

Correlation-Informed Permutation of Qubits for Reducing Ansatz Depth in the Variational Quantum Eigensolver

Nikolay V. Tkachenko^{1,§}, James Sud^{2,3,§}, Yu Zhang^{3,*}, Sergei Tretiak³, Petr M. Anisimov⁴, Andrew T. Arrasmith³, Patrick J. Coles³, Lukasz Cincio^{3,†} and Pavel A. Dub^{1,‡}

¹Chemistry Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

²Department of Physics, University of California, Berkeley, Berkeley, California 94720, USA

³Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁴Accelerators and Electrodynamics Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA



(Received 14 September 2020; revised 22 January 2021; accepted 19 April 2021; published 9 June 2021)

The variational quantum eigensolver (VQE) is a method of choice to solve the electronic structure problem for molecules on near-term gate-based quantum computers. However, the circuit depth is expected to grow significantly with the problem size. Increased depth can both degrade the accuracy of the results and reduce trainability. In this work, we propose an approach to reduce ansatz circuit depth. Our approach, called “PermVQE,” adds an additional optimization loop to the VQE that permutes qubits in order to solve for the qubit Hamiltonian that maximally localizes correlations in the ground state. The choice of permutations is based on mutual information, which is a measure of interaction between electrons and/or holes in spin-orbitals. Encoding strongly entangled spin-orbitals into proximal qubits on a quantum chip naturally reduces the circuit depth needed to prepare the ground state. For representative molecular systems, LiH, H₂, (H₂)₂, H₄⁺, H₃⁺, and N₂, we demonstrate that placing entangled qubits in close proximity leads to shallower depth circuits required to reach a given eigenvalue-eigenvector accuracy. The approach is designed for hardware-efficient ansatz of any qubit connectivity, and examples are demonstrated for linear and two-dimensional grid architectures. The main ideas can also be applied to simulate molecules with other ansatz as well as variational quantum algorithms beyond the VQE. In particular, we demonstrate the beneficial effect of qubit permutations to build fermionic-adaptive derivative assembled pseudo-Trotter ansatz on a linear qubit connectivity architecture with nearly a twofold reduction of the number of controlled NOT gates.

DOI: [10.1103/PRXQuantum.2.020337](https://doi.org/10.1103/PRXQuantum.2.020337)

I. INTRODUCTION

Quantum computing is expected to revolutionize computational chemistry by achieving polynomial scaling in both the number of quantum particles and the quality of the description of the system (e.g., the number of orbital basis functions or numerical grid points) [1,2]. The variational quantum eigensolver (VQE) has emerged as a viable algorithm [3] to find asymptotically exact lowest

eigenvalues for solutions to the Schrödinger equation on noisy intermediate-scale quantum (NISQ) devices. VQE-based ground-state electronic energy calculations of small molecular systems (e.g., H₂, BeH₂, H₂O, alkali metal hydrides, and H₁₂) in minimal basis sets (e.g., the contracted Gaussian STO-*n*G family) were experimentally implemented using superconducting circuits [4–8] and trapped ions [9] as physical qubits. As the age of quantum supremacy dawns [10], demonstrating quantum advantage for chemistry will naturally involve considering larger molecules and/or basis sets. In addition to quantum circuit width (number of qubits), this will, in turn, increase the quantum circuit depth, which typically grows polynomially with the problem size [11].

The growing circuit depth with problem size causes two main issues. One issue is the accumulation of hardware noise, which impacts the accuracy of the results. The other issue is the trainability of the ansatz parameters (i.e., whether the parameters have large enough gradients to

*zhy@lanl.gov

†lcincio@lanl.gov

‡pdub@lanl.gov

§These authors contributed equally.

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

allow progress in the optimization). This concern arises since increased circuit depth leads to smaller gradients [12–18], and accurate estimation of small gradients on a quantum computer requires a large number of runs (or “shots”). These two issues are related as the accumulation of hardware noise also leads to smaller gradients [19]. This highlights the importance of keeping the circuit depth shallow in variational ansatz.

Some strategies partially address these problems, such as improved classical optimizers [20–22], parameter initialization strategies [23–25], error mitigation [26–28], and noise resilience [29]. On the other hand, the most direct strategy would be to somehow reduce ansatz circuit depth. Some promising “iterative VQE” approaches have been proposed for this purpose, largely focusing on adaptive ansatz [30–40]. Given that reducing ansatz depth will reduce the training complexity and improve the accuracy of the VQE, it is a crucial research direction that could bring us closer to realizing quantum advantage for electronic structure calculations.

It is worth emphasizing that lack of complete qubit connectivity on NISQ devices [41] makes it necessary to increase the depth of quantum circuits. Thus, any efforts to minimize ansatz depth in the VQE should account for the connectivity of the specific NISQ hardware [42]. For instance, if two qubits are highly entangled in the exact solution of an electronic structure problem, but the qubits are physically distant on a device with limited connectivity, a deeper circuit is required for accurate simulation. Long sequences of noisy two-qubit gates are necessary to entangle distant qubits, which can degrade the fidelity and reduce trainability.

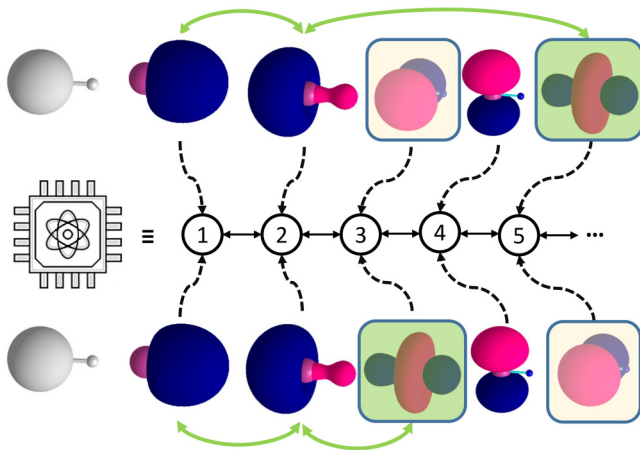


FIG. 1. Correlation-informed permutation schematics for the example of a LiH molecule in a reduced active space of the STO-3G basis: direct encoding of five α spin-orbitals into five physical qubits on a linear architecture quantum chip before (top) and after (bottom) permutation based on orbital entanglement information (green double-headed arrows show strongly interacting orbitals).

In this work, we introduce an approach to reduce ansatz depth by permuting the pattern with which the Hamiltonian is embedded onto qubits. Specifically, for hardware-efficient ansatz developed in the context of electronic structure [5,43,44], we use a correlation-informed approach, where the permutation is chosen on the basis of the mutual information defining entanglement between two individual spin-orbitals [45,46]. Our main idea is illustrated in Fig. 1 for the example of selected spin-orbitals of a LiH molecule. Entanglement of spin-orbitals triggers permutation of the embedding to ensure that these orbitals are encoded into qubits that are placed physically close to each other on an actual quantum chip.

We numerically implement such qubit permutations assuming a linear as well as two-dimensional (2D) grid architecture qubit connectivity for the LiH, H₂, (H₂)₂, H₄[±], H₃[±], and N₂ molecular systems. In all cases, permutations significantly reduce the circuit depth required to reach a given energy accuracy, relative to the unpermuted case. To realize this permutation technique, we introduce the PermVQE algorithm, an added layer on top

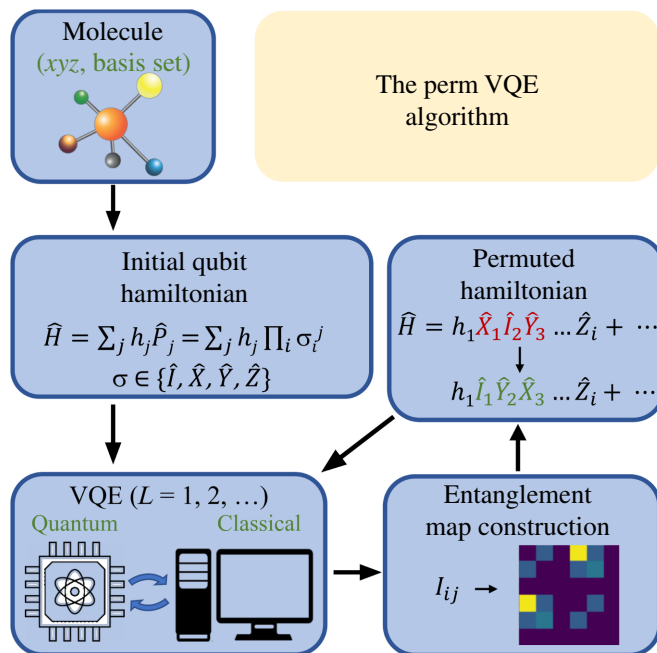


FIG. 2. The PermVQE algorithm. After the approximate ground-state wave function is obtained from the initial Hamiltonian at a fixed circuit depth L through the VQE, the entanglement map is produced on the basis of mutual information I_{ij} . Minimization of a cost function (which quantifies the amount of long-range correlations) over possible permutations of qubits returns the permuted Hamiltonian, which is then fed back into the VQE algorithm. This procedure is iteratively repeated until the best-permuted Hamiltonian is generated, from which the best-possible (within L) ground-state wave function and energy are obtained.

of the original VQE posed in Ref. [3]. At a fixed circuit depth L , this algorithm, as depicted in Fig. 2, starts from the approximate ground-state wave function of a given initial Hamiltonian, and then uses it to calculate an optimal reordering of qubits and builds a permuted Hamiltonian. The algorithm iteratively repeats this process until an adequate reordering of qubits is obtained, and the best-permuted qubit Hamiltonian is used to variationally calculate the best-possible (within L) ground-state wave function and associated energy.

We also demonstrate the beneficial effect of qubit permutations to build fermionic-adaptive derivative assembled pseudo-Trotter (ADAPT) ansatz [30] for the implementation on a linear nearest-neighbor quantum chip architecture [47].

