Schrödinger-Heisenberg Variational Quantum Algorithms

Zhong-Xia Shang,[1,2,3] Ming-Cheng Chen,[1,2,3] Xiao Yuan,[4,5] Chao-Yang Lu,[1,2,3] and Jian-Wei Pan[1,2,3]
[1]Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China
[2]Shanghai Branch, CAS Centre for Excellence and Synergetic Innovation Centre in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai 201315, China
[3]Shanghai Research Center for Quantum Sciences, Shanghai 201315, China
[4]Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China
[5]School of Computer Science, Peking University, Beijing 100871, China

(Received 4 March 2022; accepted 14 July 2023; published 11 August 2023)

Recent breakthroughs have opened the possibility of intermediate-scale quantum computing with tens to hundreds of qubits, and shown the potential for solving classical challenging problems, such as in chemistry and condensed matter physics. However, the high accuracy needed to surpass classical computers poses a critical demand on the circuit depth, which is severely limited by the non-negligible gate infidelity, currently around 0.1%–1%. The limited circuit depth places restrictions on the performance of variational quantum algorithms (VQA) and prevents VQAs from exploring desired nontrivial quantum states. To resolve this problem, we propose a paradigm of Schrödinger-Heisenberg variational quantum algorithms (SHVQA). Using SHVQA, the expectation values of operators on states that require very deep circuits to prepare can now be efficiently measured by rather shallow circuits. The idea is to incorporate a virtual Heisenberg circuit, which acts effectively on the measurement observables, into a real shallow Schrödinger circuit, which is implemented realistically on the quantum hardware. We choose a Clifford virtual circuit, whose effect on the Hamiltonian can be seen as efficient classical processing. Yet, it greatly enlarges the state's expressivity, realizing much larger unitary t designs. Our method enables accurate quantum simulation and computation that otherwise are only achievable with much deeper circuits or more accurate operations conventionally. This has been verified in our numerical experiments for a better approximation of random states, higher-fidelity solutions to the XXZ model, and the electronic structure Hamiltonians of small molecules. Thus, together with effective quantum error mitigation, our work paves the way for realizing accurate quantum computing algorithms with near-term quantum devices.

DOI: 10.1103/PhysRevLett.131.060406

Almost four decades after Richard Feynman put forward the idea of quantum computing [1], the quantum advantage has been experimentally tested recently in the solid state systems [2–4] and photonic systems [5,6]. However, those quantum computational advantage works focused on well-defined quantum sampling problems which were not designed to be practically useful. Therefore, the next important near-term milestone is to find algorithms for noisy intermediate-scale quantum (NISQ) [7] devices to solve nontrivial practical problems that are intractable for classical computation.

One of the most promising NISQ applications is using variational quantum algorithms (VQA) [8,9] such as the variational quantum eigensolver (VQE) [10] and the variational quantum simulation (VQS) [11] where a quantum circuit is optimized classically to approximate the eigenstate state energy and to simulate the dynamics of a Hamiltonian respectively for tasks that are widely considered in combinatorial optimization problems [12], condensed matter physics [13], and quantum chemistry [14,15]. A practical advantage of hybrid algorithms is their certain degree of resilience to noise in the optimization and quantum hardware [8,16,17].

Considering the limitations of NISQ devices, VQAs generally use a shallow local unitary circuit (LUC) [Fig. 1(a)] to approximate the target quantum states. States prepared by shallow LUCs however, could be trivial, obeying the entanglement area law [18] which can be well captured by classical tensor networks [19]. Indeed, the Lieb-Robinson bound [20] indicates that the entanglement light cone restricts the propagation of correlations, and, therefore, shallow LUC cannot generate long-range entanglement. However, the ground states of some Hamiltonians of interest could be highly nontrivial and require a relatively deep LUC with a depth that has linear or even higher scaling with the qubit number [20,21] such as interacting spins at critical points [22,23], topological quantum orders [24,25], and interacting fermions in complex molecules [15]. This is a big challenge for NISQ devices. Indeed, without an effective quantum error correction, the final

