Absence of Barren Plateaus in Finite Local-Depth Circuits with Long-Range Entanglement

Hao-Kai Zhang,[1,*,†] Shuo Liu,[1,*] and Shi-Xin Zhang[2]
[1]Institute for Advanced Study, Tsinghua University, Beijing 100084, China
[2]Tencent Quantum Laboratory, Tencent, Shenzhen, Guangdong 518057, China

(Received 6 November 2023; revised 24 January 2024; accepted 13 March 2024; published 9 April 2024)

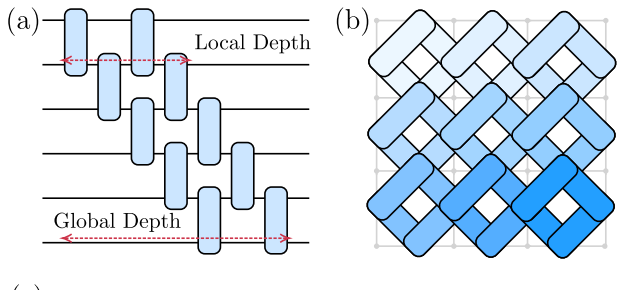
Ground state preparation is classically intractable for general Hamiltonians. On quantum devices, shallow parametrized circuits can be effectively trained to obtain short-range entangled states under the paradigm of variational quantum eigensolver, while deep circuits are generally untrainable due to the barren plateau phenomenon. In this Letter, we give a general lower bound on the variance of circuit gradients for arbitrary quantum circuits composed of local 2-designs. Based on our unified framework, we prove the absence of barren plateaus in training finite local-depth circuits (FLDC) for the ground states of local Hamiltonians. FLDCs are allowed to be deep in the conventional circuit depth to generate long-range entangled ground states, such as topologically ordered states, but their local depths are finite, i.e., there is only a finite number of gates acting on individual qubits. This characteristic sets FLDC apart from shallow circuits: FLDC in general cannot be classically simulated to estimate local observables efficiently by existing tensor network methods in two and higher dimensions. We validate our analytical results with extensive numerical simulations and demonstrate the effectiveness of variational training using the generalized toric code model.

DOI: 10.1103/PhysRevLett.132.150603

Introduction.—Predicting the ground state properties of a quantum many-body system, as a central task in modern quantum physics, generally requires exponential resources for classical computers due to the curse of dimensionality: the number of parameters needed to describe a quantum system scales exponentially with the system size. Although some successful classical algorithms have been developed in past decades [1–5] such as tensor networks [3–5], their respective limitations restrict the performance on general systems [5–8]. Quantum computers bring new hope for this problem of quantum nature [9]. Despite the limitation posed by noisy intermediate-scale quantum (NISQ) devices [10], there are many tentative quantum algorithms proposed. One of the representatives is the variational quantum eigensolver (VQE) [11–17], which trains a parametrized quantum circuit (PQC) using a classical optimizer to minimize the energy. This hybrid quantum-classical paradigm is expected as one of the most promising routes toward practical quantum advantage [18,19] in the NISQ era.

However, these variational quantum algorithms including VQE still face great challenges for large-scale applications. One of the most notorious issues is the so-called barren plateau phenomenon [20], which states that the circuit gradient vanishes exponentially with the system size under certain conditions, akin to the vanishing gradient issue in classical neural networks. The exponentially vanishing gradient will preclude the optimization progress and lead to the exponential measurement complexity. Extensive studies have been conducted to investigate barren plateau problems and possible remedies [21–49]. It is known that shallow circuits of finite or logarithmic depth are free from barren plateaus and can be trained efficiently for local Hamiltonians to obtain short-range entangled (SRE) states, while deep circuits of linear depth and beyond are in general untrainable [29,30]. By contrast, many quantum states of physical interest exhibit long-range entanglement [50–54], such as topologically ordered states, which cannot be prepared by circuits of less than linear depth [55–58]. Nevertheless, some evidence suggests that circuits corresponding to these long-range entangled (LRE) states possess characteristic architectures, such as sequential structures [57–64], for the sake of the entanglement area law. This observation motivates us to rigorously explore the general relationship between barren plateaus, area law, and long-range entanglement.

In this Letter, we identify the critical role of the local depth as a key circuit feature that determines the trainability of PQCs. The local depth refers to the number of noncommuting gates acting on individual qubits, as illustrated in Fig. 1(a), in contrast to the conventional global depth defined by the minimum number of layers. This finding is based on our rigorously proved theorems, which establish a general lower bound on the gradient variance for arbitrary circuits composed of local 2-designs. The lower bound decays exponentially with the length and width of a certain set of paths on the circuit. For finite (or logarithmic) local-depth circuits and local Hamiltonians, the length and width can be upper bounded and hence give rise to the absence of barren plateaus. These finite local-depth circuits (FLDC) have strong expressibility to generate LRE states lacking in



(c)

	FDC	FLDC	GLDC
BP-free	✓	✓	×
Area law	✓	✓	×
LRE	×	✓	✓
Non-classicality	×	✓	✓

FIG. 1. (a) and (b) Typical examples of finite local-depth circuits on 1D and 2D lattices, respectively. Darker colors in (b) indicate later action orders. (c) Compares the class of finite depth circuit (FDC), FLDC, and general linear depth circuit (GLDC) in terms of whether they are in general free from barren plateaus (BP), preserve entanglement area law, generate long-range entanglement (LRE), and can be simulated efficiently to compute local observable expectations by known classical methods (classicality). The inclusion relation is FDC ⊂ FLDC ⊂ GLDC.

shallow circuits and are hard to simulate classically in two dimensions and above. This suggests that FLDC holds promise to serve as an appropriate class of ansatzes in VQE, as listed in Fig. 1(c). The absence of barren plateaus in FLDC is verified by numerical evaluations. Using the generalized 2D toric code model, we demonstrate that FLDC indeed has prominently better performance than both finite depth circuits and general linear depth circuits.