II. THE Perm VQE ALGORITHM

A. Overview

A diagram of the PermVQE algorithm is shown in Fig. 2. As in the standard VQE for electronic structure, the first step is the initial mapping of the fermionic Fock-space states as well as creation and annihilation operators into Hilbert-space states and Pauli operators of qubits on the basis of one of the established second-quantized encoding methods; see below. This is followed by a standard VQE loop to variationally learn an approximate correlated ground-state wave function at a fixed circuit depth L . The next step is to perform local tomography on two-qubit subsets so as to generate a mutual-information matrix, which we call the “entanglement map.” This step also minimizes a cost function that quantifies the amount of long-range correlations (for the specific hardware’s connectivity), while varying the qubit index labels. As a result we arrive at an optimal permutation of the qubits and hence a new Hamiltonian, which is then fed back to the VQE subroutine for another iteration. We now provide more details on these various subroutines.

B. Initial qubit mapping

In PermVQE, the first step is an initial spin-orbital to qubit mapping, subject to a chosen transformation [1,2]. As discussed in Sec. III, the performance of PermVQE can be affected by the initial qubit mapping. In particular, it may be desirable to select an initial qubit mapping that leads to a sparser entanglement map (i.e., minimizing the number of qubits being entangled). This would then allow qubit permutations to have a more significant effect. In Sec. III, we compare three popular second-quantized basis set encoding methods: the Jordan-Wigner (JW) [48], Bravyi-Kitaev (BK) [49], and parity [50,51] methods. We find that the Jordan-Wigner transformation, as expected, facilitates a sparser entanglement map, and hence improves the performance of PermVQE.

C. Entanglement map

After the initial qubit Hamiltonian is obtained and the VQE is run under fixed circuit depth L , the next step of PermVQE is to produce an entanglement map reflecting electronic correlations in the approximate ground state. For this purpose, we calculate the quantum mutual information for all pairs of qubits, which provides a measure of the total correlation, including both quantum and classical correlations. Previous results obtained for the orbital ordering problem in the density matrix renormalization group method in classical quantum chemistry calculations [45] showed that the quantum mutual information is a reliable parameter to quantify the correlation between two quantum particles. The quantum mutual information between qubits i and j is defined as follows:

$$I_{ij} = \frac{1}{2}(S_i + S_j - S_{ij})(1 - \delta_{ij}), \quad (1)$$

where S_i and S_{ij} are the single-qubit and two-qubit von Neumann entropies, respectively. The Kronecker δ sets all diagonal elements I_{ii} to zero. The single-qubit von Neumann entropy S_i is given by

$$S_i = -\text{Tr}(\rho_i \log \rho_i) = -\sum_{\alpha} \lambda_i^{(\alpha)} \log \lambda_i^{(\alpha)}, \quad (2)$$

and S_{ij} is defined analogously. Here ρ_i is the reduced one-body density matrix of qubit i and $\{\lambda_i^{(\alpha)}\}_{\alpha}$ are its eigenvalues.

We note that ρ_{ij} (and its marginals ρ_i and ρ_j) can be obtained from two-qubit tomography, which involves measuring 15 Pauli operators. There are $n(n-1)/2$ pairs of qubits, with n being the number of qubits. Hence, the number of operator measurements needed to construct the entanglement map is $15n(n-1)/2$. Some of these operators will commute and thus can be measured simultaneously.

On the basis of the mutual-information values for each pair of qubits, we build an $n \times n$ matrix. This matrix $\mathbf{I} = \{I_{ij}\}$, called the “entanglement map,” is useful to illustrate the amount and the length scale of the correlations in the approximate ground state. Moreover, we use it to define a cost function, as described next.

D. Cost function

To quantify the amount of long-range correlations, we introduce a cost function as follows. For a NISQ device with a given connectivity, let d_{ij} denote the distance between qubits i and j , which can be precisely defined as the number of edges in the shortest path through the connectivity graph between these qubits. Alternatively, one can view d_{ij} as the minimum number of SWAP gates (plus one) needed to make qubits i and j nearest neighbors.

Then, for any qubit connectivity, a cost function can be defined as

$$C(\mathbf{I}) = \sum_{i < j} f(d_{ij}) I_{ij}, \quad (3)$$

where $f(\cdot)$ is a monotonously increasing function of d_{ij} . As an example, one could choose a power-law function: $f(d_{ij}) = d_{ij}^\beta$ for some $\beta > 0$. When $\beta = 2$, we have

$$C(\mathbf{I}) = \sum_{i < j} d_{ij}^2 I_{ij}, \quad (4)$$

being the choice of cost function used in our numerical implementations below. Specifically, most of the examples that we discuss here implement a linear connectivity. In this case, we have $d_{ij} = |i - j|$, and the cost function becomes $C(\mathbf{I}) = \sum_{i < j} |i - j|^2 I_{ij}$. We use $d_{i,j} = |i_1 - j_1| + |i_2 - j_2|$, where $i = (i_1, i_2)$ and $j = (j_1, j_2)$ in the example that studies the N_2 molecule on a quantum chip with a square-lattice, nearest-neighbor qubit connectivity; see Sec. III D. The choice of the cost function is dictated by the optimization procedure described below.

E. Cost-function minimization

We define a permutation P as a bijection from the set of qubit indices to itself. The action of P will affect the entanglement map \mathbf{I} , and we are interested in solving the optimization problem:

$$P_{\text{opt}} = \arg \min_P C(P\mathbf{I}P^{-1}). \quad (5)$$

After solving this optimization problem, the PermVQE algorithm uses P_{opt} to permute the qubit indices and produce a new Hamiltonian. For example, if the original Hamiltonian is $\hat{H} = \hat{X}_1 \hat{Y}_2 \hat{Z}_3$ and P_{opt} transforms indices as $1 \mapsto 3, 2 \mapsto 2, \text{ and } 3 \mapsto 1$, then we produce the permuted Hamiltonian $\hat{H} = \hat{X}_3 \hat{Y}_2 \hat{Z}_1$.

Several methods are suitable for minimization of $C(P\mathbf{I}P^{-1})$ over qubit permutations. In principle, one could take a brute-force approach. That is, one could explicitly construct the $n!$ different permutations and check which one of them produces the lowest value of $C(P\mathbf{I}P^{-1})$. While this approach will certainly find the best permutation for a given entanglement map, the computational cost grows factorially in n , and hence scalability to large problem sizes is problematic. Instead, we focus on more practical techniques, as described below.

1. Spectral graph algorithm

The mutual-information matrix \mathbf{I} can be considered as a weighted graph where the weighted edge represents the mutual information between two qubits. In particular, the minimization of a quadratic cost function, $C(\mathbf{I}) =$

$\sum_{i < j} |i - j|^2 I_{ij}$, can be related to spectral graph theory [52,53]. For a given \mathbf{I} , we can define the graph Laplacian \mathbf{L} as follows:

$$\mathbf{L} = \mathbf{D} - \mathbf{I}, \quad D_{ij} = \delta_{ij} \sum_k I_{ik}. \quad (6)$$

Spectral graph theory has shown [52] that the Fiedler vector, or the second smallest eigenvector of \mathbf{L} , is the solution that can provide low values of the following function:

$$F(x) = x^\dagger \mathbf{L} x = \sum_{ij} I_{ij} (x_i - x_j)^2, \quad (7)$$

which is exactly the cost function that we defined above. Sorting the entries of the Fiedler vector in either ascending or descending order provides the optimized ordering of qubits [52]. Note that the diagonalization of the Laplacian matrix \mathbf{L} scales polynomially with the number of qubits n . Hence, the cost-function minimization based on the spectral graph algorithm provides an efficient way of finding optimal qubit ordering, with polynomial scaling in n .

The entries of the mutual-information matrix in Eq. (6) are known only to some precision. This is because one works with a finite number of shots when performing tomography. If the mutual-information matrix were known exactly, the matrix \mathbf{L} would be positive definite. This condition could be enforced by shifting the spectrum properly for imperfect mutual-information matrices. This procedure may lead to some numerical instabilities and hence we propose an alternative method as described below.

2. Heuristic permutation search

In this section we describe an alternative method to find (approximate) solutions of the minimization problem of Eq. (5). The method is iterative and is typically initialized by a random permutation. That permutation is subsequently modified during the optimization procedure that minimizes $C(P\mathbf{I}P^{-1})$. The method keeps track of the current permutation P_{curr} as well as the best overall permutation P_{best} . The approach tries to modify P_{curr} to further lower the value of $C(P\mathbf{I}P^{-1})$.

At a given iteration, the method proposes an update of P_{curr} denoted by P_m . The update is random but the preference is given to small corrections. The modification is implemented with some number of transpositions that are applied to P_{curr} . The number of transpositions is random and is chosen from an exponentially decaying probability distribution. If the updated permutation leads to a smaller cost [i.e., if $C(P_m \mathbf{I} P_m^{-1}) < C(P_{\text{curr}} \mathbf{I} P_{\text{curr}}^{-1})$], it is accepted as the new current permutation and the optimization moves to the next iteration. If $C(P_m \mathbf{I} P_m^{-1}) > C(P_{\text{curr}} \mathbf{I} P_{\text{curr}}^{-1})$, P_m is

also accepted with probability

$$\exp \left\{ - \left[C(P_m \mathbf{I} P_m^{-1}) - C(P_{\text{curr}} \mathbf{I} P_{\text{curr}}^{-1}) \right] / T \right\}, \quad (8)$$

following the standard simulated annealing approach. Here T is the temperature, which influences the likelihood of accepting permutation leading to a higher cost. It is lowered during the optimization to make the probability of accepting worse permutation smaller. The algorithm also records the best-seen permutation P_{best} and compares it with P_{curr} at the end of the search.

The iterations are repeated until convergence in $C(P \mathbf{I} P^{-1})$ is found. The optimization is restarted multiple times to minimize the impact of getting trapped in a local minimum.

We benchmark our algorithm against a brute-force search approach for all systems studied in this paper except for the 16-qubit N_2 molecule. The largest search is performed for the 14-qubit LiH molecule. In all cases, it takes up to 20 s to find permutation that is within 0.01% of the best solution found by brute force (as measured by the cost function C). In most cases, several thousand iterations are needed to identify a close-to-optimal solution of Eq. (5). The performance of our heuristic approach depends on the temperature schedule and can be further improved by modifying it on a case-by-case basis. The permutation search algorithm described in this section runs in a much shorter time than the VQE step that is performed later.

