathrm{eff}}$ is decomposed as follows:

$$\mathrm{Spec}(\tilde{H}_{\mathrm{eff}}) = \bigcup_{D \subset \Lambda} \mathrm{Spec}\left([\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D} + \sum_{i \in D} J_i \right), \quad (C10)$$

in which we count the eigenvalues with including their degeneracy. Here, $[\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D}$, described by

$$[\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D} = \sum_{i \notin D} \tilde{H}_i + \sum_{\alpha} \sum_{i \neq j: i, j \notin D} v_{ij}^{\alpha} \tilde{V}_i^{\alpha} \otimes \tilde{W}_j^{\alpha},$$

$$(C11)$$

and $\sum_{i \in D} J_i$ are considered as $(K^{N_{\mathrm{sub}}-|D|} \times M^{|D|})$-dimensional matrices defined on $\tilde{\mathcal{H}}_D$.

Proof. Let Q_i denote a projection operator to the auxiliary M-dimensional subspace on the ith subsystem:

$$Q_i \equiv (0_K \oplus I_M)_i. \qquad (C12)$$

Since $[Q_i, \tilde{H}_{\mathrm{eff}}] = 0$ and $[Q_i, Q_j] = 0$ are satisfied for any $i, j \in \Lambda$, the Hamiltonian $\tilde{H}_{\mathrm{eff}}$ can be block-diagonalized by adopting the eigenstates of Q_i as the basis. Then, each block becomes the $(K^{|D|} \times M^{N_{\mathrm{sub}}-|D|})$-dimensional Hilbert space $\tilde{\mathcal{H}}_D$ since the projection to $\tilde{\mathcal{H}}_D$ is

$$P_{\tilde{\mathcal{H}}_D} = \prod_{i \notin D} (1 - Q_i) \cdot \prod_{i \in D} Q_i. \qquad (C13)$$

As a result, the Hamiltonian $\tilde{H}_{\mathrm{eff}}$ on the extended $(K + M)^{N_{\mathrm{sub}}}$-dimensional Hilbert space $\tilde{\mathcal{H}}_{\mathrm{eff}}$ is decomposed as

$$\tilde{H}_{\mathrm{eff}} = \bigoplus_{D \subset \Lambda} \tilde{H}_{\mathrm{eff}}|_{\tilde{\mathcal{H}}_D}, \qquad (C14)$$

where $\tilde{H}_{\mathrm{eff}}|_{\tilde{\mathcal{H}}_D}$ is obtained by projecting $\tilde{H}_{\mathrm{eff}}$ to $\tilde{\mathcal{H}}_D$. When we consider the simplest case $D = \{i\}$, the restricted Hamiltonian

is given by

$$\tilde{H}_{\mathrm{eff}}|_{\tilde{\mathcal{H}}_D} = \sum_{j: j \neq i} \tilde{H}_j + J_i + \sum_{\alpha} \sum_{j,k: j,k \neq i} v_{jk}^{\alpha} \tilde{V}_j^{\alpha} \otimes \tilde{W}_k^{\alpha}. \quad (C15)$$

since $\tilde{H}_i$, $\tilde{V}_i^{\alpha}$, and $\tilde{W}_i^{\alpha}$ are respectively replaced to J_i, 0_M, and 0_M by the projection $P_{\tilde{\mathcal{H}}_D}$. We note that omitted identity operators in Eq. (C15) are properly chosen from I_K or I_M to give a matrix on $\tilde{\mathcal{H}}_D$ here. In a similar way, we can obtain those for a generic domain $D \subset \Lambda$ as

$$\tilde{H}_{\mathrm{eff}}|_{\tilde{\mathcal{H}}_D} = \sum_{i \notin D} \tilde{H}_i + \sum_{\alpha} \sum_{i \neq j: i, j \notin D} v_{ij}^{\alpha} \tilde{V}_i^{\alpha} \otimes \tilde{W}_j^{\alpha} + \sum_{i \in D} J_i$$

$$\equiv [\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D} + \sum_{i \in D} J_i. \qquad (C16)$$

This immediately results in Eq. (C10).

When we choose $D = \phi$ in Prop. 1, $[\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D} + \sum_{i \in D} J_i$ in Eq. (C10) is equivalent to the effective Hamiltonian $\tilde{H}$. Thus, all the eigenvalues of $\tilde{H}$ are embedded in those of $\tilde{H}_{\mathrm{eff}}$, though meaningless eigenvalues corresponding to $D \neq \phi$ appear.