Basic setup.—We start from the basic setup of VQE. The PQC can be written as $\mathbf{U}(\boldsymbol{\theta}) = \prod_{\mu=1}^{M} U_\mu(\theta_\mu)$, where $U_\mu(\theta_\mu) = e^{-i\Omega_\mu\theta_\mu}$ is a rotation gate, Ω_μ is a Pauli-string generator and θ_μ is a trainable parameter. The index μ follows the decreasing order from left to right in the product (the same below). For a given Hamiltonian H, the energy expectation $C(\boldsymbol{\theta}) = \langle H\rangle = \mathrm{tr}(\rho_0\mathbf{U}^\dagger H\mathbf{U})$ is taken as the cost function, where $\rho_0 = |\mathbf{0}\rangle\langle\mathbf{0}|$ and $|\mathbf{0}\rangle = |0\rangle^{\otimes N}$. N is the number of qubits. We denote the Pauli decomposition of the Hamiltonian as $H = \sum_j \lambda_j h_j$ and assume the support of H is within that of $\mathbf{U}$. The workflow of VQE involves running the PQC, measuring the cost function, and updating the trainable parameters iteratively using classical optimizers to minimize the cost function. In particular, the parameters are usually initialized randomly to thoroughly explore the parameter space in a probabilistic sense, rendering the PQC a random quantum circuit (RQC). A common assumption on RQCs is that the circuit is composed of blocks forming independent local 2-designs. Here a block refers to a grouped continuous series of gates, which can be seen as the elementary unit when we construct a PQC. Grouping the M gates into M' blocks,

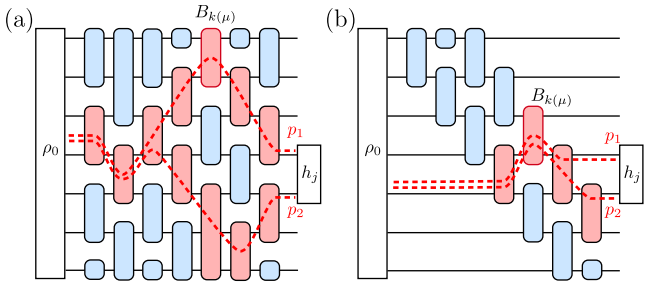


FIG. 2. (a) A possible choice of path set $P_j = \{p_1, p_2\}$ on a general linear depth circuit. (b) Depicts a path set on a finite local-depth circuit correspondingly. The length of the path set in (a) grows linearly with the system size while that in (b) is bounded by the constant local depth.

the PQC can be rewritten as $\mathbf{U} = \prod_{k=1}^{M'} B_k$. The assumption of local 2-designs will induce an ensemble of the entire circuit $\mathbf{U}$, which we denote as $\mathbb{U}$. Many statistical properties of RQCs can be analytically estimated based on $\mathbb{U}$, including the average and variance of the cost derivative $\partial_\mu C = (\partial C/\partial\theta_\mu)$. We provide preliminaries on unitary designs and the Weingarten calculus and a detailed introduction to our basic setup in Supplemental Material [65].

General lower bound.—We give an informal version of our general lower bound in Theorem 1 and leave the rigorous statement and proof to Supplemental Material [65]. The bound is closely related to a geometric concept of "'path," i.e., a time-ordered sequence of connected blocks on the circuit diagram as depicted in Fig. 2. For each Hamiltonian subterm h_j in the causal cone of the differential block $B_{k(\mu)}$ (the block containing the differential parameter θ_μ), one can draw a collection of paths from h_j to ρ_0, like $\{p_1, p_2\}$ in Fig. 2, with the right end covering h_j and at least one of the paths passing through $B_{k(\mu)}$. We call it a chosen "path set" of h_j. We define two measures of the path set: length and head width. For common circuits composed of 2-qubit blocks, the length is just the number of edges in the path set diagram and the head width is the number of blocks in the path set that are directly connected to ρ_0. Using these two measures, we can derive the following lower bound on the gradient variance.

Theorem 1 (informal version).—The gradient variance $\mathrm{Var}_{\mathbb{U}}[\partial_\mu C]$ can be lower bounded by a summation of contributions from each h_j in the causal cone of U_μ, where each contribution decays exponentially only with the length and head width of the chosen path set of h_j.

The lower bound in Theorem 1 holds for any possible choice of path sets. The path set with the minimum length and head width gives rise to the tightest bound. We remark that Theorem 1 holds for any RQCs composed of local 2-designs regardless of circuit shapes, spatial dimensions, and gate locality. Further discussions on Theorem 1 including its consistency with previous literature, alternative initial states, gate generators, the location of the differential gate U_μ in $B_{k(\mu)}$, and its extension to a path-integral-like tighter form

and application to other space-time correlators, are elaborated in Supplemental Material [65].