We also point out that PermVQE does not rely on finding the optimal solution of Eq. (5). Any permutation P that lowers the cost function [i.e., such that $C(P \mathbf{I} P^{-1}) < C(\mathbf{I})$] will lead to a better-performing ansatz. Lack of optimality in a given iteration is typically compensated in the following iteration, when the approximation to the ground state is improved and the matrix \mathbf{I} contains more accurate information about the correlations in the ground state.

III. NUMERICAL IMPLEMENTATIONS

A. Ising toy models with exact wave functions

To demonstrate the effect of permutations on circuit depth required to reach an exact solution in the VQE, we start with toy models and exact wave functions giving rise to exact entanglement maps (here only one loop for qubit reordering is needed). For that purpose we analyze several six-qubit model Ising Hamiltonians with artificially engineered entanglement. The generic Ising Hamiltonian is given by

$$\hat{H} = \sum_i \hat{Z}_i + \sum_{ij} C_{ij} \hat{X}_i \hat{X}_j + \sum_{ijk} C_{ijk} \hat{X}_i \hat{X}_j \hat{X}_k + \dots \quad (9)$$

We choose the hardware-efficient RyRz ansatz [44] (originally developed in the context of electronic structure calculations) with two layers as shown in Fig. 3, as it is

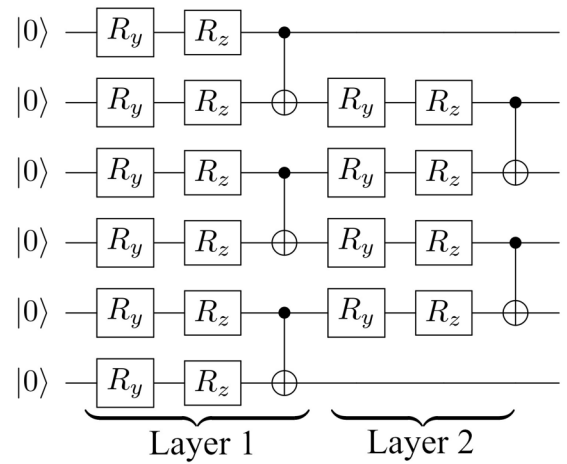


FIG. 3. The 2-depth RyRz ansatz that is used for VQE calculations using model Ising Hamiltonians.

naturally adapted for NISQ devices to ensure the nearest neighbor connectivity between qubits. The ansatz circuit with L layers is denoted as an “ L -depth” circuit. We consider five different model Hamiltonians, given in Table I. The entanglement maps for those Hamiltonians are shown in Fig. S1 in Supplemental Material [54]. From a comparison of different entanglement maps, we notice that the entanglement is more localized in the permuted case, supporting the choice of the cost function defined above. We find that for most of the Hamiltonians considered, qubit permutations significantly reduce the number of ansatz layers required to reach the exact energy of the system, as shown in Table I. For the case of the simplest Hamiltonians, \hat{H}_1 and \hat{H}_2 , the exact solution is already found at $L = 1$, where only the distinct qubit pairs are entangled. As expected, larger depth further reduces the advantage of permutations since large circuit depth allows one to solve the problem exactly for the unpermuted (default) Hamiltonian. Thus, we expect that qubit permutations could be an ideal method to variationally learn the ground-state wave functions on NISQ devices where circuit depths are limited.

TABLE I. Ansatz depth for six-qubit Ising models. For various Ising models, the ansatz depth that is required to reach the exact energy is shown for the unpermuted case and for the optimally permuted case.

Hamiltonian	Depth required to reach exact solution	
	Unpermuted	Permuted
$\hat{H}_1 = \sum_i \hat{Z}_i + \hat{X}_1 \hat{X}_6$	7	1
$\hat{H}_2 = \sum_i \hat{Z}_i + \hat{X}_1 \hat{X}_6 + \hat{X}_2 \hat{X}_5$	8	1
$\hat{H}_3 = \sum_i \hat{Z}_i + \hat{X}_1 \hat{X}_3 \hat{X}_5$	4	2
$\hat{H}_4 = \sum_i \hat{Z}_i + \hat{X}_1 \hat{X}_6 + \hat{X}_1 \hat{X}_3$	8	3
$\hat{H}_5 = \sum_i \hat{Z}_i + 2 * \hat{X}_1 \hat{X}_2 \hat{X}_5 \hat{X}_6$	5	5

For each Hamiltonian, we also compare the difference in permuted versus unpermuted cost-function values with the difference in permuted versus unpermuted depth required to converge to the true ground state; see Table S1 in Supplemental Material [54]. Notably, a larger difference in the cost function appears to be concomitant with a larger difference in the required ansatz depth, which may indicate a correlation between these parameters. However, for a large number of entangled qubits (e.g., \widehat{H}_5), this correlation fails. Nevertheless, the cost-function difference can be used as an initial estimate of the possible permutations.

B. Molecular systems with exact wave functions

For the actual molecular systems, the entanglement maps are expected to be vastly more complicated. Moreover, the type of mapping of the fermionic-to-qubit Hamiltonian can also affect the entanglement map of the system. To demonstrate the proof-of-principle advantage of qubit permutations on a circuit depth required to reach a near-to-exact solution in the VQE of chemical systems, we consider five molecular systems, LiH, H_2 , $(H_2)_2$ T-shaped van der Waals complex, H_4^{\ddagger} third-order saddle point, and H_3^+ cyclic cation, which are investigated in an ideal environment (exact wave function, exact entanglement map, no noise consideration). We compare the energy error $\Delta E = E_{\text{VQE}} - E_{\text{exact}}$ as a function of ansatz depth L , where E_{VQE} is the converged variational energy when the unpermuted or permuted Hamiltonian is used in the best-identified (JW) encoding (see below), and E_{exact} is the exact energy in the full or reduced active space (corrected by the energy shift) obtained by direct Hamiltonian matrix diagonalization in the basis of Slater determinants of spin-orbitals. Hence, the only sources of energy error in our studies are local minima and insufficiently expressive ansatz. For all systems a full active space of the STO-3G basis set [55] is used, except for LiH and H_2 , for which a reduced active space of ten spin-orbitals and a full active space of the 6-31G basis set [56] are selected, respectively. Slightly modified versions of the hardware-efficient RyRz and Ry ansatz [5,43,44] are used for all the molecules (see Figs. S2 and S3 in Supplemental Material [54]), except for H_3^+ , for which a particle-preserving hardware-efficient ansatz is used; see below. For all these systems, we present graphs of converged VQE energy with and without permutations in the main text.

1. Initial qubit mapping

There are multiple choices of spin-orbital to qubit mappings [1,2]. For the electronic structure problem, the JW [48], BK [49], and parity [50,51] methods (three popular second-quantized encoding basis set methods), differ primarily in the pattern to store information on the occupation number, the parity, and the number of qubit operations for simulating fermionic creation or annihilation operators.

While it was shown that the BK mapping can reduce the number of operations to $O(\log_2 N)$ for simulation of each fermionic operator [57], entanglement maps in this encoding contain a larger number of entangled qubits, which may reduce the advantage of permutations [Fig. 4(a)]. The same behavior was observed for the parity mapping [Fig. 4(b)]. The reason for this is likely because both the BK method and the parity method store information on the occupation number nonlocally, thus delocalizing the entanglement. In contrast, the occupation-number-locality-preserving JW mapping produces a sparse entanglement map [Fig. 4(c)]. We expect that the advantage of permutations will be more pronounced in this mapping, which is further used as the default throughout this paper. For example, LiH in the JW mapping includes six entangled qubits, unlike the BK and parity mappings, in which eight qubits are entangled; see Fig. 4. The comparison of JW entanglement maps for other molecular systems is shown in Fig.

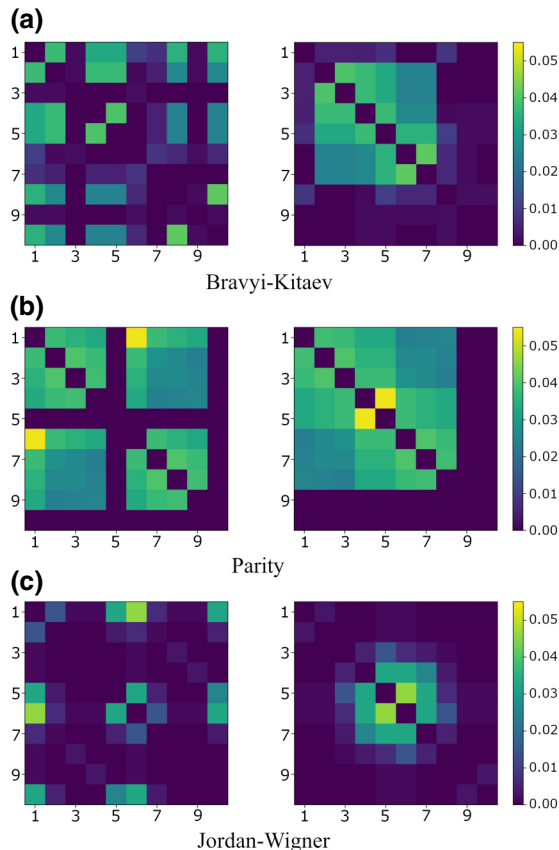


FIG. 4. Entanglement maps built from the exact wave function for LiH in the STO-3G basis in a reduced active space of ten spin-orbitals for the three commonly used mappings (a)–(c). On the left, the default Hamiltonian is used, and on the right, the permuted Hamiltonian is used. The numbers on the axes indicate qubit indices. Different values of mutual information I_{ij} are shown with different colors, with the color map shown on the right.

S4 in Supplemental Material [54]. We also compare exact entanglement maps for qubit Hamiltonians generated with IBM's Qiskit [58] and Google's Openfermion [59] frameworks for the LiH molecule in Fig. S5 in Supplemental Material [54]. Entanglement maps are identical in the JW encoding after qubit permutations, while they are different for BK and parity mappings. This result is expected taking into account that Qiskit orders (by default) spin-orbitals on the basis of spin (with "up" first and "down" next), whereas Openfermion performs even-odd ordering.