Proposition 2. Let $\tilde{H}_{\mathrm{eff}}$ be obtained from $\tilde{H}$ by means of the insertion of auxiliary dimensions described in Eqs. (C2)–(C4). When we choose the $M \times M$ matrices $\{J_i\}_{i=1}^{N_{\mathrm{sub}}}$ by

$$J_i = \lambda_i I_M, \quad \lambda_i > e_{\tilde{H}}(i) : \text{extensiveness} \qquad (C17)$$

for each subsystem $i \in \Lambda$, the ground-state energy of $\tilde{H}$ is equal to that of $\tilde{H}_{\mathrm{eff}}$, that is,

$$E_0(\tilde{H}) = E_0(\tilde{H}_{\mathrm{eff}}) \qquad (C18)$$

is satisfied.

Proof. From the result of Prop. 1, it is sufficient to show that

$$E_0(\tilde{H}) < E_0\left([\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D} + \sum_{i \in D} J_i \right) \qquad (C19)$$

is satisfied for any domain $D \neq \phi$ when J_i is designated by Eq. (C17). With the usage of Eq. (C17), we obtain

$$[\text{r.h.s of Eq. (C19)}] = E_0([\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D}) + \sum_{i \in D} \lambda_i$$

$$> E_0([\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D}) + \sum_{i \in D} e_{\tilde{H}}(i).$$

When we explicitly write the omitted identity operators in Eq. (C11), $[\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D}$ becomes

$$\left[\sum_{i \notin D} \tilde{H}_i + \sum_{\alpha} \sum_{i \neq j: i, j \notin D} v_{ij}^{\alpha} \tilde{V}_i^{\alpha} \otimes \tilde{W}_j^{\alpha} \right] \otimes \left[\bigotimes_{i \in D} (I_M)_i \right]. \quad (C20)$$

If we replace $\bigotimes_{i \in D}(I_M)_i$ by $\bigotimes_{i \in D}(I_K)_i$, we obtain $\tilde{H} - \tilde{H}(D)$, represented by a $K^{N_{\mathrm{sub}}}$-dimensional matrix. This results in the same ground-state energy,

$$E_0([\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D}) = E_0(\tilde{H} - \tilde{H}(D)), \qquad (C21)$$

although their ground states have different degeneracy. Next, we exploit the following theorem for any Hermitian operators

H and H' with the same finite dimension (see Sec. III. 2 in Ref. [52]):

$$|E_m(H) - E_m(H')| \leqslant ||H - H'||_{\mathrm{op}} \qquad (\mathrm{C22})$$

for any integer $m \geqslant 0$. Then we obtain

$$E_0(\tilde{H} - \tilde{H}(D)) \geqslant E_0(\tilde{H}) - ||\tilde{H}(D)||_{\mathrm{op}}. \qquad (\mathrm{C23})$$

From the definition of $\tilde{H}(D)$ [see Eq. (C6)] and the extensiveness [see Eq. (C8)], we arrive at the following inequality:

$$
\begin{aligned}
||H(D)||_{\mathrm{op}} &\leqslant \sum_{X: X \cap D \neq \phi} ||h_X||_{\mathrm{op}} \\
&\leqslant \sum_{i \in D} \sum_{X: X \ni i} ||h_X||_{\mathrm{op}} = \sum_{i \in D} e_H(i) \quad (\mathrm{C24})
\end{aligned}
$$

for any Hamiltonian H on the lattice Λ. As a result, we obtain

$$
\begin{aligned}
[\text{r.h.s of Eq. (C19)}] &> E_0(\tilde{H}) - ||\tilde{H}(D)||_{\mathrm{op}} + \sum_{i \in D} e_{\tilde{H}}(i) \\
&\geqslant E_0(\tilde{H}) \qquad (\mathrm{C25})
\end{aligned}
$$

for $D \neq \phi$. Under this inequality, $\min\{\mathrm{Spec}(\tilde{H}_{\mathrm{eff}})\}$ is given by $E_0(\tilde{H})$, indicating Eq. (C18).