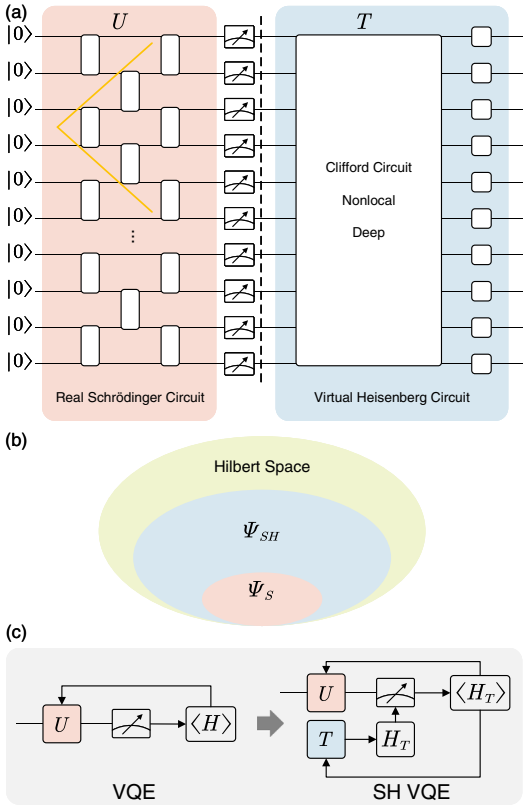

FIG. 1. SHVQE. (a) The SHVQE circuit. The circuit is composed of the Schrödinger circuit U and the Heisenberg circuit T, where U is the local unitary circuit running on real quantum computers and T is the virtual circuit acted on the Hamiltonian consisting of two parts, the Clifford part, and the single qubit layer. The architecture we use for U throughout this work is layers of parallel 2-qubit gates, which has a well-defined light cone that constrains the propagation of correlations and entanglements. (b) Improvements of SHVQE. By adding the virtual circuit, $TU|0^{\otimes n}\rangle$ is able to explore more of the Hilbert space compared with $U|0^{\otimes n}\rangle$ in conventional VQE, and the trainable Hilbert space is much larger than the conventional VQE. (c) Algorithm structure comparison between VQE and SHVQE. The transformed Hamiltonian H_T replaces H in SHVQE. We update parameters in both U and T to minimize the expectation value of H_T.

fidelity of the quantum circuits drops exponentially with the number of gates. For example, a state-of-the-art random quantum circuit with 60 qubits and 24 layers [3] ended up with a cross-entropy benchmarking fidelity as low as 0.037%. We thus need to significantly improve the NISQ hardware to implement those VQAs to the desired accuracy.

This situation can be summarized as a trade-off between the fidelity of the LUC and its expressivity [9] (i.e., the ability for the quantum circuits to "express" a sufficiently large volume of quantum states to include those nontrivial ones). To circumvent this problem, we propose a new framework of VQAs, enhanced by virtual Heisenberg circuits, which can noiselessly increase the effective circuit depth and thus simultaneously improve its expressivity and fidelity. We want to mention that there is a related work by Zhang *et al.* where their classical neural networks serve for a purpose similar to our virtual Heisenberg circuits [26]. And there is an orbital optimized unitary coupled cluster method [27] that shares a similar idea as ours where they turn single-excitation circuits into classical processing on chemical Hamiltonians. We call our scheme Schrödinger-Heisenberg (SH) VQA, which illustrates that the main idea is that, in addition to the physical unitary circuit, U, acting on the quantum states in the Schrödinger picture, we bring in a virtual circuit, T, acting on the target Hamiltonian H in the Heisenberg picture [see Fig. 1(a)]. In the following, we consider SHVQE as an example, but we note that the algorithm works for general VQAs. In this case, the energy expectation value $E(T,U) = \langle 0^{\otimes n}|U^\dagger T^\dagger H T U|0^{\otimes n}\rangle$ of the system becomes

$$E(T,U) = \langle 0^{\otimes n}|U^\dagger H_T U|0^{\otimes n}\rangle \tag{1}$$

where the classically calculated transformed Hamiltonian $H_T = T^\dagger H T$ has the same energy spectrum as H. By properly choosing a relatively deep but noiseless T, the state $TU|0^{\otimes n}\rangle$ could explore the Hilbert space far outside the range of $U|0^{\otimes n}\rangle$ [see Fig. 1(b)] and hence can obtain lower and more accurate ground-state energy than conventional VQE for nontrivial problems. We show a workflow of SHVQE together with a comparison to conventional VQE in Fig. 1(c). Compared with VQE, both the real Schrödinger circuit U and the virtual Heisenberg circuit T in SHVQE are parametrized and updated when minimizing the expectation value $E(T,U)$. The key feature of SHVQE is that only U as a shallow LUC is physically implemented, whereas the relatively deep circuit T is performed virtually and noiselessly using a classical computer.

