



Universiteit
Leiden
The Netherlands

Quantum-classical separations in shallow-circuit-based learning with and without noises

Zhang, Z.; Gong, W.; Li, W.K.; Deng, D.L.

Citation

Zhang, Z., Gong, W., Li, W. K., & Deng, D. L. (2024). Quantum-classical separations in shallow-circuit-based learning with and without noises. *Communications Physics*, 7. doi:10.1038/s42005-024-01783-7

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/4171998>

Note: To cite this publication please use the final published version (if applicable).

<https://doi.org/10.1038/s42005-024-01783-7>

Quantum-classical separations in shallow-circuit-based learning with and without noises

Check for updates

Zhihan Zhang^{1,2,6}, Weiyuan Gong^{1,2,6}, Weikang Li^{1,3} & Dong-Ling Deng^{1,4,5}✉

An essential problem in quantum machine learning is to find quantum-classical separations between learning models. However, rigorous and unconditional separations are lacking for supervised learning. Here we construct a classification problem defined by a noiseless constant depth (i.e., shallow) quantum circuit and rigorously prove that any classical neural network with bounded connectivity requires logarithmic depth to output correctly with a larger-than-exponentially-small probability. This unconditional near-optimal quantum-classical representation power separation originates from the quantum nonlocality property that distinguishes quantum circuits from their classical counterparts. We further characterize the noise regimes for demonstrating such a separation on near-term quantum devices under the depolarization noise model. In addition, for quantum devices with constant noise strength, we prove that no super-polynomial classical-quantum separation exists for any classification task defined by Clifford circuits, independent of the structures of the circuits that specify the learning models.

Quantum machine learning studies the interplay between machine learning and quantum physics^{1–6}. In recent years, a number of quantum learning algorithms have been proposed^{7–22}, which may offer potential quantum advantages over their classical counterparts. However, most of these algorithms either depend on a complexity assumption for the rigorous proof of the quantum advantage¹⁵, or are out of the limited capabilities for noisy near-term quantum devices^{7–10}. On the other hand, the difficulty of constructing fault-tolerant and deep quantum circuits motivates the study of machine learning using constant-depth (shallow) quantum circuits. Most of these algorithms assume classical access to data and use variational circuits for learning. Examples along this direction include machine learning based on variational quantum algorithms^{23–31}, which train a parameterized shallow quantum circuit with classical optimizers (see Supplementary Note 1 and Supplementary Fig. 1 for a brief introduction). For these experimental-friendly algorithms, however, there is no rigorous proof showing that they have genuine advantages over classical algorithms. These two points raise the essential problem of whether we can rigorously prove an unconditional quantum advantage in machine learning feasible for near-term quantum devices³². Here, we establish a rigorous quantum-classical representation power separation for the supervised learning model with noiseless shallow quantum circuits (as shown in Fig. 1a) and pin down the noise regimes

where the quantum-classical separation exists or disappears for shallow-circuit-based learning.

Recently, a rigorous and robust quantum advantage in supervised learning has been proved¹⁵, which relies on the assumption of the computational hardness in solving the discrete logarithm problem. Despite this exciting progress, an unconditional quantum-classical separation in supervised learning is still lacking. In the quantum computing literature, a different paradigm to rigorously demonstrate quantum advantage was proposed³³ to focus on shallow circuits, which is more feasible with noisy devices. It compared the circuit depth required for quantum and classical circuits to solve a relation problem, obtaining a separation originating from the classical hardness of simulating the intrinsic nonlocality property of quantum mechanics. In particular, it is proved that a shallow quantum circuit can solve a relation problem such that any classical circuits with bounded fan-in gates, namely gates with a constant number of input bits, solving this problem with a moderate probability requires a logarithmic depth. Later works extended and amplified this result into an exponential separation in success probability^{34–37}. Other studies further showed that it is possible to preserve the separation from noises^{38,39}. Very recently, a separation with similar intuition was proposed for sampling problems with shallow circuits⁴⁰. For classical neural networks, many convolution layers,

¹Center for Quantum Information, IIIS, Tsinghua University, Beijing, 100084, China. ²School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138, USA. ³Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands. ⁴Shanghai Qi Zhi Institute, 41st Floor, AI Tower, No. 701 Yunjin Road, Xuhui District, Shanghai, 200232, China. ⁵Hefei National Laboratory, Hefei, 230088, China. ⁶These authors contributed equally: Zhihan Zhang, Weiyuan Gong. ✉ e-mail: dldeng@tsinghua.edu.cn