2. LiH molecule

For the analysis of LiH in the STO-3G basis, the lowest-energy spatial orbital is assumed to be doubly occupied and its contribution is integrated out to an effective field felt by the active space [60] of the remaining five orbitals; Fig. 1. The latter span ten spin-orbitals, and therefore the system can be described with a ten-qubit circuit.

For RyRz and Ry ansatz, the effect of qubit permutations on circuit depth is shown in Figs. 5(a) and 5(e), respectively. For both ansatz at depth $L = 2$, the VQE converges to a correlated wave function whose energy is comparable to the Hartree-Fock (HF) energy. For depth $L = 4$, approximately 70% of pseudocorrelation energy ΔE_{PC} [Fig. 5(a)] is captured for the permuted Hamiltonian, where ΔE_{PC} is defined as the difference between the HF energy E_{HF} and the exact ground-state energy E_{exact} for the reduced active space ($\Delta E_{PC} = 12.2$ kcal/mol). In contrast, for the unpermuted (default) qubit Hamiltonian, the VQE energy is still comparable to the HF energy even for the depth $L = 8$, indicating that it generates insufficient entanglement of the qubits. Similar results are observed with the Ry ansatz; see Fig. 5(e).

3. H_2 , $(H_2)_2$, and $H_4^\#$ molecular species

The improvement in energy convergence due to qubit permutations for eight-qubit circuit depth are well

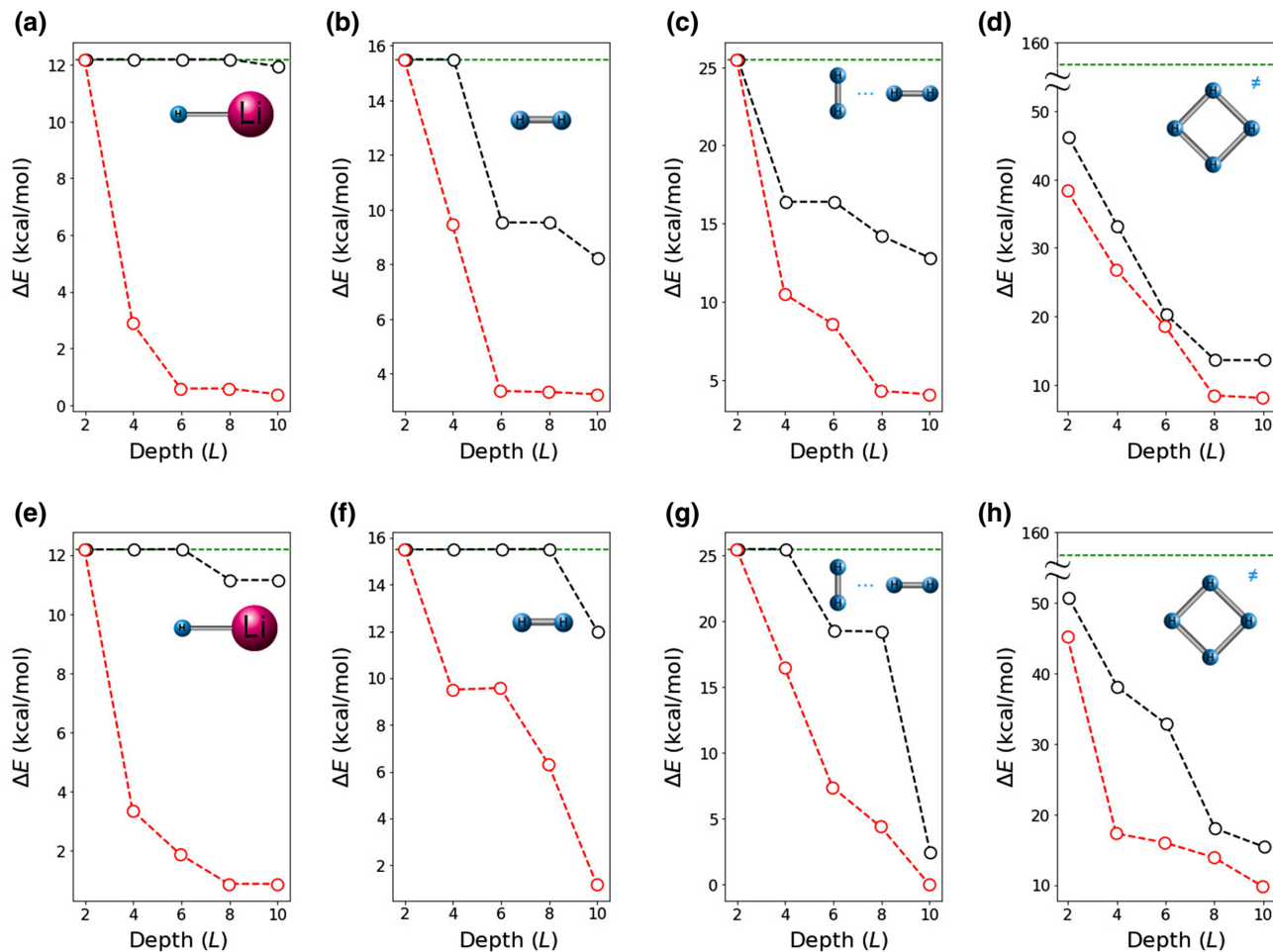


FIG. 5. Proof-of-principle energy error $\Delta E = E_{VQE} - E_{exact}$ versus ansatz depth L for various molecular species based on default (black) and permuted qubit Hamiltonians (red) built from the exact wave functions. (a)–(d) and (e)–(h) show the cases for the hardware-efficient RyRz and Ry ansatz, respectively. Dotted green lines correspond to pseudo-correlation energy $\Delta E_{PC} = E_{HF} - E_{exact}$.

observed for H_2 (6-31G basis), the non-covalently-bound (van der Waals) complex $(H_2)_2$ (STO-3G basis), and the H_4^\ddagger third-order saddle point (three imaginary frequencies in the Hessian matrix, STO-3G basis); see Figs. 5(b)–5(d) and Figs. 5(f)–5(h) for the RyRz and Ry ansatz, respectively. The level of correlations in these molecular systems is expected to increase as follows: $H_2 < (H_2)_2 \ll H_4^\ddagger$. Consequently, a smaller difference between permuted and unpermuted E_{VQE} can be observed in the corresponding panels. H_4^\ddagger is so strongly correlated that already at depth $L = 2$ the converged VQE corresponds to a highly correlated wave function even when an unpermuted Hamiltonian is used, $E_{HF} - E_{VQE}(\text{unpermuted}) \sim 100$ kcal/mol; see Figs. 5(d) and 5(h).

4. H_3^+ cation

Since the VQE performs an unconstrained energy optimization in the Fock space of the original electronic problem, calculation of the exact energy for charged molecules (ions) could be a challenging task [61]. Indeed, as charge-neutral molecules are usually energetically more preferable over their ions, the VQE will invariably collapse to a lower energy of the corresponding neutral form. Among other options [61], one could design a quantum circuit that conserves the number of electrons, and force the search to be in the expected subspace [62].

The JW mapping, which relates fixed-particle-number subspaces to the fixed-excitation subspaces in the qubit Hilbert space, seems to be the simplest method to approach such a circuit. A complication, however, is the large number of two-qubit gates that act on state $|00\rangle$ or $|11\rangle$ in the HF reference state, in which case electron-preserving gates must act as the identity. Thus, we remove these ineffective gates in our hardware-efficient ansatz when reporting the depth of our circuit to get a more appropriate comparison between unpermuted and permuted Hamiltonians. We display this result for our H_3^+ cation in Figs. 6(a) and 6(b) along with a decomposition of the electron-preserving gates, which we take from Ref. [62]. The HF layer was initialized depending on the permutation used. Thus, if qubit 1 corresponds to an occupied orbital and after the permutation it swaps with unoccupied qubit 5, we initialize qubit 5 in a $|1\rangle$ state and qubit 1 in a $|0\rangle$ state. Using this approach, we again find that permutations help us to reduce the circuit depth required to reach lower ΔE , although the effect is less profound compared with that for neutral molecules; see Fig. 7. At a higher depth $L = 8$, the results equalize.

C. Molecular systems with approximate wave functions (PermVQE)

So far we have demonstrated the proof-of-concept beneficial effect of qubit permutations on circuit depth for toy Ising models and various molecular systems based

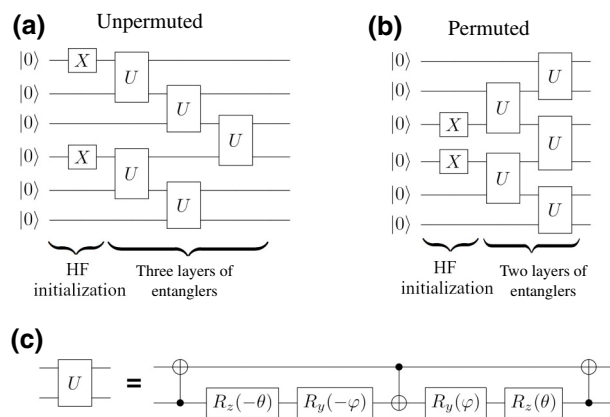


FIG. 6. Example of particle-preserving hardware-efficient ansatz that are used in H_3^+ simulation. (a) The 3-depth ansatz used with the unpermuted Hamiltonian. (b) The 2-depth ansatz used with the permuted Hamiltonian. (c) An electron-preserving entangler.

on ideally permuted Hamiltonians where exact entanglement maps were built from the exact wave functions under noise-free conditions. We now describe PermVQE presented in Fig. 2 for the example of a LiH molecule based on an approximate wave function under noise-free and noisy conditions without resorting to an exact solution. Each iteration uses the current best approximation of the ground state. Because von Neumann entropy S is a function of density-matrix values, the approximate wave function for molecular systems must be at a minimum correlated (e.g.,

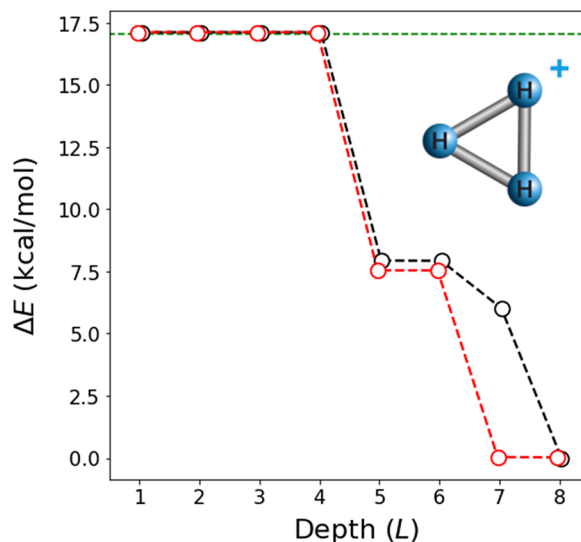


FIG. 7. Proof-of-principle energy error $\Delta E = E_{VQE} - E_{\text{exact}}$ versus ansatz depth L for the six-qubit H_3^+ cation based on default (black) and permuted qubit Hamiltonians (red) built from the exact wave function. Particle-preserving ansatz are used as shown in Fig. 6. The dotted green line corresponds to pseudo-correlation energy defined as $\Delta E_{PC} = E_{HF} - E_{\text{exact}}$.

improved over the uncorrelated HF wave function) before S is calculated.