We first show how to effectively measure H_T. In general, the target Hamiltonian H could be expressed as a linear sum of multiqubit Pauli terms $H = \sum_{i=1}^m g_i P_i$, where $P_i \in \{\sigma_I, \sigma_X, \sigma_Y, \sigma_Z\}^{\otimes n}$. Then we can measure each P_i with a total number of samples $(m/\epsilon^2)\sum_i g_i^2 \mathrm{Var}[P_i]$, proportional to the number of terms m in the Hamiltonian [28], to evaluate the energy expectation value within an error of ϵ. Here $\mathrm{Var}[P_i] = \langle P_i^2\rangle - \langle P_i\rangle^2$. We can similarly measure H_T, by similarly decomposing each $T^\dagger P_i T$ into Pauli strings. While most practical Hamiltonians H only contain a polynomial number of terms, this might not be the case for $T^\dagger P_i T$ or H_T, after the transformation [29].

Here we propose a structure of the Heisenberg circuit that also leads to efficiently measurable $T^\dagger P_i T$ or H_T. The circuit consists of two parts (Fig. 1(a)), where the first part is an arbitrary Clifford circuit that can be decomposed into a sequence of $O(n^2)$ basic gates from the set $\{\mathrm{H}, \mathrm{S}, \mathrm{CNOT}\}$, and the second part is a layer of single-qubit gates. The first

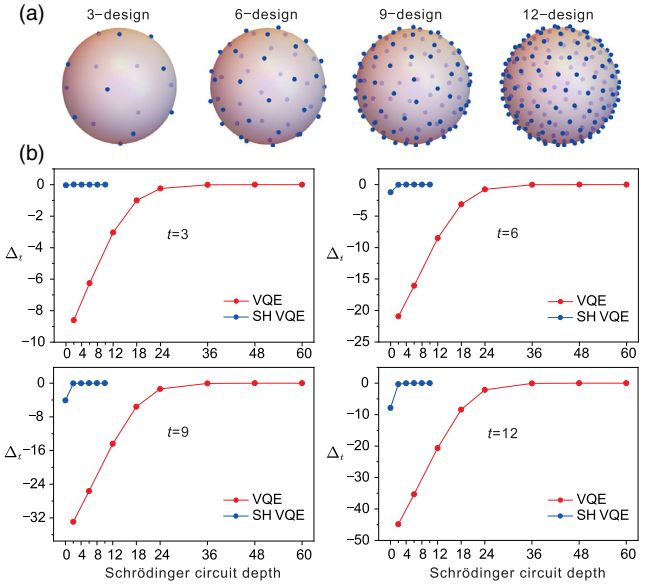

FIG. 2. SHVQE expressivity. (a) Relationship between expressivity and the t design. We show the point distribution on the Bloch sphere of different design orders $t = 3, 6, 9, 12$. (b) Comparison of the expressivity measure Δ_t between VQE and SHVQE. The structure of the Clifford part is formed by 500 randomly picked basic Clifford gates in the set $\{\mathrm{H}, \mathrm{S}, \mathrm{CNOT}\}$. The other parts including the two-qubit blocks in Schrödinger LUC and gates in the Heisenberg single qubit layer are random gates drawn from the Haar measure. The zero-depth setting in SHVQE can be understood as the performance of the Clifford circuit. Since Clifford circuit can generate 3-design, Δ_t approaches 0 for $t = 3$ whereas below 0 in other cases. Schrödinger circuit of depth greater than 6 combined with the Heisenberg circuit is believed to generate the maximally scrambled states since values of Δ_t from $t = 3$ to $t = 12$ are all zero [35].

part realizes discrete gates such as CNOT to build correlations between any 2 qubits, and the second part makes them continuous. The Clifford circuit maps the multiqubit Pauli group to itself, which conserves the number of terms of the Hamiltonian. Also, the Gottesman-Knill theorem [30] indicates that calculating the transformed Hamiltonian is easy. While the second part might increase the number of terms of the Hamiltonian, the overhead is polynomial for Hamiltonians H consisting of only k-weight terms, i.e., the Pauli operators $\{\sigma_X, \sigma_Y, \sigma_Z\}$ act on at most k qubits since the weight remains unchanged. We note that one can change this part into other easier or more complex circuits for different Hamiltonians, considering the trade-off between the circuit power and the measurement cost.