As discussed in the main text, this proposition is used for the modified way to construct the effective Hamiltonian $\tilde{H}_{\mathrm{eff}}$, which properly gives the ground state with arbitrary variational quantum states. In order to evaluate low-energy eigenstates, Prop. 2 is extended to the following proposition, which is referred to as **Proposition** in Sec. III B of the main text.

Proposition 3. We consider the same setup as the one in Prop. 2. For a certain integer $n \geqslant 0$, let us choose the $M \times M$ matrices $\{J_i\}_{i=1}^{N_{\mathrm{sub}}}$ by

$$J_i = \lambda_i I_M, \quad \lambda_i > e_{\tilde{H}}(i) + E_n(\tilde{H}) - E_0(\tilde{H}), \quad {}^{\forall}i \in \Lambda. \qquad (\mathrm{C26})$$

Then, the low-energy eigenvalues of $\tilde{H}$ are equal to those of $\tilde{H}_{\mathrm{eff}}$ respectively up to the nth level, that is,

$$E_m(\tilde{H}) = E_m(\tilde{H}_{\mathrm{eff}}), \quad m = 0, 1, \dots, n \qquad (\mathrm{C27})$$

is satisfied.

Proof. We prove this in the same way for Prop. 2. From the result of Prop. 2, we show that

$$E_n(\tilde{H}) < E_0\left([\tilde{H} - \tilde{H}(D)]_{\tilde{\mathcal{H}}_D} + \sum_{i \in D} J_i\right) \qquad (\mathrm{C28})$$

is satisfied for any domain $D \neq \phi$ when J_i is designated by Eq. (C26). By the same calculation as Eq. (C25), we obtain

$$
\begin{aligned}
[\text{r.h.s of Eq. (C28)}] &> E_0(\tilde{H}) - ||\tilde{H}(D)||_{\mathrm{op}} + \sum_{i \in D} \lambda_i \\
&> E_0(\tilde{H}) + |D|\{E_n(\tilde{H}) - E_0(\tilde{H})\} \\
&\geqslant E_n(\tilde{H}) \qquad (\mathrm{C29})
\end{aligned}
$$

for $D \neq \phi$. We use $|D| \geqslant 1$ under $D \neq \phi$. Under this inequality and Eq. (C10), Eq. (C27) is satisfied. ∎

Note that we should know the gap $E_n(\tilde{H}) - E_0(\tilde{H})$ in advance when we construct $\tilde{H}_{\mathrm{eff}}$ following Eq. (C26) different from the case when evaluating the ground state. In practice,

we know typical energy scale of the gap by some other methods in advance, and can properly choose J_i as discussed in the main text. From the mathematical point of view, we can construct $\tilde{H}_{\mathrm{eff}}$ without the knowledge of the gap $E_n(\tilde{H}) - E_0(\tilde{H})$ in advance even when we evaluate excited states of $\tilde{H}$, by the following proposition.

Proposition 4. We consider the same setup as the one in Prop. 2. Let us choose the $M \times M$ matrices $\{J_i\}_{i=1}^{N_{\mathrm{sub}}}$ by

$$J_i = \lambda_i I_M, \quad \lambda_i > e_{\tilde{H}}(i) + 2\sum_{j \in \Lambda} e_{\tilde{H}}(j), \quad {}^{\forall}i \in \Lambda. \qquad (\mathrm{C30})$$

Then, all of the eigenvalues of $\tilde{H}$ appear in the beginning of $\mathrm{Spec}(\tilde{H}_{\mathrm{eff}})$, that is,

$$E_m(\tilde{H}) = E_m(\tilde{H}_{\mathrm{eff}}), \quad m = 0, 1, \dots, K^{N_{\mathrm{sub}}} - 1 \qquad (\mathrm{C31})$$

is satisfied.