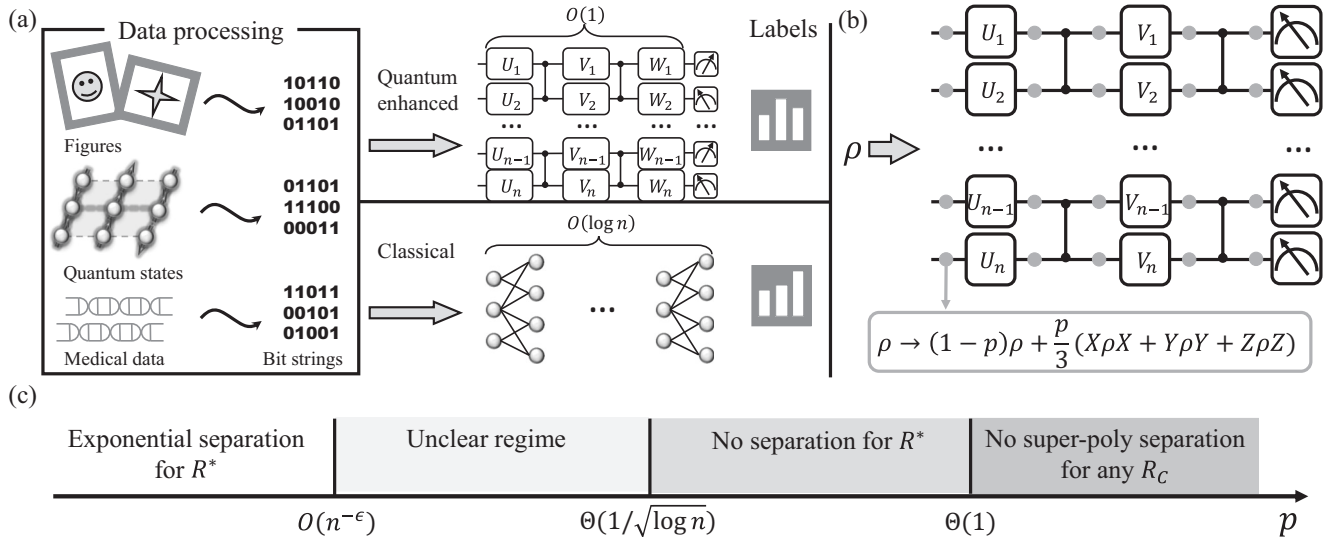


Fig. 1 | Illustration of our settings and results. **a** An illustration of classical and quantum supervised learning models using shallow quantum circuits and neural networks with bounded connectivity. The data samples are first encoded into bit strings \mathbf{x} . These bit strings are input into a variational quantum circuit in the quantum-enhanced scenario and a classical neural network in the classical scenario. Each neuron in the classical neural network has bounded connectivity to the neurons in the previous layer. The outputs for two classifiers are distributions on all possible labels encoded in bit strings \mathbf{y} . **b** An implementation of a noisy quantum circuit, where an identical amount p of depolarizing noise is added to each qubit at each step.

c The characterization of noise rate regimes for the existence and absence of quantum-classical separation in learning classification tasks defined by shallow quantum circuits. Here, classification tasks R^* and R_c are relations defined in Theorem 1 and Theorem 2, respectively. The classification task R^* can provide a super-polynomial separation in the prediction accuracy when the noise rate is bounded above by $O(n^{-\epsilon})$, while such separation will vanish in the noise regime of $p > \Omega(1/\sqrt{\log n})$. In addition, any classification defined by a shallow Clifford circuit possesses no super-polynomial quantum-classical separation when the quantum circuit is implemented with a constant noise rate.

ReLU layers, and pooling layers satisfy the local connectivity assumptions (i.e., bounded connectivity). Therefore, these unconditional separations provide the intuition that the nonlocality of quantum mechanics can provide an unconditional quantum advantage in machine learning tasks as well.

In this paper, we investigate the separations between shallow variational quantum circuits and classical neural networks for supervised learning tasks. For noiseless quantum devices, we prove that there exists a classification task defined by the input bits and output distributions of a shallow quantum circuit such that any classical neural network with bounded connectivity requires logarithmic depth to output correctly even with exponentially small probability $\exp(-O(n^{1-\epsilon}))$, where ϵ is a positive constant and n is the system size. This indicates an unconditional near-optimal separation between the representation power of shallow-circuit-based quantum classifiers and shallow-network-based classical neural networks. This quantum-classical separation originates from the classical hardness of simulating Bell nonlocality⁴¹. In addition, we consider a scenario with noises as well (Fig. 1b, c). We show that when the noise rate is bounded above by $O(n^{-\epsilon})$, the quantum-classical separation will persist. On the contrary, if the noise rate exceeds $\Omega(1/\sqrt{\log n})$, this separation will no longer exist. In addition, we prove that any classification task defined by a noiseless shallow Clifford circuit provides no super-polynomial quantum-classical separation if implemented on quantum devices with a constant noise rate. We will outline our results and leave details to Supplementary Methods.