We begin to study the expressivity of the circuit in SHVQE. We consider the expressivity measure using the method of quantum complex projective t design [31], which means that the distribution of the output states has equal moments up to the tth order to a Haar uniform distributed states from the whole Hilbert space. Intuitively, as illustrated in Fig. 2(a) [32], a higher t design indicates a more uniform and denser state distribution in the Hilbert space, and vice versa. In general, a LUC of depth $O(nt^{10})$ is needed to generate a t design [33], and the Clifford circuits can produce a three design [34]. Using the tight Page's theorem [35], we define the logarithmic difference of entanglement entropy as

$$\Delta_t = \log\left(E_{\mathrm{Haar}}[\mathrm{Tr}(\rho_{n/2}^t)]\right) - \log\left(E_{\mathrm{SH}}[\mathrm{Tr}(\rho_{n/2}^t)]\right) \quad (2)$$

to identify the order of expressivity of SHVQE, where $\rho_{n/2}$ is the reduced half system density matrix, E_{Haar} is the average over Haar random states, and E_{SH} is the average over the quantum states $TU|0^{\otimes n}\rangle$. If Δ_t increases and approaches 0, it means that $TU|0^{\otimes n}\rangle$ is a t design.

Figure 2(b) shows a comparison of the expressivity of SHVQE versus VQE through a numerical experiment on a 12-qubit system. In the VQE setting, we run a random LUC at different depths and calculate Δ_t to characterize the t design. In the SHVQE setting, we implement both the real Schrödinger circuits U and the virtual Heisenberg circuits T which are pure Clifford consisting of 500 random gates from $\{\mathrm{H}, \mathrm{S}, \mathrm{CNOT}\}$. The key observation for both cases is the critical depths when the Δ_t measure increases to and saturates at around 0. It is evident that the Δ_t curves for SHVQE rise much more rapidly than that for VQE for all t values from 3 to 12. The rising curve for SHVQE quickly hits the saturation point at a Schrödinger circuit depth of $\sim$2, while the VQE curve arrives at a much deeper depth of $\sim$36. This indicates that SHVQE can effectively reduce the gate depth by more than 1 order of magnitude to achieve the same level of expressivity. For a higher number of qubits, we expect an even more dramatic advantage, which can be inferred from a qubit-size dependent test of depth reduction as shown in Fig. D3 in the Supplemental Material [29]. The above results indicate that we can use current NISQ hardware to effectively run deep quantum circuits while maintaining high fidelity. Particularly, based on a 2-qubit gate fidelity of 99.5%, the SHVQE can allow us to run, for instance, a 12-qubit four-depth quantum circuit with an output fidelity of 90%, which would otherwise demand a 2-qubit gate fidelity of 99.95% (currently unrealistic) and depth of 40 in conventional VQE (Fig. D2 in the Supplemental Material [29]). Note that shallow LUCs or Clifford circuits alone can only generate small design orders, but a combination of them can achieve high expressivity.

We consider an example of the XXZ spin model with a periodic boundary condition,

$$H_{XXZ} = \sum_{i=1}^{n} [\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z], \quad (3)$$

to demonstrate a kind of working flow of SHVQE. At the critical point $\Delta = 1$, the XXZ model is equivalent to the Heisenberg model whose ground state has a logarithmic