Proof. Under the choice of Eq. (C30),

$$
\begin{aligned}
\lambda_i &> e_{\tilde{H}}(i) + 2||\tilde{H}||_{\mathrm{op}} \\
&\geqslant e_{\tilde{H}}(i) + E_{K^{N_{\mathrm{sub}}}-1}(\tilde{H}) - E_0(\tilde{H}) \qquad (\mathrm{C32})
\end{aligned}
$$

is satisfied. We use the result of Prop. 3 with $n = K^{N_{\mathrm{sub}}} - 1$, and then we obtain Eq. (C31). ∎

We note that, although Prop. 4 provides a mathematically right way to evaluate excited states without the knowledge of the gap in advance, it is not practical compared to the one exploiting Prop. 3. When we choose J_i by Eq. (C30), its energy scale becomes linear in the system size $N_{\mathrm{sub}} \times N_{\mathrm{qubit}}$. When considering a large system, J_i becomes dominant in the effective Hamiltonian $\tilde{H}_{i,\mathrm{eff}}$ [see Eqs. (C2)–(C4)] compared to the original one $\tilde{H}_i$. Then, errors originating from J_i buries the information of $\tilde{H}_i$, which is genuinely of interest, and hence the insertion of auxiliary dimensions based on Prop. 4 is not practical, though it is mathematically correct.

We finally discuss the generalization of our results. In the above discussion, we designate the decomposition of the Hamiltonian $\tilde{H}$ by Eq. (C1). When we choose another decomposition, we obtain a different extended Hamiltonian $\tilde{H}_{\mathrm{eff}}$ [Eqs. (C2)–(C4)] and a different extensiveness $e_{\tilde{H}}(i)$ [Eq. (C9)]. However, the same results, Props. 1–4, are valid due to the relations Eqs. (C21) and (C24). Thus, the decomposition is not essential. As well, we can consider intersubsystem interactions which simultaneously involve more than three subsystems. In such cases, we introduce auxiliary dimensions for intrasubsystem terms by a nonzero matrix J_i and for intersubsystem terms by a zero matrix $0_{2^{N_{\mathrm{eff}}}-K}$ as well as Eqs. (C2)–(C4). The extensiveness $e_H(i)$ is given by the same definition Eq. (C8). Then, using the relations Eqs. (C21) and (C24), Props. 1–4 are valid in the cases with interactions involving more than three subsystems.

APPENDIX D: DETAILS OF NUMERICAL SIMULATIONS

Here, we describe the details of numerical simulations in the main text, especially focusing on the variational quantum circuits and the cost function. First, we identify the variational quantum circuits in VQE in Step 1 and Step 3 of Deep VQE. We employ the hardware-efficient-type ansatz [20,37], where the parametric circuit with the depth D is
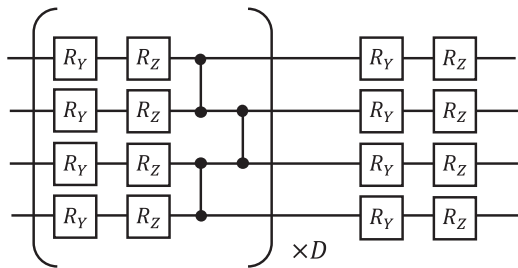
FIG. 5. Variational quantum circuit based on Hardware-efficient ansatz. We prepare the reference states by $|0\ldots00\rangle$ for the VQE and $|0\ldots00\rangle$, $|0\ldots01\rangle$ for the SSVQE.

described by

$$U(\vec{\theta}) = U_1(\vec{\theta}_{D+1}, \vec{\theta}'_{D+1}) \prod_{d=1}^{D} [U_2 \cdot U_1(\vec{\theta}_d, \vec{\theta}'_d)], \quad \text{(D1)}$$

$$U_1(\vec{\theta}_d, \vec{\theta}'_d) = \prod_{n=1}^{N} [R_{Z,n}(\theta'_{d,n}) \cdot R_{Y,n}(\theta_{d,n})], \quad \text{(D2)}$$

$$U_2 = \prod_{n=1}^{N-1} CZ_{n,n+1}, \quad \text{(D3)}$$

for a N-qubit system (see Fig. 5). Here, $R_{Y,n}(\theta)$ [$R_{Z,n}(\theta)$] represents rotation around Y-axis [Z-axis] with an angle θ on the nth qubit, and $CZ_{n,m}$ is a controlled-Z gate on the nth and mth qubits. The parameter set $\vec{\theta}$ is composed of $\{\vec{\theta}_d\}_{d=1}^{D+1}$ and $\{\vec{\theta}'_d\}_{d=1}^{D+1}$, whose initial values are chosen uniformly from $[0, 2\pi)^{2N(D+1)}$ at random. When we evaluate the ground state in Step 1, we employ the cost function $\langle 0\ldots0|U(\vec{\theta})^\dagger H_i U(\vec{\theta})|0\ldots0\rangle$ for each subsystem-Hamiltonian H_i. For evaluating the first-excited states in Step 3 of Deep VQE, we employ SSVQE [7], in which the cost function is given by