Results

Generalized classification task

Consider a generalized classification task in the context of supervised learning, where we assign multiple labels $\mathbf{y}_1, \dots, \mathbf{y}_K \in \mathcal{O}$ to an input data sample $\mathbf{x} \in \mathcal{I}$, with \mathcal{I} and \mathcal{O} denoting the set of all possible samples and labels. Without loss of generality, we encode the samples and labels into bit strings of n and m bits, respectively. In essence, such a classification task can be described by a relation $R: \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}$, such that a label \mathbf{y} is a

legitimate label for an input sample \mathbf{x} if and only if $R(\mathbf{x}, \mathbf{y}) = 1$. For a supervised learning algorithm, we can also describe the inputs and outputs through a relation (called a hypothesis relation) $R_h: \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}$. For an input sample \mathbf{x} , we assign $R_h(\mathbf{x}, \mathbf{y}) = 1$ if \mathbf{y} is a possible output for the learning algorithm and $R_h(\mathbf{x}, \mathbf{y}) = 0$ otherwise. Given an input \mathbf{x} , the learning algorithm outputs a label $\mathbf{y} \in \{0, 1\}^m$ with probability $\Pr[\mathbf{y}|\mathbf{x}, R_h]$. To evaluate the performance of a learning algorithm, we use a loss function defined as the probability of the learning algorithm outputting a wrong label averaged uniformly over all possible input strings $\mathbf{x} \in \{0, 1\}^n$:

$$\mathcal{L}_p(R, R_h) = 1 - \frac{1}{2^n} \sum_{\mathbf{x}, \mathbf{y}} \Pr[\mathbf{y}|\mathbf{x}, R_h] \mathbb{1}_{R(\mathbf{x}, \mathbf{y})=1}, \quad (1)$$

where $\mathbb{1}_A$ takes value 1 when the argument A is satisfied and 0 otherwise. We say a relation can be represented by a neural network or a quantum circuit if the loss function can be reduced to zero by choosing a set of suitable variational parameters.

We note that in our setting m can be of the same scale with n , which enables an exponential number of legitimate labels for each input. This generalized classification model fits naturally with the recently flourishing field of computer vision and natural language processing, particularly image segmentation^{42,43} and sequence-to-sequence learning tasks⁴⁴ such as text translation⁴⁵. Take image segmentation for example: the goal is to assign labels to every pixel in an image, which can result in potentially exponentially many valid outputs, especially when certain pixels contain some ambiguity and are allowed to be assigned to even exponentially many labels. Similarly, in translation tasks, there may exist numerous grammatically correct translations, all of which could be considered valid.

Separations for noiseless quantum circuits

We compare the representation power of shallow quantum circuits and shallow classical neural networks equipped with neurons of bounded connectivity. We start with variational quantum circuits without noise. We

focus on the qubit encoding scheme throughout this paper for the relation $R_h: \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}$. Therefore, the quantum circuit takes n -qubit quantum states as input and output m -bit labels after measurements. In particular, we prove the following theorem, which shows a near-optimal quantum-classical separation between these two types of classifiers.

Theorem 1. *There exists a classification task described by a relation $R^*: \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}$ such that a constant-depth parametrized variational quantum circuit with n -qubit input using single- and two-qubit gates can represent this relation (i.e., with zero loss). However, any classical neural network with neurons of bounded connectivity that can represent a hypothesis relation R_h with*

$$\mathcal{L}_p(R^*, R_h) \leq 1 - \exp(-\gamma n^{1-\epsilon}) \quad (2)$$

for some constant $\gamma > 0$ requires depth at least $\Omega(\epsilon \log n)$ for any positive constant $\epsilon < 1$.