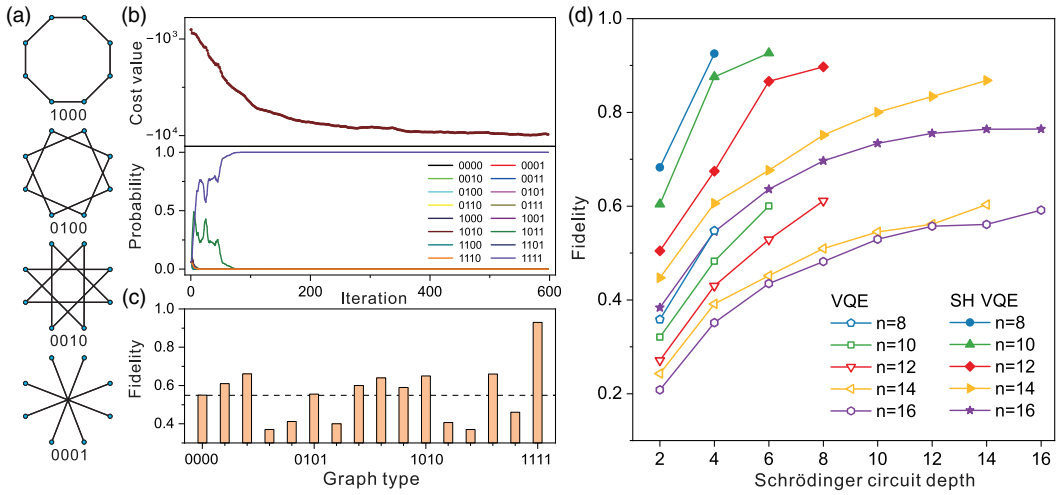

FIG. 3. Searching the Clifford circuit for the XXZ model. (a): The 4 elementary graphs and their corresponding code strings for $n = 8$ TI graphs. (b): Upper Panel: Minimizing the cost function to search for the best graph. Parameters contain both gate parameters and probability parameters. The cost function is the sum of the Hamiltonian expectation values of circuits sampled from $\vec{\alpha}$ under the same gate parameters. The number of samples at each iteration is 800. Lower Panel: Probabilities of all 16 graphs as functions of iteration times. All graphs have the same probabilities at the beginning. The probability of the fully connected graph "1111" becomes 1 as the iteration times grow. The optimization algorithm used for circuit structure searching is adam-SPSA [40]. (c): Direct comparisons between different graphs. The dashed line is the result of the graph type: "0000" i.e. without the Heisenberg circuit. The fully connected graph is indeed the best choice. There exist graphs that have worse performance than "0000". (d): Comparison of solved ground state fidelities. We use VQE and SHVQE to solve 8, 10, 12,14, and 16-qubit XXZ models. Fully connected graphs are used as the Clifford layer. VQE and SHVQE of the same Schrödinger circuit depth share the same color. Each point is the best result obtained from 20 sets of random initial parameters.

scaling of entanglement entropy [22,23], and hence cannot be prepared by a constant-depth LUC. Since we aim to boost the performance of the NISQ experiments, we use the hardware efficient ansatz [14] for the real Schrödinger circuit even though this may lead to barren plateau problems [36], where each circuit layer composes a layer of CZ gates and a layer of parametrized arbitrary single-qubit gates (denoted as $\vec{\theta}$) .

For the Heisenberg circuit, the single-qubit gate layer is parametrized with parameters $\vec{\phi}$. And we restrict the Clifford part to graph circuits [37] where only commuting CZ gates are used. We separate the graph circuit into patterns of different connectivity with the same translational invariant (TI) symmetry as H_{XXZ}. More concretely, for an n-qubit circuit, we can set $\lfloor n/2 \rfloor$ elementary graphs (For the j^{th} elementary graph, each node I is connected with node $i + j$. $\lfloor \cdot \rfloor$ is the floor function.). As each elementary graph can be turned on or turned off, the total number of possible patterns is $2^{\lfloor n/2 \rfloor}$, and we use a $\lfloor n/2 \rfloor$-bit string to label all the possible patterns such as "01001…," where 0 means the corresponding elementary graph is turned on whereas 1 means off [Fig. 3(a)]. To efficiently search through an exponentially large space of Clifford gate patterns, we borrow the idea from differentiable quantum architecture search [38], where each elementary TI graph is turned on independently according to a probability described by a two-parameter softmax function [39]. Thus, only $\lfloor n/2 \rfloor \times 2$ parameters (denoted as $\vec{\alpha}$) are needed to implement the discrete search of the huge Clifford patterns. Therefore, the circuit ansatz for the SHVQE is

$$T(\vec{\alpha}, \vec{\phi})U(\vec{\theta})|0^{\otimes n}\rangle, \tag{4}$$

where $\vec{\alpha}$ and $\vec{\phi}$ represent all configurations of the Heisenberg circuit T and $\vec{\theta}$ are the continuous parameters in the single-qubit gates inside the Schrödinger circuit U. The parameters $\vec{\alpha}$ are used to generate samples of different circuits, and the cost function is the average of the Hamiltonian expectation values of these circuits under the same gate parameters $\vec{\theta}$ and $\vec{\phi}$. The SHVQE method then optimizes over all the parameters to search for the ground state of the Hamiltonian.