$$w_0 \langle 0\ldots00|U(\vec{\theta})^\dagger \tilde{H}_{\mathrm{eff}} U(\vec{\theta})|0\ldots00\rangle$$
$$+ w_1 \langle 0\ldots01|U(\vec{\theta})^\dagger \tilde{H}_{\mathrm{eff}} U(\vec{\theta})|0\ldots01\rangle$$

with $w_0 > w_1 > 0$. After the optimization, we can obtain the approximate ground (first-excited) state by $U(\vec{\theta}^*)|0\ldots00\rangle$ $(U(\vec{\theta}^*)|0\ldots01\rangle)$. We employ the BFGS optimizer for optimizing the circuit parameters with numerical differentiation.

For the spin chain in Sec. III D, we set the depth D by 10,15,20 for 4,6,8-qubit systems in Step 1, and by 15,20,25 for 4,6,8-qubit systems in Step 3, respectively. The set of weights for SSVQE is $(w_0, w_1) = (2, 1)$. For the periodic hydrogen chain in Sec. IV, we employ the depth $D = 20$ in Step 1 and $D = 80$ in Step 3, and choose the weights by $(w_0, w_1) = (7, 2)$.

APPENDIX E: RELATION BETWEEN DEEP VQE WITH HIGHER-ORDER EXCITATION AND QSE RESULTS

In Sec. IV B, the local basis by the double-particle excitation $\mathcal{W}_d$ gives better low-energy eigenstates than the one by single-particle excitation $\mathcal{W}_s$. Here, we discuss the role of higher-order excitations in Deep VQE. Concretely, we show that Deep VQE with higher-order excitations overwhelms the combination of QSE and Deep VQE with lower-order

excitations—calculating the ground state by Deep VQE with lower-order excitations and performing QSE on the obtained ground state with lower-order excitations. This gives the advantage of considering higher-order excitations in Deep VQE.

Let us compare the above two methods by the active subspaces in them. For simplicity, we consider single-particle (the first-order) and double-particle (the second-order) excitations. First, we focus on the combination of QSE and Deep VQE with single-particle excitation $\mathcal{W}_s$ [see Eq. (27)]. The Deep VQE with $\mathcal{W}_s$ searches low-lying states from the subspace $\tilde{\mathcal{H}}_s$, spanned by

$$|\Psi_0\rangle, \ R_i|\Psi_0\rangle, \ R_iR_j|\Psi_0\rangle, \ R_iR_jR_k|\Psi_0\rangle, \ldots. \quad \text{(E1)}$$

Here, R_i is given by c'_i or $c'^\dagger_i$ and the indices $i, j, k, \ldots$ represent certain sites of the whole system belonging to different subsystems each other. Let $|\Psi_0^s(\text{Deep VQE})\rangle \in \tilde{\mathcal{H}}_s$ denote the approximate ground state by the Deep VQE with $\mathcal{W}_s$. The QSE with the reference state $|\Psi_0^s(\text{Deep VQE})\rangle$ and the single-particle excitations in the whole-system $R_i = c'_i$ or $c'^\dagger_i$ searches low-energy eigenstates from the subspace

$$\tilde{\mathcal{H}}_s^{\mathrm{QSE}} = \mathrm{span}\big(\big\{\,|\Psi_0^s(\text{Deep VQE})\rangle, \ R_i\,|\Psi_0^s(\text{Deep VQE})\rangle\,\big\}\big). \quad \text{(E2)}$$

On the other hand, we next focus on the Deep VQE result with the complete double-particle excitations,

$$\mathcal{W}'_d = \mathcal{W}_s \cup \{c'^\dagger_i c'_j, \ c'_i c'_j, \ c'^\dagger_i c'^\dagger_j \,|\, i, j \in \Lambda_e\}, \quad \text{(E3)}$$