For the LiH molecule in the reduced active space of the STO-3G basis, some correlations are observed at $L = 1$ and $L = 2$ for the Ry ansatz on the basis of the entanglement maps shown in Fig. S6 in Supplemental Material [54]. Despite the apparent correction from $L = 1$ to $L = 2$, the system is not correlated enough for further improvement. As expected, on increasing L under noise-free conditions ($L \geq 3$), one can clearly see the profound effect of qubit permutations on the converged VQE energy; see Fig. 8(a). For example, at $L = 7$, the converged VQE energy is at the level of chemical accuracy for the permuted Hamiltonian ($\Delta E = 1.2$ kcal/mol). In contrast, the converged VQE energy is at the level of the HF energy for the unpermuted case. We present the tabulated number of cost-function evaluations per PermVQE iteration for the converged VQE per L for unpermuted and permuted Hamiltonians as well as selected PermVQE iterations for $L = 1, 2, 5,$ and 10 (up to three for the system under study) including the corresponding JW entanglement maps in Table S2 and Fig. S7 in Supplemental Material [54].

We also note the following empirical observation. One could intuitively expect that the approximate wave function to build the best-permuted Hamiltonian (within given L) should ideally have a reasonable overlap with the exact wave function. Results obtained for the VQE based on the unpermuted versus the permuted Hamiltonian for $L = 2$ to $L = 7$ demonstrate a weak dependence on the accuracy of the initial wave function. The only relevant criterion seems to be that the variational state develops sufficient correlations, such that the correlation pattern is noticeable. Our experiments suggest that correlations do not have to be accurately computed for the purpose of finding favorable permutation of qubits. Indeed, the wave function is further refined in each iteration under fixed L ; for example, see Fig. S7 in Supplemental Material [54].

Finally we study the effects of qubit permutations on a noisy simulator. Optimized ansatz parameters under a noisy environment are used for the energy evaluation. No error mitigation is performed here and error rates are kept very low to obtain meaningful results, being 5×10^{-5} for one-qubit gates and 5×10^{-4} for two-qubit gates. The average energies after ten trials are shown in Fig. 8(b). For the unpermuted (default) Hamiltonian, the energy is not decreased below the HF value, and an increase of ansatz depth leads to a greater error in energy ΔE due to the additional noise that those layers introduce. Similar behavior is observed for several initial points of the permuted Hamiltonian ($L \leq 5$). However, a noticeable improvement is observed from $L \geq 6$ for the case of the permuted Hamiltonian. After $L = 7$, ΔE further increases as expected. We stress that the final permutation (for a given L) is obtained in several steps of the algorithm outlined in Fig. 2. In each step, the current best approximation of the ground

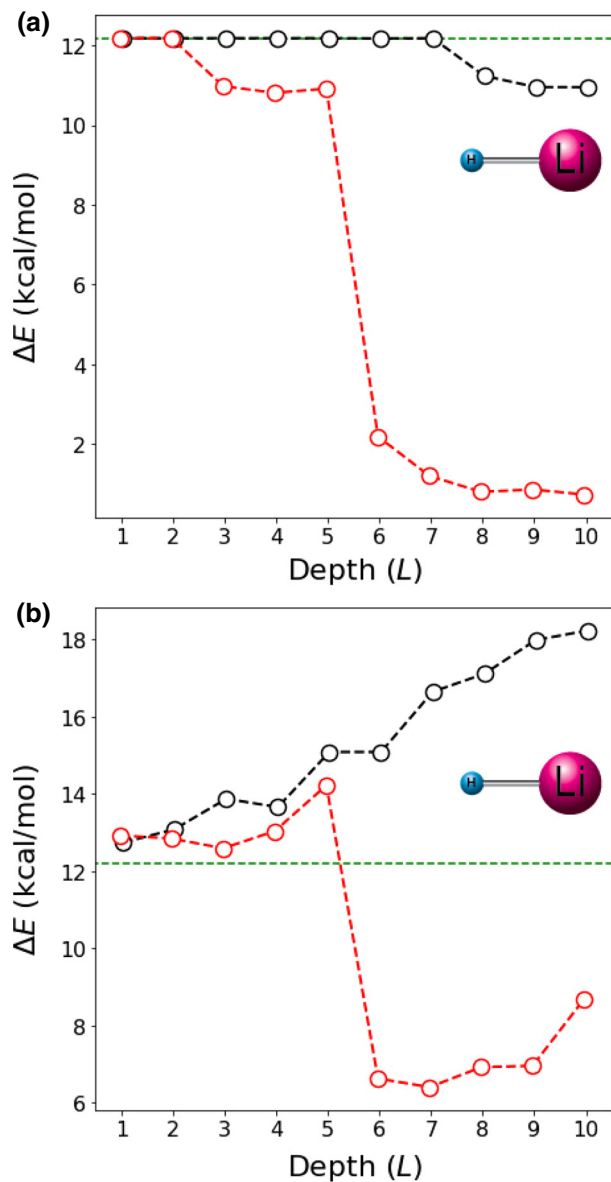


FIG. 8. Energy error $\Delta E = E_{\text{VQE}} - E_{\text{exact}}$ versus Ry ansatz depth L for the ten-qubit LiH in the STO-3G basis in a reduced active space based on default (black) and permuted qubit Hamiltonians (red) built from approximate wave functions (up to three PermVQE iterations are used to build the best entanglement maps). Results are based on (a) a state vector (noiseless) simulator and (b) the Qasm (noisy) simulator. Dotted green lines correspond to pseudo-correlation energy defined as $\Delta E_{\text{PC}} = E_{\text{HF}} - E_{\text{exact}}$.

state is used (as opposed to the exact ground state used for the proof-of-concept calculations presented in previous sections).

D. PermVQE for 2D grid architecture

Up to this point, we have demonstrated the beneficial effect of qubit permutations for small molecular systems

implemented on a linear qubit connectivity quantum chip. In this section we illustrate our approach for the N_2 molecule described by the STO-3G basis (16 qubits) on a 2D grid architecture chip. We assume that qubits are arranged on a 4×4 square lattice and every qubit is connected to its nearest neighbor only. This is a realistic setup and it represents a quantum chip that can be currently manufactured. The cost function in Eq. (4) is defined with the following measure of the distance between qubits i and j :

$$d_{ij} = |i_1 - j_1| + |i_2 - j_2|, \quad (10)$$

where qubits are labeled by their coordinates on a 2D square lattice, $i = (i_1, i_2)$ and $j = (j_1, j_2)$.

We construct a 2D generalization of the alternating layered ansatz shown in Fig. 3. Here, a given layer of the 2D ansatz takes one of two forms (A or B) shown in Fig. 9. The full ansatz takes the form $A-B-A-B-\dots-A-B$ to reach a given depth L . Layer A consists of four four-qubit gates acting on 2×2 blocks of qubits. Layer B is constructed such that the gates act on the boundary of 2×2 blocks defined in layer A . This is achieved with the combination of one four-qubit gate and four two-qubit gates, as shown in Fig. 9. Multiqubit gates are decomposed into controlled NOT (CNOT) gates and single-qubit rotations using known techniques [63–65]. The optimization is performed sequentially, gate by gate, until convergence in the energy is reached. A given gate in the ansatz is replaced by the one that minimizes the total energy, while all the other gates remain fixed. This variant of the optimization method can be implemented in a simulator in a more efficient way than the general techniques described in Appendix B. We use the methods implemented in Ref. [65] for fast simulation of large noiseless circuits. Qubit permutations are found with use of the heuristic method described in Sec. II E 2.

Our results for the N_2 molecule are summarized in Fig. 10. The beneficial effect of qubit permutations is already substantial at $L = 4$. One reaches a factor of 10 in reduction of the energy error at depths $L \geq 12$.

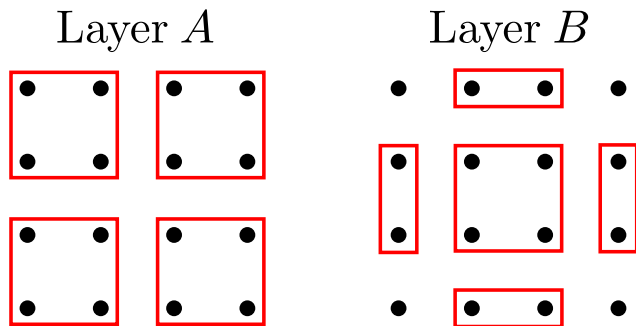


FIG. 9. Two-dimensional hardware-efficient ansatz used in this work to simulate the N_2 molecule in the STO-3G basis (16 qubits).