In our numerical simulation, we consider an eight-spin XXZ model with a four-depth circuit U and four elementary TI graphs of the Clifford layer as shown in Fig. 3(a). We show the energy expectation and the evolution of the possibilities of all 16 configurations during the optimization as functions of the number of iterations in Fig. 3(b). When the energy expectation is converged, the probabilities of the candidate circuit structures concentrate on the optimal configuration, the fully connected graph "1111." In Fig. 3(c), we show the optimal energies of all the 16 candidate circuit configurations, which verifies that the

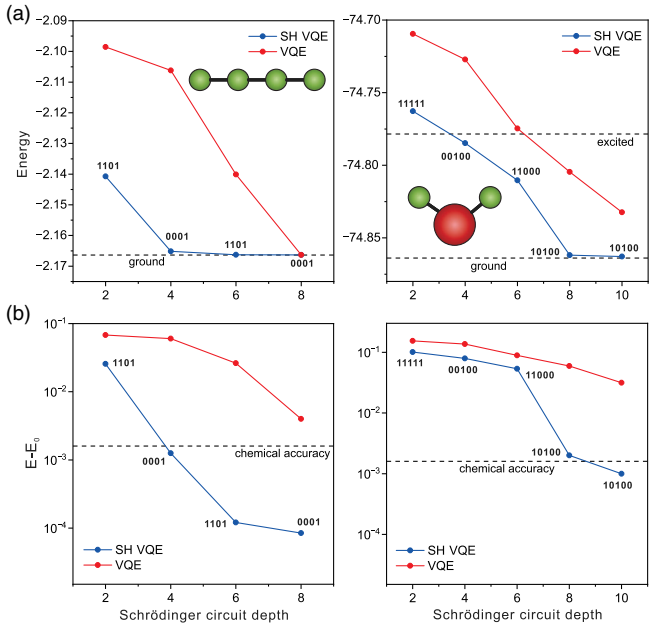

FIG. 4. Comparisons of SHVQE and VQE on solving small molecules. Using VQE and SHVQE to solve the 8-qubit Hamiltonian of the H_4 molecule of the bond distance 1.0 Angstrom and the 10-qubit Hamiltonian of the H_2O molecule of the bond distance 1.5 Angstrom. The binary string around the SHVQE data label the searched optimal graph circuit. (a): Solved energy as a function of Schrödinger circuit depth. (b): Absolute energy differences as functions of Schrödinger circuit depth.

optimal configuration is indeed the fully connected graph "1111." We further solve larger models up to 16 spins to show the improvement of SHVQE compared with conventional VQE using the same Schrödinger circuits [Fig. 3(d)]. For SHVQE, we directly use the generalized fully connected graph circuits as the Clifford part. We can find under the same circuit depth, the SHVQE obtains higher fidelities than the VQE (an average improvement of 25.2%).

To further demonstrate the practical values of our algorithm, we implement our algorithm to solve the electronic structure problems of H_4 and H_2O molecules following the same workflow as above. The H_4 molecule corresponds to an 8-qubit Hamiltonian. For the H_2O molecule, we use the active space method [41] to create an effective 10-qubit Hamiltonian containing ten spin orbitals and six electrons. Since the SHVQA has the Pauli weight restriction, we use the Bravyi-Kitaev mapping which transforms an M-mode fermionic Hamiltonian to a spin Hamiltonian of $O(\log_2 M)$ Pauli weight [41,42]. Note that the ground states of these molecule Hamiltonians have the correct number of electrons. The results are shown in Fig. 4, where we can see SHVQE can reach the chemical accuracy (1.6×10^{-3}) with Schrödinger circuits of much shallower depth than VQE.