We provide an intuitive physical picture behind Theorem 1. It is known that local quantum information will be scrambled [101–103] when passing through the random gates in an RQC. The more gates it passes through, the more severe the scrambling becomes. If we consider a local term h_j as a piece of information, finding a short path through the RQC will allow effective information transfer, so that adjusting parameters can make an effective difference in the expectation value, resulting in nonvanishing gradients. Conversely, if such a short path does not exist, local information will be scrambled globally, leaving no useful information for optimization.

Finite local-depth circuits.—Before presenting Theorem 2, we first clarify some relevant quantities. The maximum interaction range of a Hamiltonian is the maximum value of the support sizes of all h_j. An r-local Hamiltonian means that the maximum interaction range is fixed as r that does not scale with N. The maximum block size β is the maximum value of the support sizes of all blocks in the circuit. The local depth of a qubit is the number of blocks (or gates) acting on the qubit. We use χ to denote the maximum value of the local depths over all qubits, distinguished from the global depth D which refers to the minimum number of layers where blocks within each layer commute with each other. An FLDC is defined as a circuit whose χ does not scale with N, without any other constraints such as circuit shapes, spatial dimensions, and gate locality. Based on Theorem 1, we have the following theorem.

Theorem 2.—Suppose the maximum local depth of $\mathbf{U}$ is χ and the maximum block size is β. Then for any r-local Hamiltonian, the gradient variance is lower bounded by

$$\mathrm{Var}_{\mathbb{U}}[\partial_\mu C] \geq 4^{-r\chi\beta} \sum_j 2\lambda_j^2, \tag{1}$$

where j runs over h_j that is nontrivial on the support of the differential block $B_{k(\mu)}$.

Proof.—The detailed proof is left in Supplemental Material [65]. The main idea is choosing the path sets in Theorem 1 to be the straight wires on the support of h_j, and hence the length and head width can be upper bounded in terms of r, χ and β. The contribution from h_j that is trivial on the support of $B_{k(\mu)}$ is just neglected. ∎

Theorem 2 elegantly integrates the factors related to barren plateaus in a concise manner, i.e., the block locality β [20], the Hamiltonian locality r [29,30] and the circuit deepness χ. It is vitally important to note that the relevant quantity characterizing the circuit deepness is the local depth χ, instead of the global depth D. These two depths may coincide [29,30], but they are distinct in general and can differ significantly as in Fig. 1. This implies that the circuit class free from barren plateaus can be enlarged to logarithmic local-depth circuits (Log-LDC), which is a superclass of circuit architectures proven previously, such as finite or logarithmic depth brickwall circuits [29,30], quantum convolutional neural networks (QCNN) [31,41], multiscale entanglement renormalization ansatzes (MERA) [41–44], tree tensor networks [41–44], matrix product states (MPS) [40–44], and high-dimensional isometric tensor network states [47]. We focus on FLDC in this Letter. Log-LDC class involves states beyond the area law, e.g., gapless topologically ordered states, which is also interesting to study in the future.

A significant feature of FLDCs composed of spatially local gates is that the generated quantum states satisfy the entanglement area law (or say boundary law) because the number of gates acting across any simple partition boundary entangling the two sides can be upper bounded by the local depth times the size of the boundary [65]. This feature makes them form a subclass of the projected entangled paired states (PEPS) [3,5] of the corresponding spatial dimension, where the local depth χ plays the role of bond dimension. Note that PEPS can represent LRE states because the nonunitary projectors in PEPS enable quantum teleportation, while FLDC relies on large global depth. Previously proposed circuits of tensor network states [41–44,57,61,62,64,104–107] including sequential quantum circuits [57], isometric tensor network states [47,61,64,108–110], and plaquette PEPS [62], can all be seen as subclasses of FLDC. This implies that FLDC covers a wide range of physical ground states such as string-net states with anyons [64,111] and fracton-ordered states [57].

Nonclassicality of FLDC.—A matter of recent concern is the classical simulability of the tasks with the provable absence of barren plateaus [112]. Previous results that are proven free from barren plateaus mainly focus on finite or logarithmic depth circuits [29–31,41–44], which can be efficiently simulated to compute local observable expectations due to the existence of small causal cones and small tree widths of the corresponding tensor networks [113]. Nevertheless, the causal cone in FLDC can be extensive due to the large global depth, and the loop structures in FLDC of two dimensions and above can lead to polynomially large tree width, rendering FLDC in general hard to simulate classically for local observable expectations (the 1D case has constant tree width and can be efficiently simulated via MPS methods). In fact, even for the subclasses of 2D FLDC such as isometric tensor network states [61,64,108] and plaquette PEPS [62], there is no known efficient method to compute the expectation values of arbitrary local observables with controllable error, not to mention nonlocal observables of interest such as few-body long-distance correlators, nonlocal order parameters, dynamical correlations and so forth. In particular, a recent work [114] rigorously proves that computing local expectation values in isometric tensor networks is BQP complete, i.e., is hard to simulate classically unless BQP = BPP.