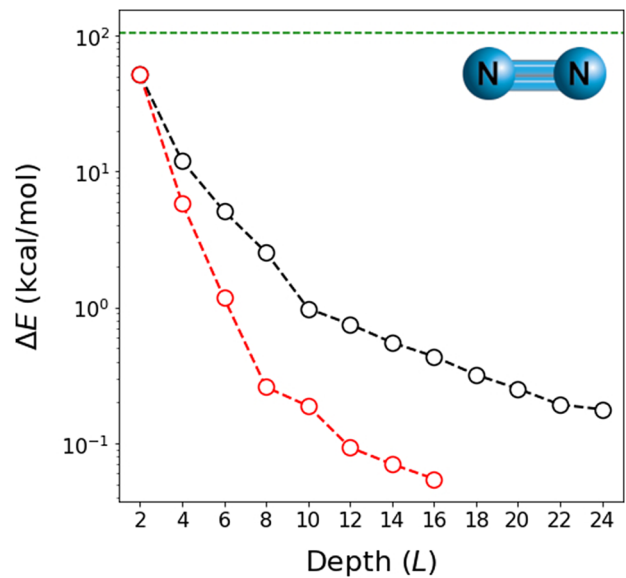


FIG. 10. Energy error $\Delta E = E_{VQE} - E_{\text{exact}}$ versus 2D hardware-efficient ansatz depth L for the 16-qubit N_2 in the STO-3G basis based on default (black) and permuted qubit Hamiltonians (red) built from approximate wave functions. The results are based on a noiseless simulator.

E. Beneficial effect of qubit permutations to build fermionic-ADAPT ansatz

In contrast to hardware-efficient ansatz [5,43,44,66,67], the beneficial effect of qubit permutations is not expected for another commonly used trial wave function in electronic structure, namely, the chemically inspired unitary coupled cluster for single and double (UCCSD) excitation ansatz [44,68], which naturally establishes important couplings between qubits. Permutations of qubits can therefore be regarded as a strategy to improve simulations of molecules on NISQ devices based on hardware-efficient ansatz for any qubit connectivity at the limit.

However, one might anticipate a beneficial effect of qubit permutations to build adaptive ansatz based on the ADAPT VQE [30] implemented on architectures with limited qubit connectivity. Using fermionic operators as a pool, Ref. [30] demonstrated that the ADAPT VQE substantially outperforms the UCCSD excitation ansatz for electronic structure simulations in terms of both the number of variational parameters and accuracy. The ADAPT VQE can be used with a variety of operator pools [69]. For example, there are a total of 24 single and double excitation operators for the LiH molecule described by a ten-qubit circuit (two electrons to be distributed on ten spin-orbitals). At convergence with a gradient norm threshold of $\epsilon_m = 10^{-m}$ ($m = 4$), the ADAPT VQE algorithm gives $|\psi^{\text{ADAPT}}(\epsilon_4)\rangle = e^{\theta_{05}^{16} \hat{\tau}_{05}^{16}} e^{\theta_{05}^{38} \hat{\tau}_{05}^{38}} e^{\theta_{05}^{27} \hat{\tau}_{05}^{27}} e^{\theta_{05}^{46} \hat{\tau}_{05}^{46}} e^{\theta_{05}^{19} \hat{\tau}_{05}^{19}} e^{\theta_{05}^{49} \hat{\tau}_{05}^{49}} |\psi^{\text{HF}}\rangle$ and

$|\psi^{\text{ADAPT}(\epsilon_4)}\rangle = e^{\theta_{05}^{16,16}} e^{\theta_{05}^{19,19}} e^{\theta_{05}^{46,46}} e^{\theta_{05}^{49,49}} |\psi^{\text{HF}}\rangle$ ansatz when noiseless and noisy simulators are used, respectively; $\hat{\tau}_{ij}^{ab} = (a_a^\dagger a_b^\dagger a_i a_j - a_i^\dagger a_j^\dagger a_a a_b)$, the anti-Hermitian sum of double excitation and de-excitation operators. Each $\hat{\tau}_{ij}^{ab}$ can be represented by a sum of eight Pauli words (e.g., Table S3 in Supplemental Material) [54], each of which is further used to construct an exponentiation quantum circuit; see Fig. S8 in Supplemental Material [54] for a representative example. The whole circuit contains multiple CNOT gates connecting distant qubits. For linear qubit connectivity architectures, they need to be decomposed into a series of multiple nearest-neighbor CNOT gates [70]. To decrease the number of such gates, one can find a permutation that minimizes the number of distant qubits connected with two-qubit gates. Even though such qubit permutations are not correlation-informed anymore, the proposed approach can significantly decrease the number of nearest-neighbor CNOT gates required to build the adaptive ansatz.

Figure 11 compares the number of CNOT gates versus the number of iterations used in the ADAPT VQE to iteratively build the $|\psi^{\text{ADAPT}(\epsilon_4)}\rangle$ ansatz. We show the unpermuted (default) and permuted qubit JW Hamiltonians of the LiH molecule in a reduced active space of the STO-3G basis (ten qubits) by using noiseless [Fig. 11(a)] and

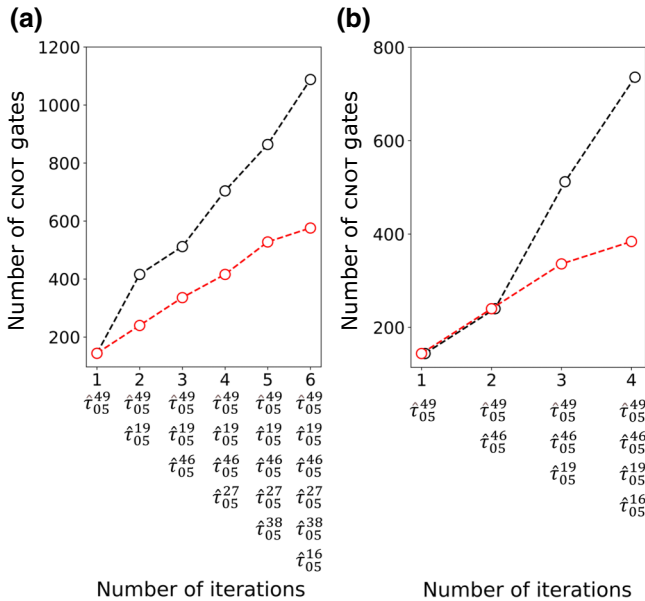


FIG. 11. Number of nearest-neighbor CNOT gates versus the number of iterations (with operator set included) used in the ADAPT VQE to iteratively build the $|\psi^{\text{ADAPT}(\epsilon_4)}\rangle$ ansatz. We consider the default (black) and permuted (red) qubit Hamiltonians of the LiH molecule in a reduced active space of the STO-3G basis (ten qubits) assuming a linear qubit connectivity. The set of operators is generated on the basis of (a) a state vector (noiseless) simulator and (b) the Qasm (noisy) simulator; $\hat{\tau}_{ij}^{ab} = (a_a^\dagger a_b^\dagger a_i a_j - a_i^\dagger a_j^\dagger a_a a_b)$.

noisy [Fig. 11(b)] simulators. The benefit of permutation is found almost in every iteration. Particularly, one needs 1088 nearest-neighbor CNOT gates to build the $|\psi^{\text{ADAPT}(\epsilon_4)}\rangle$ ansatz with six excitation operators for the unpermuted JW Hamiltonian. The number of CNOT gates drops to 576 when the permuted JW Hamiltonian is used instead (noiseless simulator). Similarly, for noisy calculations, the $|\psi^{\text{ADAPT}(\epsilon_4)}\rangle$ ansatz with four excitation operators includes 736 nearest-neighbor CNOT gates, while for the best permutation the ansatz can be built with only 384 of them. The reduction of the number of nearest-neighbor CNOT gates is approximately twofold therefore, which could be a noticeable advance to simulate molecules on NISQ devices with the ADAPT VQE.

IV. CONCLUSIONS

In this work we show that encoding strongly interacting spin-orbitals of molecular systems into proximal qubits on a linear or 2D grid architecture chain architecture quantum chip naturally reduces the circuit depth needed to prepare the ground state for the quantum chemistry electronic structure problem with hardware-efficient ansatz. If we assume direct mapping between qubits and canonical spin-orbitals (eigenfunctions of the Fock operator)—that is, each qubit represents the occupation number of a particular spin-orbital—the two-qubit mutual information describes the entanglement between two individual qubits, whereas JW encoding is the most suitable for permuting qubits. With this observation, we develop the PermVQE algorithm, which uses default (unpermuted) Hamiltonians generated with IBM’s Qiskit or Google’s Openfermion frameworks. PermVQE then iteratively converts inputs into permuted Hamiltonians on the basis of mutual information from the approximate wave function, and performs the VQE with increased precision at a given circuit depth L . We remark that a different approach based on mutual-information maximization to select entangling ansatz (rather than our mutual-information minimization to select permutations) was reported in Ref. [71] during the preparation of this paper.

We believe that the proposed method will facilitate the simulation of larger molecular systems or basis sets with NISQ devices by reducing the depth, and will contribute to the demonstration of chemical advantage. Furthermore, our correlation-informed permutation approach can be combined within the other variational quantum algorithms [72] beyond the VQE, for example, to reduce ansatz depth in variational quantum algorithms for simulating dynamics [73–76], solving linear systems [77,78], and compiling [29,79]. As an example, we demonstrate the beneficial effect of qubit permutations to build fermionic-ADAPT ansatz on a linear qubit connectivity architecture with a nearly twofold reduction of the number of CNOT gates.

Finally, PermVQE can be used for any hardware connectivity with hardware-efficient ansatz. Future study might include extending the permutation approach to take advantage of more highly connected qubit architectures, such as a hexagonal lattice, which would require modified and more complicated cost functions.

ACKNOWLEDGMENTS

The research presented in this paper was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory (LANL) under Project No. 20200056DR. LANL is operated by Triad National Security LLC for the National Nuclear Security Administration of the U.S. Department of Energy (Contract No. 89233218CNA000001). We thank the LANL Institutional Computing Program for access to high-performance computing resources. J.S. acknowledges support from the U.S. Department of Energy through a quantum computing program sponsored by the LANL Information Science & Technology Institute. This work was conducted in part at the Center for Integrated Nanotechnologies, a U.S. Department of Energy, Office of Basic Energy Sciences user facility. P.J.C. and L.C. were partially supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under the Accelerated Research in Quantum Computing program.