We now give some discussions for SHVQA. First, we want to emphasize that the states $TU|0^{\otimes n}\rangle$ are both hard to prepare on NISQ devices, as it requires implementing the relatively deep T circuit, and hard to simulate on classical computers, as it can be treated as Clifford circuits with nonstabilizer input states. However, interestingly, within the SHVQE framework, the operator expectation values under these states can be efficiently evaluated as long as U is classically tractable. Second, we want to talk about the trainability of SHVQA. A known result is that in general, an ansatz with high expressivity may lead to low trainability [43]. We want to emphasize that the expressivity benchmarked under the very random settings in Fig. 2 should be understood as the achievable expressivity of the NISQ devices enhanced by Heisenberg circuits but not the actual expressivity of the ansatz within the SHVQA framework for specific problems. Thus, SHVQA can be understood as a general methodology for improving existing variational algorithms within which biased and trainable ansatzes can be tested. We summarize some strategies in the Supplemental Material [29].

In summary, we have introduced a novel variational quantum algorithm, the SHVQA, to efficiently extend the circuit depth of near-term noisy quantum processors. By virtually introducing relatively deep and nonlocal Clifford circuits, we show that the expressivity of shallow quantum circuits can be significantly enhanced, without sacrificing the fidelity. We use the XXZ model to demonstrate the workflow of SHVQA and further demonstrate the practical values of SHVQA by solving small molecules. Our method is directly applicable to current quantum hardware and is compatible with most existing quantum algorithms. Leveraging quantum error mitigation, our work pushes near-term quantum hardware into wide nontrivial applications.

We use the QULACS [44] and the QISKIT [45] packages for parts of simulations.

This work was supported by the National Natural Science Foundation of China (No. 91836303 and No. 11805197), the National Key R&D Program of China, the Chinese Academy of Sciences, the Anhui Initiative in Quantum Information Technologies, and the Science and Technology Commission of Shanghai Municipality (2019SHZDZX01).

Z.-X. S and M.-C. C contributed equally to this work.

[1] Richard P. Feynman, Simulating physics with computers, Int. j. Theor. phys **21**, 133 (2018).

[2] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[3] Qingling Zhu, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, Ming Gong *et al.*, Quantum computational