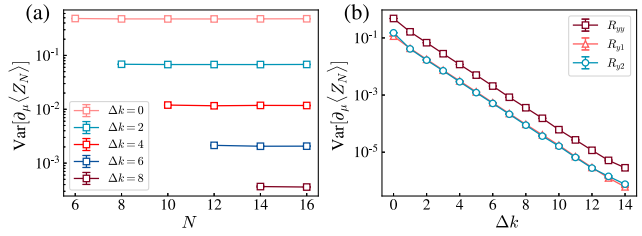


FIG. 3. (a) The variance of the derivative vs the system size N in a 1D FLDC instance. The observable is chosen as Z_N. Δk is the distance between the differential block and the last block, proportional to the path length. (b) The variance vs Δk by fixing $N = 16$. R_{yy}, R_{y1}, and R_{y2} represent the different choices of the differential gate [65].

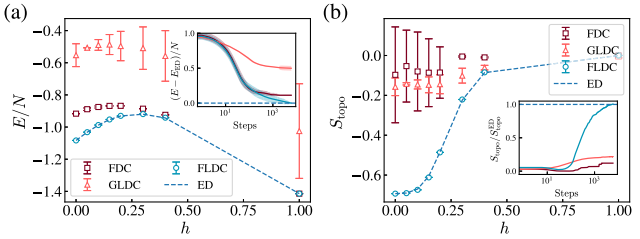


FIG. 4. VQE performance comparison of the FDC, FLDC, and GLDC ansatzes using the generalized toric code model under the external field $h^z = h^x = h$ with $N = 12$. The data are averaged over the best half of the 100 training trajectories starting from different initializations. (a) The converged energy E/N vs h. The inset depicts the energy training dynamics at $h = 0.1$. The dashed lines represent the exact values obtained from ED. The (shaded) error bar represents the standard deviation. (b) The topological entanglement entropy S_{topo} correspondingly.

The same naturally holds true for FLDC because the states generated by FLDC form a superclass of isometric tensor network states.

Therefore, FLDC (or Log-LDC) is a circuit class that is proven to be barren-plateau-free and at the same time generally cannot be efficiently simulated to estimate local observables by existing classical methods. On the contrary, it can be accomplished within polynomial time by running FLDCs on quantum devices and measuring corresponding observables. This suggests that FLDC is potentially relevant to quantum advantage in the ground state preparation task. A detailed discussion and a numerical demonstration of the computational overhead for contracting tensor networks of FDLC are provided in Supplemental Material [65].

Numerical experiments.—FLDC has stronger expressibility than its subclass finite depth circuits (FDC), e.g., brickwall circuits of constant depth [29,30], because FDC can only generate SRE states such as symmetry-protected topological states [115]. FLDCs have less expressibility than its superclass general linear depth circuits (GLDC), e.g., brickwall circuits of linear depth, as typical GLDCs lead to entanglement volume law [28]. But FLDC has better trainability than GLDC. We will compare the variational performance of the three circuit classes to see the advantages and the good trade off between trainability and expressibility brought by FLDC.

To demonstrate the absence of barren plateaus in FLDC, we estimate the cost gradient in a 1D FLDC ansatz with a ladder layout as in Fig. 1(a). The two-qubit block template is chosen as the Cartan decomposition [65]. The Hamiltonian is chosen as a single Pauli Z operator on the last qubit. All the numerical experiments are implemented using TensorCircuit [116]. As depicted in Fig. 3, the gradient variance is almost constant with the system size, while it decays exponentially with the path length Δk. This resembles the phenomenon found in isometric tensor networks recently [40,43,44,47]. However, we clarify that the exponential decay with Δk does not indicate poor trainability in practice, because as long as the gradients of some circuit parameters do not vanish, the optimization could still proceed successfully.

As an example of training FLDCs for LRE ground states, we use the generalized 2D toric code model under the external field $\boldsymbol{h} = (h^x, h^y, h^z)$ with open boundary conditions. The ground state near the zero-field limit is topologically ordered and then experiences a quantum phase transition to an SRE state with increasing $\boldsymbol{h}$ [117,118]. The ground state at $\boldsymbol{h} = 0$ can be constructed by applying the Hadamard and CNOT gates sequentially [56,57], which belongs to the FLDC class. Possible generalization to $h^z \neq 0$ has also been proposed [58]. However, unlike in Ref. [58], we will not utilize any prior information about the exact ground state except the entanglement area law. Namely, we choose our ansatz to be an FLDC similar to Fig. 1(b), with each two-qubit block being the general Cartan decomposition. We also conducted the same simulation using typical ansatzes in FDC and GLDC for comparison. As shown in Fig. 4(a), the energies of FLDC almost coincide with the exact values from the exact diagonalization (ED). By contrast, the energies of GLDC severely deviate due to poor trainability. On the other hand, although FDC does not suffer from barren plateaus, it lacks the expressibility to represent LRE states faithfully, so FDC works well in the large field limit but deviates near the zero-field limit. We also show the results of the topological entanglement entropy S_{topo} of these variational states in Fig. 4(b) correspondingly. The technical details and additional numerical results can be found in Supplemental Material [65].