By construction, the restricted subspace $\tilde{\mathcal{H}}'_d$, which is active in the Deep VQE with $\mathcal{W}'_d$, includes not only states in $\tilde{\mathcal{H}}_s$ but also

$$R_i|\Psi\rangle \quad \text{for any state} \quad |\Psi\rangle \in \tilde{\mathcal{H}}_s, \quad \text{(E4)}$$

with single-particle excitations $R_i = c'_i$ or $c'^\dagger_i$. Considering the fact $|\Psi_0^s(\text{Deep VQE})\rangle \in \tilde{\mathcal{H}}_s$, we obtain

$$\tilde{\mathcal{H}}_s^{\mathrm{QSE}} \subset \tilde{\mathcal{H}}'_d. \quad \text{(E5)}$$

This means that the Deep VQE results with double-particle excitations are more accurate than the QSE results on the Deep VQE with single-particle excitations. We stress that this statement on the difference of the accuracy between single- and double-particle excitations is more strict than the one by simply comparing the Deep VQE results with $\mathcal{W}_s$ and $\mathcal{W}'_d$ from the relation $\tilde{\mathcal{H}}_s \subset \tilde{\mathcal{H}}'_d$.

In general, the Deep VQE results with local excitations up to nth order always give better low-lying energy eigenvalues than the results obtained by the QSE with up to $(n - m)$th order excitations after performing the Deep VQE with up to mth order excitations ($0 \leqslant m \leqslant n$). Although our calculation with $\mathcal{W}_d$ in the main text [Eq. (28)] does not strictly satisfy the above discussion due to the absence of $c'_i c'_j$ and $c'^\dagger_i c'^\dagger_j$, we expect that the restricted space by $\mathcal{W}_d$, similar to $\tilde{\mathcal{H}}'_d$, explains the improved accuracy compared to the single-particle excitations $\mathcal{W}_s$. For larger systems, we can exploit a complete set of higher-order excitations for constructing the local basis with keeping the reduction of qubits. In that case, we expect much improved accuracy of the Deep VQE results by higher-order excitations from their superiority to the combination of QSE and Deep VQE with lower-order excitations.

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

[2] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, Quantum optimization using variational algorithms on near-term quantum devices, Quantum Sci. Technol. **3**, 030503 (2018).

[3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[4] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[5] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[6] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Phys. Rev. A **95**, 042308 (2017).

[7] K. M. Nakanishi, K. Mitarai, and K. Fujii, Subspace-search variational quantum eigensolver for excited states, Phys. Rev. Research **1**, 033062 (2019).

[8] R. M. Parrish, E. G. Hohenstein, P. L. McMahon, and T. J. Martínez, Quantum Computation of Electronic Transitions Using a Variational Quantum Eigensolver, Phys. Rev. Lett. **122**, 230401 (2019).

[9] O. Higgott, D. Wang, and S. Brierley, Variational quantum computation of excited states, Quantum **3**, 156 (2019).

[10] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Variational quantum algorithms for discovering Hamiltonian spectra, Phys. Rev. A **99**, 062304 (2019).

[11] P. J. Ollitrault, A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, and I. Tavernelli, Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor, Phys. Rev. Research **2**, 043140 (2020).

[12] S. Endo, I. Kurata, and Y. O. Nakagawa, Calculation of the Green's function on near-term quantum computers, Phys. Rev. Research **2**, 033281 (2020).

[13] I. Rungger, N. Fitzpatrick, H. Chen, C. H. Alderete, H. Apel, A. Cowtan, A. Patterson, D. M. Ramo, Y. Zhu, N. H. Nguyen *et al.*, Dynamical mean field theory algorithm and experiment on quantum computers, arXiv:1910.04735.

[14] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, Phys. Rev. X **7**, 021050 (2017).

[15] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum **3**, 191 (2019).

[16] S. Endo, J. Sun, Y. Li, S. C. Benjamin, and X. Yuan, Variational Quantum Simulation of General Processes, Phys. Rev. Lett. **125**, 010501 (2020).

[17] N. Yoshioka, Y. O. Nakagawa, K. Mitarai, and K. Fujii, Variational quantum algorithm for nonequilibrium steady states, Phys. Rev. Research **2**, 043289 (2020).

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

[19] 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).

[20] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature (London) **549**, 242 (2017).

[21] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, Computation of Molecular Spectra on a Quantum Processor with an Error-Resilient Algorithm, Phys. Rev. X **8**, 011021 (2018).

[22] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature (London) **567**, 491 (2019).