advantage via 60-qubit 24-cycle random circuit sampling, Sci. Bull. **67**, 240 (2022).
[4] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan *et al.*, Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, Phys. Rev. Lett. **127**, 180501 (2021).
[5] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).
[6] H. S. Zhong, Y. H. Deng, J. Qin, H. Wang, M. C. Chen, L. C. Peng, Y. H. Luo, D. Wu, S. Q. Gong, H. Su *et al.*, Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light, Phys. Rev. Lett. **127**, 180502 (2021).
[7] John Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).
[8] Suguru Endo, Zhenyu Cai, Simon C. Benjamin, and Xiao Yuan, Hybrid quantum-classical algorithms and quantum error mitigation, J. Phys. Soc. Jpn. **90**, 032001 (2021).
[9] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio *et al.*, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).
[10] 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).
[11] Xiao Yuan, Suguru Endo, Qi Zhao, Ying Li, and Simon C. Benjamin, Theory of variational quantum simulation, Quantum **3**, 191 (2019).
[12] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.
[13] Dave Wecker, Matthew B. Hastings, and Matthias Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).
[14] 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 (London) **549**, 242 (2017).
[15] 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).
[16] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, Phys. Rev. Lett. **119**, 180509 (2017).
[17] Ying Li and Simon C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, Phys. Rev. X **7**, 021050 (2017).
[18] Fernando G. S. L. Brandao and Michał Horodecki, Exponential decay of correlations implies area law, Commun. Math. Phys. **333**, 761 (2015).
[19] Frank Verstraete, Valentin Murg, and J. Ignacio Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, Adv. Phys. **57**, 143 (2008).
[20] Sergey Bravyi, Matthew B. Hastings, and Frank Verstraete, Lieb-Robinson Bounds and the Generation of Correlations and Topological Quantum Order, Phys. Rev. Lett. **97**, 050401 (2006).
[21] Wen Wei Ho and Timothy H. Hsieh, Efficient variational simulation of non-trivial quantum states, SciPost Phys. **6**, 029 (2019).
[22] Guifre Vidal, José Ignacio Latorre, Enrique Rico, and Alexei Kitaev, Entanglement in Quantum Critical Phenomena, Phys. Rev. Lett. **90**, 227902 (2003).
[23] José Ignacio Latorre, Enrique Rico, and Guifré Vidal, Ground state entanglement in quantum spin chains, Quantum Inf. Comput. **4**, 48 (2004).
[24] Yichen Huang, Xie Chen *et al.*, Quantum circuit complexity of one-dimensional topological phases, Phys. Rev. B **91**, 195143 (2015).
[25] Xie Chen, Zheng-Cheng Gu, and Xiao-Gang Wen, Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order, Phys. Rev. B **82**, 155138 (2010).
[26] Shi-Xin Zhang, Zhou-Quan Wan, Chee-Kong Lee, Chang-Yu Hsieh, Shengyu Zhang, and Hong Yao, Variational Quantum-Neural Hybrid Eigensolver, Phys. Rev. Lett. **128**, 120502 (2022).
[27] Wataru Mizukami, Kosuke Mitarai, Yuya O. Nakagawa, Takahiro Yamamoto, Tennin Yan, and Yu-ya Ohnishi, Orbital optimized unitary coupled cluster theory for quantum computer, Phys. Rev. Res. **2**, 033421 (2020).
[28] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New J. Phys. **18**, 023023 (2016).
[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.131.060406 for the measurement cost, the structure of parametrized circuit, the expressivity and the trainability of SH-VQA, and the circuit expressivity and hardware requirements.
[30] Daniel Gottesman, The Heisenberg representation of quantum computers, arXiv:quant-ph/9807006.
[31] Stuart G. Hoggar, t-designs in projective spaces, Eur. J. Combinatorics **3**, 233 (1982).
[32] Data source, https://www.polyu.edu.hk/ama/staff/xjchen/sphdesigns.html.
[33] Fernando G. S. L. Brandao, Aram W. Harrow, and Michał Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).
[34] Huangjun Zhu, Multiqubit Clifford groups are unitary 3-designs, Phys. Rev. A **96**, 062336 (2017).
[35] Zi-Wen Liu, Seth Lloyd, Elton Zhu, and Huangjun Zhu, Entanglement, quantum randomness, and complexity beyond scrambling, J. High Energy Phys. 07 (2018) 041.
[36] 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).

[37] Marc Hein, Wolfgang Dür, Jens Eisert, Robert Raussendorf, M. Nest, and H.-J. Briegel, Entanglement in graph states and its applications, arXiv:quant-ph/0602096.

[38] Shi-Xin Zhang, Chang-Yu Hsieh, Shengyu Zhang, and Hong Yao, Differentiable quantum architecture search, Quantum Sci. Technol. **7**, 045023 (2022).

[39] The possibility of picking a graph $\vec{k} = k_1 k_2 \ldots k_{\lfloor n/2 \rfloor}$ is described by the product of $\lfloor n/2 \rfloor$ independent softmax functions $\prod_{i=1}^{\lfloor n/2 \rfloor} \{\exp(\alpha_{i,k_i})/[\exp(\alpha_{i,0}) + \exp(\alpha_{i,1})]\}$.

[40] Yazan Arouri and Mohammad Sayyafzadeh, An accelerated gradient algorithm for well control optimization, J. Pet. Sci. Eng. **190**, 106872 (2020).

[41] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[42] Sergey B. Bravyi and Alexei Yu Kitaev, Fermionic quantum computation, Ann. Phys. (Amsterdam) **298**, 210 (2002).

[43] Zoë Holmes, Kunal Sharma, Marco Cerezo, and Patrick J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, PRX Quantum **3**, 010313 (2022).

[44] Yasunari Suzuki, Yoshiaki Kawase, Yuya Masumura, Yuria Hiraga, Masahiro Nakadai, Jiabao Chen, Ken M. Nakanishi, Kosuke Mitarai, Ryosuke Imai, Shiro Tamiya *et al.*, Qulacs: A fast and versatile quantum circuit simulator for research purpose, Quantum **5**, 559 (2021).

[45] Andrew Cross, The IBM Q experience and qiskit open-source quantum computing software, in *APS March Meeting Abstracts* (2018), Vol. 2018, pp. L58–003.