Discussion.—In this Letter, we prove a general lower bound on the gradient variance for arbitrary quantum circuits composed of local 2-designs, which unifies the known gradient scaling behaviors of various architectures. An intuitive physical picture emerges that relates the nonvanishing gradients with the information scrambling in RQCs along certain path sets. We further prove the

absence of barren plateaus for local Hamiltonians in a new circuit class—finite local-depth circuits, which can generate LRE states thanks to large global depths. FLDCs composed of spatially local gates preserve the entanglement area law, which makes it form a powerful and accessible subclass of PEPS that covers a wide range of physical ground states. Importantly, FLDC cannot be classically simulated efficiently in two and higher dimensions by the known tensor network methods. We remark that the indication of local depth is also instructive in developing quantum architecture search schemes [119–122]. Finally, we point out that the absence of barren plateaus is a necessary but not sufficient condition for the effectiveness of training. There are other challenging issues such as the local minimum problem [123–125]. Enhancing the VQE performance of FLDCs in more general systems requires further exploration in future studies.

We acknowledge stimulating discussions with Xie Chen, Jing-Yu Zhao, Zhi-Yuan Wei, Yu-Jie Liu, and Zhehao Dai. This work was supported in part by the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302502)

*These two authors contributed equally to this work.
†zhk20@mails.tsinghua.edu.cn

[1] D. Ceperley and B. Alder, Quantum Monte Carlo, Science **231**, 555 (1986).
[2] S. R. White, Density matrix formulation for quantum renormalization groups, Phys. Rev. Lett. **69**, 2863 (1992).
[3] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Criticality, the area law, and the computational power of projected entangled pair states, Phys. Rev. Lett. **96**, 220601 (2006).
[4] G. Vidal, Entanglement renormalization, Phys. Rev. Lett. **99**, 220405 (2007).
[5] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. (Amsterdam) **326**, 96 (2011).
[6] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Sign problem in the numerical simulation of many-electron systems, Phys. Rev. B **41**, 9301 (1990).
[7] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Computational complexity of projected entangled pair states, Phys. Rev. Lett. **98**, 140506 (2007).
[8] J. D. Biamonte, J. Morton, and J. Turner, Tensor network contractions for #SAT, J. Stat. Phys. **160**, 1389 (2015).
[9] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).
[10] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).
[11] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).
[12] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature (London) **549**, 242 (2017).
[13] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).
[14] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods and best practices, Phys. Rep. **986**, 1 (2022).
[15] S.-X. Zhang, Z.-Q. Wan, C.-K. Lee, C.-Y. Hsieh, S. Zhang, and H. Yao, Variational quantum-neural hybrid eigensolver, Phys. Rev. Lett. **128**, 120502 (2022).
[16] S. Liu, S.-X. Zhang, C.-Y. Hsieh, S. Zhang, and H. Yao, Probing many-body localization by excited-state variational quantum eigensolver, Phys. Rev. B **107**, 024204 (2023).
[17] C. Tabares, A. Muñoz de las Heras, L. Tagliacozzo, D. Porras, and A. González-Tudela, Variational quantum simulators based on waveguide QED, Phys. Rev. Lett. **131**, 073602 (2023).
[18] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).
[19] H.-S. Zhong *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).
[20] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 1 (2018).
[21] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).
[22] K. M. Nakanishi, K. Fujii, and S. Todo, Sequential minimal optimization for quantum-classical hybrid algorithms, Phys. Rev. Res. **2**, 043158 (2020).
[23] M. Ostaszewski, E. Grant, and M. Benedetti, Structure optimization for parameterized quantum circuits, Quantum **5**, 391 (2021).
[24] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, Quantum **3**, 214 (2019).
[25] A. Arrasmith, M. Cerezo, P. Czarnik, L. Cincio, and P. J. Coles, Effect of barren plateaus on gradient-free optimization, Quantum **5**, 558 (2021).
[26] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, Nat. Commun. **12**, 6961 (2021).
[27] M. Cerezo and P. J. Coles, Higher order derivatives of quantum neural networks with barren plateaus, Quantum Sci. Technol. **6**, 035006 (2021).
[28] C. Ortiz Marrero, M. Kieferová, and N. Wiebe, Entanglement-induced barren plateaus, PRX Quantum **2**, 040316 (2021).
[29] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nat. Commun. **12**, 1791 (2021).