[23] K. Mitarai, Y. O. Nakagawa, and W. Mizukami, Theory of analytical energy derivatives for the variational quantum eigensolver, Phys. Rev. Research **2**, 013129 (2020).

[24] T. E. O'Brien, B. Senjean, R. Sagastizabal, X. Bonet-Monroig, A. Dutkiewicz, F. Buda, L. DiCarlo, and L. Visscher, Calculating energy derivatives for quantum chemistry on a quantum computer, npj Quantum Inf. **5**, 113 (2019).

[25] S. Tamiya, S. Koh, and Y. O. Nakagawa, Calculating nonadiabatic couplings and Berry's phase by variational quantum eigensolvers, Phys. Rev. Research **3**, 023244 (2021).

[26] J. Liu, L. Wan, Z. Li, and J. Yang, Simulating periodic systems on a quantum computer using molecular orbitals, J. Chem. Theory Comput. **16**, 6904 (2020).

[27] D. Z. Manrique, I. T. Khan, K. Yamamoto, V. Wichitwechkarn, and D. M. Ramo, Momentum-space unitary coupled cluster and translational quantum subspace expansion for periodic systems on quantum computers, arXiv:2008.08694.

[28] N. Yoshioka, Y. O. Nakagawa, Y. ya Ohnishi, and W. Mizukami, Variational quantum simulation for periodic materials, arXiv:2008.09492.

[29] 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.

[30] W. Yang and T. Lee, A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules, J. Chem. Phys. **103**, 5674 (1995).

[31] M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, Fragmentation methods: A route to accurate calculations on large systems, Chem. Rev. **112**, 632 (2012).

[32] T. Yamazaki, S. Matsuura, A. Narimani, A. Saidmuradov, and A. Zaribafiyan, Towards the practical application of near-term quantum computers in quantum chemistry simulations: A problem decomposition approach, arXiv:1806.01305.

[33] Y. Kawashima, M. P. Coons, Y. Nam, E. Lloyd, S. Matsuura, A. J. Garza, S. Johri, L. Huntington, V. Senicourt, A. O. Maksymov *et al.*, Efficient and accurate electronic structure simulation demonstrated on a trapped-ion quantum computer, arXiv:2102.07045.

[34] M. Sajjan, S. H. Sureshbabu, and S. Kais, Quantum machine-learning for eigenstate filtration in two-dimensional materials, J. Am. Chem. Soc. **143**, 18426 (2021).

[35] H. Werner and E. Reinsch, The self-consistent electron pairs method for multiconfiguration reference state functions, J. Chem. Phys. **76**, 3144 (1982).

[36] H. Werner and P. J. Knowles, An efficient internally contracted multiconfiguration–reference configuration interaction method, J. Chem. Phys. **89**, 5803 (1988).

[37] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning, Phys. Rev. A **98**, 032309 (2018).

[38] Qulacs (2018), https://github.com/qulacs/qulacs.

[39] 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, Quantum **5**, 559 (2021).

[40] J. R. McClean, N. C. Rubin, K. J. Sung, I. D. Kivlichan, X. Bonet-Monroig, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby *et al.*, OpenFermion: The electronic structure package for quantum computers, Quantum Sci. Technol. **5**, 034014 (2020).

[41] K. Bharti and T. Haug, Iterative quantum assisted eigensolver, arXiv:2010.05638.

[42] P. Jordan and E. Wigner, Über das paulische äquivalenzverbot, Z. Phys. **47**, 631 (1928).

[43] 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).

[44] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, PySCF: the Python-based simulations of chemistry framework, WIREs Comput. Mol. Sci. **8**, e1340 (2018).

[45] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui *et al.*, Recent developments in the PySCF program package, J. Chem. Phys. **153**, 024109 (2020).

[46] G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, Phys. Rev. Lett. **91**, 147902 (2003).

[47] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, Phys. Rev. Lett. **93**, 040502 (2004).

[48] U. Schollwöck, The density-matrix renormalization group, Rev. Mod. Phys. **77**, 259 (2005).

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

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

[51] S. McArdle, X. Yuan, and S. Benjamin, Error-Mitigated Digital Quantum Simulation, Phys. Rev. Lett. **122**, 180501 (2019).

[52] R. Bhatia, *Matrix Analysis* (Springer, New York, 1997).