APPENDIX A: HARDWARE-EFFICIENT ANSATZ

The heuristic RyRz ansatz built on the basis of IBM's Qiskit software package [58] is used to prepare quantum states for Ising toy models, whereas the partially modified RyRz and Ry ansatz are used to prepare quantum states for neutral molecules. For RyRz, additional Ry and Rz rotational final layers are included to provide more flexibility to the parameterized circuit; see Fig. S2 in Supplemental Material [54]. For Ry, all Rz gates are removed and an additional final Ry rotational layer is included; see Fig. S3 in Supplemental Material [54]. This ansatz can be used since all ground states of the systems that we analyze possess time-reversal symmetry. Thus, we expect all state coefficients to be purely real, eliminating the need for Z rotations. The electron-preserving ansatz shown in Fig. 6 is used to model the H_3^+ cation.

APPENDIX B: COMPUTATIONAL METHODS

All quantum simulations (except for those reported in Sec. III D) are performed with the Qiskit 0.19.6 package [58]. For the VQE and the noisy VQE, Qiskit's state vector and Qasm simulators are used, respectively. For a classical optimization, the COBYLA protocol [80] is used with 200 000 maximum iterations for the molecular systems

investigated. The lowest eigenvalue is obtained from multiple trial runs (five to ten) for each depth of the ansatz. For model Ising Hamiltonians, the same classical optimizer is used with a maximum of 10 000 iterations. The PermVQE calculations with approximate reference wave functions are performed with the same classical optimizer with a maximum of 100 000 iterations and a maximum of three consecutive permutations of the corresponding Hamiltonians. The noise model is incorporated by our running the final step of the VQE calculation using Qasm simulator with 10 000 shots per Pauli-word evaluation and error rates of 5×10^{-5} for one-qubit gates and 5×10^{-4} for two-qubit gates. The geometries of the molecular systems are initially preoptimized (in the gas phase) at the restricted Hartree-Fock or full configurational interaction level for $(H_2)_2$ with use of the electronic structure package GAUSSIAN 16 (revision B.01) [81]. Table S4 in Supplemental Material [54] contains additional information (x - y - z coordinates, energy data). The molecular orbitals are visualized with ChemCraft [82]. The entanglement maps for each run are built with use of the exact or approximate wave function of a given system. The brute-force approach is applied to find the best permutation of a given entanglement map. The results in Sec. III D are obtained with a simulator implemented in Ref. [65] and with the modified optimization technique described in the main text. The ADAPT VQE algorithm using the Qasm simulator with 1024 shots per Pauli-word evaluation and the L_BFGS_B optimizer with 15 000 maximum iterations is used to define a set of the most important single and double excitation operators for the LiH molecule in a reduced active space of ten spin-orbitals. The threshold criterion for operator selection is set to 10^{-4} . The same error rates are used as specified above. The number of CNOT gates is calculated from the explicit form of the ansatz circuit following Ref. [70].

-
- [1] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, and Nicolas P. D. Sawaya *et al.*, Quantum chemistry in the age of quantum computing, *Chem. Rev.* **119**, 10856 (2019).
 - [2] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
 - [3] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213 (2014).
 - [4] Peter J. J. O'Malley, Ryan Babbush, Ian D. Kivlichan, Jonathan Romero, Jarrod R. McClean, Rami Barends, Julian Kelly, Pedram Roushan, Andrew Tranter, and Nan Ding *et al.*, Scalable Quantum Simulation of Molecular Energies, *Phys. Rev. X* **6**, 031007 (2016).

- [5] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature* **549**, 242 (2017).
- [6] Abhinav Kandala, Kristan Temme, Antonio D Córcoles, Antonio Mezzacapo, Jerry M. Chow, and Jay M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, *Nature* **567**, 491 (2019).
- [7] Alexander J. McCaskey, Zachary P. Parks, Jacek Jakowski, Shirley V. Moore, Titus D. Morris, Travis S. Humble, and Raphael C. Pooser, Quantum chemistry as a benchmark for near-term quantum computers, *Npj Quantum Inf.* **5**, 99 (2019).
- [8] Frank Arute *et al.*, Hartree-fock on a superconducting qubit quantum computer, *Science* **369**, 1084 (2020).
- [9] Yunseong Nam, Jwo-Sy Chen, Neal C. Pseni, Kenneth Wright, Conor Delaney, Dmitri Maslov, Kenneth R. Brown, Stewart Allen, Jason M. Amini, and Joel Apisdorf *et al.*, Ground-state energy estimation of the water molecule on a trapped-ion quantum computer, *Npj Quantum Inf.* **6**, 1 (2020).
- [10] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, and David A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
- [11] Joonho Lee, William J. Huggins, Martin Head-Gordon, and K. Birgitta Whaley, Generalized unitary coupled cluster wave functions for quantum computation, *J. Chem. Theory Comput.* **15**, 311 (2019).
- [12] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [13] Marco Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles, Cost-function-dependent barren plateaus in shallow quantum neural networks, [arXiv:2001.00550](https://arxiv.org/abs/2001.00550) (2020).
- [14] Kunal Sharma, Marco Cerezo, Lukasz Cincio, and Patrick J. Coles, Trainability of dissipative perceptron-based quantum neural networks, [arXiv:2005.12458](https://arxiv.org/abs/2005.12458) (2020).
- [15] Marco Cerezo and Patrick J. Coles, Impact of barren plateaus on the Hessian and higher order derivatives, [arXiv:2008.07454](https://arxiv.org/abs/2008.07454) (2020).
- [16] Andrew Arrasmith, M. Cerezo, Piotr Czarnik, Lukasz Cincio, and Patrick J. Coles, Effect of barren plateaus on gradient-free optimization, [arXiv preprint arXiv:2011.12245](https://arxiv.org/abs/2011.12245) (2020).
- [17] Zoe Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, [arXiv preprint arXiv:2101.02138](https://arxiv.org/abs/2101.02138) (2021).
- [18] Alexey Uvarov, Jacob D. Biamonte, and Dmitry Yudin, Variational quantum eigensolver for frustrated quantum systems, *Phys. Rev. B* **102**, 075104 (2020).
- [19] Samson Wang, Enrico Fontana, Marco Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles, Noise-induced barren plateaus in variational quantum algorithms, [arXiv:2007.14384](https://arxiv.org/abs/2007.14384) (2020).
- [20] Jonas M. Kübler, Andrew Arrasmith, Lukasz Cincio, and Patrick J. Coles, An adaptive optimizer for measurement-frugal variational algorithms, *Quantum* **4**, 263 (2020).
- [21] Andrew Arrasmith, Lukasz Cincio, Rolando D. Somma, and Patrick J. Coles, Operator sampling for shot-frugal optimization in variational algorithms, [arXiv:2004.06252](https://arxiv.org/abs/2004.06252) (2020).
- [22] Kevin Jeffery Sung, Jiahao Yao, Matthew Harrigan, Nicholas Rubin, Zhang Jiang, Lin Lin, Ryan Babbush, and Jarrod McClean, Using models to improve optimizers for variational quantum algorithms, *Quantum Sci. Technol.* **5**, 044008 (2020).
- [23] Tyler Volkoff and Patrick J. Coles, Large gradients via correlation in random parameterized quantum circuits, [arXiv:2005.12200](https://arxiv.org/abs/2005.12200) (2020).
- [24] Guillaume Verdon, Michael Broughton, Jarrod R. McClean, Kevin J. Sung, Ryan Babbush, Zhang Jiang, Hartmut Neven, and Masoud Mohseni, Learning to learn with quantum neural networks via classical neural networks, [arXiv:1907.05415](https://arxiv.org/abs/1907.05415) (2019).
- [25] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, *Quantum* **3**, 214 (2019).
- [26] Ying Li and Simon C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, *Phys. Rev. X* **7**, 021050 (2017).
- [27] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, *Phys. Rev. Lett.* **119**, 180509 (2017).
- [28] Piotr Czarnik, Andrew Arrasmith, Patrick J. Coles, and Lukasz Cincio, Error mitigation with Clifford quantum-circuit data, [arXiv:2005.10189](https://arxiv.org/abs/2005.10189) (2020).
- [29] Kunal Sharma, Sumeet Khatri, Marco Cerezo, and Patrick J. Coles, Noise resilience of variational quantum compiling, *New J. Phys.* **22**, 043006 (2020).
- [30] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, *Nat. Commun.* **10**, 3007 (2019).
- [31] Ho Lun Tang, Edwin Barnes, Harper R. Grimsley, Nicholas J. Mayhall, and Sophia E. Economou, qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansatz on a quantum processor, [arXiv:1911.10205](https://arxiv.org/abs/1911.10205) (2019).
- [32] Ilya G. Ryabinkin, Tzu-Ching Yen, Scott N. Genin, and Artur F. Izmaylov, Qubit coupled cluster method: A systematic approach to quantum chemistry on a quantum computer, *J. Chem. Theory Comput.* **14**, 6317 (2018).
- [33] Ilya G. Ryabinkin, Robert A. Lang, Scott N. Genin, and Artur F. Izmaylov, Iterative qubit coupled cluster approach with efficient screening of generators, *J. Chem. Theory Comput.* **16**, 1055 (2020).
- [34] Robert A. Lang, Ilya G. Ryabinkin, and Artur F. Izmaylov, Unitary transformation of the electronic hamiltonian with an exact quadratic truncation of the baker-campbell-hausdorff expansion, *J. Chem. Theory Comput.* **17**, 66 (2021), PMID: 33295175.
- [35] Ilya G. Ryabinkin, Artur F. Izmaylov, and Scott N. Genin, A posteriori corrections to the iterative qubit coupled cluster