[30] A. V. Uvarov and J. D. Biamonte, On barren plateaus and cost function locality in variational quantum algorithms, J. Phys. A **54**, 245301 (2021).
[31] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles, Absence of barren plateaus in quantum convolutional neural networks, Phys. Rev. X **11**, 041011 (2021).
[32] A. Arrasmith, Z. Holmes, M. Cerezo, and P. J. Coles, Equivalence of quantum barren plateaus to cost concentration and narrow gorges, Quantum Sci. Technol. **7**, 045015 (2022).
[33] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, PRX Quantum **3**, 010313 (2022).
[34] Z. Holmes, A. Arrasmith, B. Yan, P. J. Coles, A. Albrecht, and A. T. Sornborger, Barren plateaus preclude learning scramblers, Phys. Rev. Lett. **126**, 190501 (2021).
[35] J. Kim, J. Kim, and D. Rosa, Universal effectiveness of high-depth circuits in variational eigenproblems, Phys. Rev. Res. **3**, 023203 (2021).
[36] X. Liu, G. Liu, J. Huang, H.-K. Zhang, and X. Wang, Mitigating barren plateaus of variational quantum eigensolvers, arXiv:2205.13539.
[37] P. Huembeli and A. Dauphin, Characterizing the loss landscape of variational quantum circuits, Quantum Sci. Technol. **6**, 025011 (2021).
[38] H.-K. Zhang, C. Zhu, G. Liu, and X. Wang, Fundamental limitations on optimization in variational quantum algorithms, arXiv:2205.05056.
[39] H. R. Grimsley, N. J. Mayhall, G. S. Barron, E. Barnes, and S. E. Economou, Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus, npj Quantum Inf. **9**, 19 (2023).
[40] Z. Liu, L.-W. Yu, L.-M. Duan, and D.-L. Deng, Presence and absence of barren plateaus in tensor-network based machine learning, Phys. Rev. Lett. **129**, 270501 (2022).
[41] C. Zhao and X.-S. Gao, Analyzing the barren plateau phenomenon in training quantum neural networks with the ZX-calculus, Quantum **5**, 466 (2021).
[42] E. Cervero Martín, K. Plekhanov, and M. Lubasch, Barren plateaus in quantum tensor network optimization, Quantum **7**, 974 (2023).
[43] T. Barthel and Q. Miao, Absence of barren plateaus and scaling of gradients in the energy optimization of isometric tensor network states, arXiv:2304.00161.
[44] Q. Miao and T. Barthel, Isometric tensor network optimization for extensive Hamiltonians is free of barren plateaus, arXiv:2304.14320.
[45] Q. Miao and T. Barthel, Quantum-classical eigensolver using multiscale entanglement renormalization, Phys. Rev. Res. **5**, 033141 (2023).
[46] Q. Miao and T. Barthel, Convergence and quantum advantage of trotterized MERA for strongly-correlated systems, arXiv:2303.08910.
[47] Z. Liu, Q. Ye, L.-W. Yu, L. M. Duan, and D.-L. Deng, Theory on variational high-dimensional tensor networks, arXiv:2303.17452.
[48] R. J. Garcia, C. Zhao, K. Bu, and A. Jaffe, Barren plateaus from learning scramblers with local cost functions, J. High Energy Phys. 01 (2023) 090.
[49] S. Liu, S.-X. Zhang, S.-K. Jian, and H. Yao, Training variational quantum algorithms with random gate activation, Phys. Rev. Res. **5**, L032040 (2023).
[50] A. Kitaev, Fault-tolerant quantum computation by anyons, Ann. Phys. (Amsterdam) **303**, 2 (2003).
[51] A. Kitaev, Anyons in an exactly solved model and beyond, Ann. Phys. (Amsterdam) **321**, 2 (2006).
[52] X.-G. Wen, Quantum orders in an exact soluble model, Phys. Rev. Lett. **90**, 016803 (2003).
[53] A. Kitaev and J. Preskill, Topological entanglement entropy, Phys. Rev. Lett. **96**, 110404 (2006).
[54] L. Gioia and C. Wang, Nonzero momentum requires long-range entanglement, Phys. Rev. X **12**, 031007 (2022).
[55] S. Bravyi, M. B. Hastings, and F. Verstraete, Lieb-Robinson bounds and the generation of correlations and topological quantum order, Phys. Rev. Lett. **97**, 050401 (2006).
[56] K. J. Satzinger *et al.*, Realizing topologically ordered states on a quantum processor, Science **374**, 1237 (2021).
[57] X. Chen, A. Dua, M. Hermele, D. T. Stephen, N. Tantivasadakarn, R. Vanhove, and J.-Y. Zhao, Sequential quantum circuits as maps between gapped phases, Phys. Rev. B **109**, 075116 (2024).
[58] R.-Y. Sun, T. Shirakawa, and S. Yunoki, Parametrized quantum circuit for weight-adjustable quantum loop gas, Phys. Rev. B **107**, L041109 (2023).
[59] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf, Sequential generation of entangled multiqubit states, Phys. Rev. Lett. **95**, 110503 (2005).
[60] C. Schön, K. Hammerer, M. M. Wolf, J. I. Cirac, and E. Solano, Sequential generation of matrix-product states in cavity QED, Phys. Rev. A **75**, 032311 (2007).
[61] L. Slattery and B. K. Clark, Quantum circuits for two-dimensional isometric tensor networks, arXiv:2108.02792.
[62] Z.-Y. Wei, D. Malz, and J. I. Cirac, Sequential generation of projected entangled-pair states, Phys. Rev. Lett. **128**, 010607 (2022).
[63] Y. Zhang, S. Jahanbani, A. Riswadkar, S. Shankar, and A. C. Potter, Sequential quantum simulation of spin chains with a single circuit QED device, Phys. Rev. A **109**, 022606 (2024).
[64] Y.-j. Liu, K. Shtengel, A. Smith, and F. Pollmann, Methods for simulating string-net states and anyons on a digital quantum computer, PRX Quantum **3**, 040315 (2022).
[65] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.132.150603 for preliminaries on unitary designs and the Weingarten calculus, basic setup, theoretical derivations including rigorous definitions and complete proofs, and additional numerical results with technical details, which includes Refs. [66–100].
[66] B. Collins and P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic group, Commun. Math. Phys. **264**, 773 (2006).
[67] M. B. Hastings, An area law for one-dimensional quantum systems, J. Stat. Mech. (2007) P08024.
[68] J. Haferkamp, D. Hangleiter, J. Eisert, and M. Gluza, Contracting projected entangled pair states is average-case hard, Phys. Rev. Res. **2**, 013010 (2020).

[69] S. Bravyi, D. Gosset, and R. Movassagh, Classical algorithms for quantum mean values, Nat. Phys. **17**, 337 (2021).
[70] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, Stripe order in the underdoped region of the two-dimensional Hubbard model, Science **358**, 1155 (2017).
[71] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, and S. Zhang, Absence of superconductivity in the pure two-dimensional Hubbard model, Phys. Rev. X **10**, 031016 (2020).
[72] E. W. Huang, C. B. Mendl, S. Liu, S. Johnston, H.-C. Jiang, B. Moritz, and T. P. Devereaux, Numerical evidence of fluctuating stripes in the normal state of high-T_c cuprate superconductors, Science **358**, 1161 (2017).
[73] G. Vidal, Class of quantum many-body states that can be efficiently simulated, Phys. Rev. Lett. **101**, 110501 (2008).
[74] D. Malz, G. Styliaris, Z.-Y. Wei, and J. I. Cirac, Preparation of matrix product states with log-depth quantum circuits, Phys. Rev. Lett. **132**, 040404 (2024).
[75] Z. Landau, U. Vazirani, and T. Vidick, A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians, Nat. Phys. **11**, 566 (2015).
[76] X. Chen, Z.-C. Gu, and X.-G. Wen, Complete classification of one-dimensional gapped quantum phases in interacting spin systems, Phys. Rev. B **84**, 235128 (2011).
[77] X. Chen, Z.-C. Gu, and X.-G. Wen, Classification of gapped symmetric phases in one-dimensional spin systems, Phys. Rev. B **83**, 035107 (2011).
[78] N. Schuch, D. Pérez-García, and I. Cirac, Classifying quantum phases using matrix product states and projected entangled pair states, Phys. Rev. B **84**, 165139 (2011).
[79] B. Zeng, X. Chen, D.-L. Zhou, and X.-G. Wen, *Quantum Information Meets Quantum Matter*, Quantum Science and Technology (Springer, New York, NY, 2019).
[80] J. Gray and S. Kourtis, Hyper-optimized tensor network contraction, Quantum **5**, 410 (2021).
[81] A. Anshu, I. Arad, and A. Jain, How local is the information in tensor networks of matrix product states or projected entangled pairs states, Phys. Rev. B **94**, 195143 (2016).
[82] M. Schwarz, O. Buerschaper, and J. Eisert, Approximating local observables on projected entangled pair states, Phys. Rev. A **95**, 060102(R) (2017).
[83] S.-J. Ran, E. Tirrito, C. Peng, X. Chen, L. Tagliacozzo, G. Su, and M. Lewenstein, *Tensor Network Contractions*, Lecture Notes in Physics Vol. 964 (Springer International Publishing, Cham, 2020).
[84] R. Orús, Tensor networks for complex quantum systems, Nat. Rev. Phys. **1**, 538 (2019).
[85] D. Bluvstein, H. Levine, G. Semeghini, T. T. Wang, S. Ebadi, M. Kalinowski, A. Keesling, N. Maskara, H. Pichler, M. Greiner, V. Vuletić, and M. D. Lukin, A quantum processor based on coherent transport of entangled atom arrays, Nature (London) **604**, 451 (2022).
[86] D. Bluvstein *et al.*, Logical quantum processor based on reconfigurable atom arrays, Nature (London) **626**, 58 (2023).
[87] S.-k. Chu, G. Zhu, and A. V. Gorshkov, Entanglement renormalization circuits for chiral topological order, arXiv:2304.13748.
[88] A. Anshu, N. P. Breuckmann, and C. Nirkhe, NLTS Hamiltonians from good quantum codes, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, USA, 2023), pp. 1090–1096.
[89] S. Bravyi, A. W. Cross, J. M. Gambetta, D. Maslov, P. Rall, and T. J. Yoder, High-threshold and low-overhead fault-tolerant quantum memory, arXiv:2308.07915.
[90] Q. Xu, J. P. B. Ataides, C. A. Pattison, N. Raveendran, D. Bluvstein, J. Wurtz, B. Vasic, M. D. Lukin, L. Jiang, and H. Zhou, Constant-overhead fault-tolerant quantum computation with reconfigurable atom arrays, arXiv:2308.08648.
[91] F. Pan, P. Zhou, S. Li, and P. Zhang, Contracting Arbitrary tensor networks: General approximate algorithm and applications in graphical models and quantum circuit simulations, Phys. Rev. Lett. **125**, 060503 (2020).
[92] F. Pan, K. Chen, and P. Zhang, Solving the sampling problem of the sycamore quantum circuits, Phys. Rev. Lett. **129**, 090502 (2022).
[93] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Time-evolution methods for matrix-product states, Ann. Phys. (Amsterdam) **411**, 167998 (2019).
[94] R. Raussendorf, S. Bravyi, and J. Harrington, Long-range quantum entanglement in noisy cluster states, Phys. Rev. A **71**, 062313 (2005).
[95] L. Piroli, G. Styliaris, and J. I. Cirac, Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter, Phys. Rev. Lett. **127**, 220503 (2021).
[96] R. Verresen, N. Tantivasadakarn, and A. Vishwanath, Efficiently preparing Schrödinger's cat, fractons and non-Abelian topological order in quantum devices, arXiv:2112.03061.
[97] N. Tantivasadakarn, R. Thorngren, A. Vishwanath, and R. Verresen, Long-range entanglement from measuring symmetry-protected topological phases, arXiv:2112.01519.
[98] S. Bravyi, I. Kim, A. Kliesch, and R. Koenig, Adaptive constant-depth circuits for manipulating non-Abelian anyons, arXiv:2205.01933.
[99] N. Tantivasadakarn, R. Verresen, and A. Vishwanath, Shortest route to non-Abelian topological order on a quantum processor, Phys. Rev. Lett. **131**, 060405 (2023).
[100] N. Tantivasadakarn, A. Vishwanath, and R. Verresen, Hierarchy of topological order from finite-depth unitaries, measurement, and feedforward, PRX Quantum **4**, 020339 (2023).
[101] Y. Sekino and L. Susskind, Fast scramblers, J. High Energy Phys. 10 (2008) 065.
[102] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Towards the fast scrambling conjecture, J. High Energy Phys. 04 (2013) 022.
[103] X. Mi *et al.*, Information scrambling in quantum circuits, Science **374**, 1479 (2021).
[104] M. C. Bañuls, D. Pérez-García, M. M. Wolf, F. Verstraete, and J. I. Cirac, Sequentially generated states for the study of two-dimensional systems, Phys. Rev. A **77**, 052306 (2008).