- method to minimize the use of quantum resources in large-scale calculations, [arXiv:2009.13622](https://arxiv.org/abs/2009.13622) [quant-ph] (2020).
- [36] Yordan S. Yordanov, V. Armaos, Crispin H. W. Barnes, and David R. M. Arvidsson-Shukur, Iterative qubit-excitation based variational quantum eigensolver, [arXiv:2011.10540](https://arxiv.org/abs/2011.10540) [quant-ph] (2020).
- [37] Sukin Sim, Jonathan Romero, Jerome F. Gonthier, and Alexander A. Kunitsa, Adaptive pruning-based optimization of parameterized quantum circuits, [arXiv:2010.00629](https://arxiv.org/abs/2010.00629) [quant-ph] (2020).
- [38] Daniel Claudino, Jeremiah Wright, Alexander J. McCaskey, and Travis S. Humble, Benchmarking adaptive variational quantum eigensolvers, *Front. Chem.* **8**, 606863 (2020).
- [39] Atsushi Matsuo, Yudai Suzuki, and Shigeru Yamashita, Problem-specific parameterized quantum circuits of the vqe algorithm for optimization problems, [arXiv:2006.05643](https://arxiv.org/abs/2006.05643) [quant-ph] (2020).
- [40] Rongxin Xia and Sabre Kais, Qubit coupled cluster singles and doubles variational quantum eigensolver ansatz for electronic structure calculations, *Quantum Sci. Technol.* **6**, 015001 (2020).
- [41] Swamit S. Tannu and Moinuddin K. Qureshi, in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS '19 (Association for Computing Machinery, New York, NY, USA, 2019), p. 987.
- [42] Adam Holmes, Sonika Johri, Gian Giacomo Guerreschi, James S. Clarke, and A. Y. Matsuura, Impact of qubit connectivity on quantum algorithm performance, *Quantum Sci. Technol.* **5**, 025009 (2020).
- [43] Pauline J. Ollitrault, Alberto Baiardi, Markus Reiher, and Ivano Tavernelli, Hardware efficient quantum algorithms for vibrational structure calculations, *Chem. Sci.* **11**, 6842 (2020).
- [44] Panagiotis Kl. Barkoutsos, Jerome F. Gonthier, Igor Sokolov, Nikolaj Moll, Gian Salis, Andreas Fuhrer, Marc Ganzhorn, Daniel J. Egger, Matthias Troyer, Antonio Mezzacapo, Stefan Filipp, and Ivano Tavernelli, Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wave-function expansions, *Phys. Rev. A* **98**, 022322 (2018).
- [45] Jörg Rissler, Reinhard M. Noack, and Steven R. White, Measuring orbital interaction using quantum information theory, *Chem. Phys. Phys. Lett.* **323**, 519 (2006).
- [46] Zhen Huang and Sabre Kais, Entanglement as measure of electron-electron correlation in quantum chemistry calculations, *Chem. Phys. Lett.* **413**, 1 (2005).
- [47] Ian D. Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, *Phys. Rev. Lett.* **120**, 110501 (2018).
- [48] P. Jordan and E. Wigner, Über das Paulische Äquivalenzverbot, *Zeitschrift für Physik* **47**, 631 (1928).
- [49] Sergey B. Bravyi and Alexei Yu. Kitaev, Fermionic quantum computation, *Ann. Phys.* **298**, 210 (2002).
- [50] Jacob T. Seeley, Martin J. Richard, and Peter J. Love, The bravyi-kitaev transformation for quantum computation of electronic structure, *J. Chem. Phys.* **137**, 224109 (2012).
- [51] Sergey Bravyi, Jay M. Gambetta, Antonio Mezzacapo, and Kristan Temme, Tapering off qubits to simulate fermionic hamiltonians, [arXiv:1701.08213](https://arxiv.org/abs/1701.08213) (2017).
- [52] Miroslav Fiedler, A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory, *Czech. Math. J.* **25**, 619 (1975).
- [53] G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher, Quantum-information analysis of electronic states of different molecular structures, *Phys. Rev. A* **83**, 012508 (2011).
- [54] See the Supplemental Material at <http://link.aps.org/supplemental/10.1103/PRXQuantum.2.020337> for additional results including entanglement maps, cost-function values, number of cost-function evaluations as well as Cartesian coordinates and corresponding energies of molecular systems studied in the main text.
- [55] W. J. Hehre, R. F. Stewart, and J. A. Pople, Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of slater-type atomic orbitals, *J. Chem. Phys.* **51**, 2657 (1969).
- [56] R. Ditchfield, W. J. Hehre, and J. A. Pople, Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules, *J. Chem. Phys.* **54**, 724 (1971).
- [57] Andrew Tranter, Peter J. Love, Florian Mintert, and Peter V. Coveney, A comparison of the bravyi-kitaev and jordan-wigner transformations for the quantum simulation of quantum chemistry, *J. Chem. Theory Comput.* **14**, 5617 (2018).
- [58] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen *et al.*, Qiskit: An open-source framework for quantum computing, (2019).
- [59] Jarrod R. McClean *et al.*, OpenFermion: The electronic structure package for quantum computers, *Quantum Sci. Technol.* **5**, 034014 (2020).
- [60] Tyler Takeshita, Nicholas C. Rubin, Zhang Jiang, Eunseok Lee, Ryan Babbush, and Jarrod R. McClean, Increasing the Representation Accuracy of Quantum Simulations of Chemistry Without Extra Quantum Resources, *Phys. Rev. X* **10**, 011004 (2020).
- [61] Ilya G. Ryabinkin, Scott N. Genin, and Artur F. Izmaylov, Constrained variational quantum eigensolver: Quantum computer search engine in the fock space, [arXiv:1806.00461](https://arxiv.org/abs/1806.00461) (2018).
- [62] Bryan T. Gard, Linghua Zhu, and George S. Barron *et al.*, Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm, *Npj Quantum Inf.* **6**, 1 (2020).
- [63] Marc G. Davis, Ethan Smith, Ana Tudor, Koushik Sen, Irfan Siddiqi, and Costin Iancu, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2020), p. 223.
- [64] Ed Younis, Koushik Sen, Katherine Yelick, and Costin Iancu, QFAST: Quantum synthesis using a hierarchical continuous circuit space, [arXiv:2003.04462](https://arxiv.org/abs/2003.04462) (2020).
- [65] Ed Younis and Lukasz Cincio, <https://github.com/edyounis/qfactor>.
- [66] Arthur G. Rattew, Shaohan Hu, Marco Pistoia, Richard Chen, and Steve Wood, A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver, [arXiv:1910.09694](https://arxiv.org/abs/1910.09694) [quant-ph] (2020).

- [67] Stefano Barison, Davide Emilio Galli, and Mario Motta, Quantum simulations of molecular systems with intrinsic atomic orbitals, [arXiv:2011.08137](https://arxiv.org/abs/2011.08137) [quant-ph] (2020).
- [68] Jonathan Romero, Ryan Babbush, Jarrod R. McClean, Cornelius Hempel, Peter J. Love, and Alán Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, *Quantum Sci. Technol.* **4**, 014008 (2018).
- [69] Ho Lun Tang, V. O. Shkolnikov, George S. Barron, Harper R. Grimsley, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou, qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansatzes on a quantum processor, [arXiv:1911.10205](https://arxiv.org/abs/1911.10205) [quant-ph] (2020).
- [70] Md. Mazder Rahman and Gerhard W. Dueck, Synthesis of linear nearest neighbor quantum circuits, [arXiv:1508.05430](https://arxiv.org/abs/1508.05430) [cs.ET] (2015).
- [71] Zi-Jian Zhang, Thi Ha Kyaw, Jakob S. Kottmann, Matthias Degroote, and Alán Aspuru-Guzik, Mutual information-assisted adaptive variational quantum eigensolver ansatz construction, [arXiv:2008.07553](https://arxiv.org/abs/2008.07553) (2020).
- [72] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles, Variational quantum algorithms, [arXiv preprint arXiv:2012.09265](https://arxiv.org/abs/2012.09265) (2020).
- [73] Cristina Cirstoiu, Zoe Holmes, Joseph Iosue, Lukasz Cincio, Patrick J. Coles, and Andrew Sornborger, Variational fast forwarding for quantum simulation beyond the coherence time, [arXiv:1910.04292](https://arxiv.org/abs/1910.04292) (2019).
- [74] Benjamin Commeau, Marco Cerezo, Zoe Holmes, Lukasz Cincio, Patrick J. Coles, and Andrew Sornborger, Variational Hamiltonian diagonalization for dynamical quantum simulation, [arXiv:2009.02559](https://arxiv.org/abs/2009.02559) (2020).
- [75] Xiao Yuan, Suguru Endo, Qi Zhao, Ying Li, and Simon C. Benjamin, Theory of variational quantum simulation, *Quantum* **3**, 191 (2019).
- [76] Suguru Endo, Jinzhao Sun, Ying Li, Simon C. Benjamin, and Xiao Yuan, Variational Quantum Simulation of General Processes, *Phys. Rev. Lett.* **125**, 010501 (2020).
- [77] Carlos Bravo-Prieto, Ryan LaRose, Marco Cerezo, Yigit Subasi, Lukasz Cincio, and Patrick J. Coles, Variational quantum linear solver: A hybrid algorithm for linear systems, [arXiv:1909.05820](https://arxiv.org/abs/1909.05820) (2019).
- [78] Xiaosi Xu, Jinzhao Sun, Suguru Endo, Ying Li, Simon C. Benjamin, and Xiao Yuan, Variational algorithms for linear algebra, [arXiv:1909.03898](https://arxiv.org/abs/1909.03898) (2019).
- [79] Sumeet Khatri, Ryan LaRose, Alexander Poremba, Lukasz Cincio, Andrew T. Sornborger, and Patrick J. Coles, Quantum-assisted quantum compiling, *Quantum* **3**, 140 (2019).
- [80] M. J. D. Powell, Direct search algorithms for optimization calculations, *Acta Numer.* **7**, 287 (1998).
- [81] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, G. A. Petersson V. Barone, and H. Nakatsuji *et al.*, *Gaussian 16, Revision b.01*, (Gaussian, Inc., Wallingford CT, 2016).
- [82] <https://www.chemcraftprog.com/>.