[105] S.-J. Ran, Encoding of matrix product states into quantum circuits of one- and two-qubit gates, Phys. Rev. A **101**, 032310 (2020).
[106] P. F. Zhou, R. Hong, and S. J. Ran, Automatically differentiable quantum circuit for many-qubit state preparation, Phys. Rev. A **104**, 042601 (2021).
[107] R. Haghshenas, J. Gray, A. C. Potter, and G. K.-L. Chan, Variational power of quantum circuit tensor networks, Phys. Rev. X **12**, 011047 (2022).
[108] M. P. Zaletel and F. Pollmann, Isometric tensor network states in two dimensions, Phys. Rev. Lett. **124**, 037201 (2020).
[109] T. Soejima, K. Siva, N. Bultinck, S. Chatterjee, F. Pollmann, and M. P. Zaletel, Isometric tensor network representation of string-net liquids, Phys. Rev. B **101**, 085117 (2020).
[110] L. Haller, W.-t. Xu, Y.-j. Liu, and F. Pollmann, Quantum phase transition between symmetry enriched topological phases in tensor-network states, Phys. Rev. Res. **5**, 043078 (2023).
[111] Y.-j. Liu, K. Shtengel, and F. Pollmann, Topological quantum phase transitions in 2D isometric tensor networks, arXiv:2312.05079.
[112] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing, arXiv:2312.09121.
[113] I. L. Markov and Y. Shi, Simulating quantum computation by contracting tensor networks, SIAM J. Comput. **38**, 963 (2008).
[114] D. Malz and R. Trivedi, Computational complexity of isometric tensor network states, arXiv:2402.07975.
[115] D. Azses, R. Haenel, Y. Naveh, R. Raussendorf, E. Sela, and E. G. Dalla Torre, Identification of symmetry-protected topological states on noisy quantum computers, Phys. Rev. Lett. **125**, 120502 (2020).
[116] S.-X. Zhang, J. Allcock, Z.-Q. Wan, S. Liu, J. Sun, H. Yu, X.-H. Yang, J. Qiu, Z. Ye, Y.-Q. Chen, C.-K. Lee, Y.-C. Zheng, S.-K. Jian, H. Yao, C.-Y. Hsieh, and S. Zhang, TensorCircuit: A quantum software framework for the NISQ Era, Quantum **7**, 912 (2023).
[117] S. Trebst, P. Werner, M. Troyer, K. Shtengel, and C. Nayak, Breakdown of a topological phase: Quantum phase transition in a loop gas model with tension, Phys. Rev. Lett. **98**, 070602 (2007).
[118] F. Wu, Y. Deng, and N. Prokof'ev, Phase diagram of the toric code model in a parallel magnetic field, Phys. Rev. B **85**, 195104 (2012).
[119] S.-X. Zhang, C.-Y. Hsieh, S. Zhang, and H. Yao, Differentiable quantum architecture search, Quantum Sci. Technol. **7**, 045023 (2022).
[120] Y. Du, T. Huang, S. You, M.-H. Hsieh, and D. Tao, Quantum circuit architecture search for variational quantum algorithms, npj Quantum Inf. **8**, 62 (2022).
[121] Z. Lu, P.-X. Shen, and D.-L. Deng, Markovian quantum neuroevolution for machine learning, Phys. Rev. Appl. **16**, 044039 (2021).
[122] S.-X. Zhang, C.-Y. Hsieh, S. Zhang, and H. Yao, Neural predictor based quantum architecture search, Mach. Learn. **2**, 045027 (2021).
[123] L. Bittel and M. Kliesch, Training variational quantum algorithms is NP-hard, Phys. Rev. Lett. **127**, 120502 (2021).
[124] H.-K. Zhang, C. Zhu, M. Jing, and X. Wang, Statistical analysis of quantum state learning process in quantum neural networks, arXiv:2309.14980.
[125] E. R. Anschuetz and B. T. Kiani, Quantum variational algorithms are swamped with traps, Nat. Commun. **13** (2022).