Proof Sketch. We provide the main idea here. The detailed proof is technically involved and thus left to the Supplementary Note 2. Specifically, we prove the equivalence of the classical circuit model and neural network model with bounded connectivity, and a proposition indicating that there exists an elementary classification task $R_e: \{0, 1\}^{n_e} \times \{0, 1\}^{m_e} \rightarrow \{0, 1\}$ such that there exists a constant-depth quantum circuit that outputs a correct label y for any given sample \mathbf{x} with unity probability. Whereas, any (randomized) classical neural network with bounded connectivity outputs correctly with probability at least $1 - \alpha$ for some constant α requires depth $\Omega(\log n_e)$. We then exploit a direct product of t copies of the relation R_e . We show that, by choosing large enough t , one can find a relation $R^* = R_e^{\times t}$ with $n = n_e t$ and $m = m_e t$ to amplify the quantum-classical loss function separation to 0 versus $1 - \exp(-O(n^{1-\epsilon}))$ for any positive constant $\epsilon < 1$. This completes the proof for Theorem 1. \square

We remark that the improved soundness amplification technique, which considers a direct product of t copies of the relation R_e , deviates from the argument by Le Gall³⁴ and enables us to push the unconditional separation to near optimal. In particular, while the unconditional separation proved there between classical and quantum circuit is $1 - \exp(-\gamma n^{1/2})$ versus 0, we extend this separation to $1 - \exp(-\gamma n^{1-\epsilon})$ versus 0 by allowing arbitrary $0 < \epsilon < 1$. To obtain this result, we push the soundness amplification technique developed by Le Gall³⁴ to its limit by carefully tuning the parameters and refining the procedure.

Theorem 1 can be attributed to the difficulty of simulating quantum nonlocality using classical algorithms⁴⁶. More concretely, we consider supervised learning the measurement outcomes when measuring each qubit of a well-designed graph state⁴⁷ on the X -basis or the Y -basis. The graph states contain long cycles, requiring entanglement between distant sites in classical simulation. As a result, any classical neural network with bounded connectivity requires logarithmic depth to connect these distant wires. Similar ideas have been exploited to demonstrate the computational power separation between shallow classical and quantum circuits^{33–36,38}.

Theorem 1 shows an exponential quantum-classical separation in the success probability of outputting a correct label for the classification task. We note that here the exponential separation is in success probability exclusively, which is different from the typical usage as in³⁵, where the exponential separation refers to both the circuit size and the success probability. This result indicates a quantum-classical separation in the representation power between classical neural networks with bounded connectivity and variational quantum circuits in supervised learning. Shallow quantum circuits are more powerful in representing complicated relations than classical shallow neural networks. Our results indicate the existence of some relation that can be represented by a constant-depth quantum circuit but not by any constant-depth neural networks with bounded connectivity however one tunes the parameters in a rigorous fashion. Whether one can efficiently optimize the parameters in the variational quantum circuit to find this function is an independent and

challenging question that warrants further investigation. In addition, this separation also implies a constant-versus-logarithmic time separation in the inference stage of the classification task, between quantum classifiers based on constant-depth variational circuits and classical classifiers based on neural networks with bounded connectivity.

Furthermore, we rigorously prove that there exists a classical neural network of depth $c_0 \log n$ with bounded connectivity and fine-tuned parameters that can reduce the loss function to $1 - \exp(-O(n/\sqrt{\log n}))$ for any relation defined by a two-dimensional shallow quantum circuit (see Supplementary Note 3). Therefore, the quantum-classical separation shown in Theorem 1 is near-optimal. In particular, to prove the optimality we explicitly propose an algorithm. This algorithm divides the input qubits into blocks of $O(\sqrt{\log n}) \times O(\sqrt{\log n})$ qubits. We show that for shallow quantum circuits with single- and two-qubit gates, a classical simulation can output correctly if there are no inter-block gates. We then prove that the number of qubits in the input string affected by inter-block gates is bounded above by $O(n/\sqrt{\log n})$. Therefore, this classical simulation can output a correct string for all the qubits except the affected qubits. Finally, we apply a random guessing strategy on these affected qubits, and the probability of outputting a correct string is thus $\exp(-O(n/\sqrt{\log n}))$.

Theorem 1 can be extended straightforwardly to loss functions other than that defined in Eq. (1). For example, we consider the loss function defined on R^* and R_h based on the widely used Kullback-Leibler (KL) divergence⁴⁸. In Supplementary Note 3, we show that relation R^* in Theorem 1 outputs according to an equal distribution on all possible strings \mathbf{y} . If we compute the KL divergence between R^* and the probability distribution output from a learning algorithm, the loss function for the variational quantum circuit can be reduced to zero. However, any classical neural network with bounded connectivity and loss function smaller than $O(n^{1-\epsilon})$ requires depth $\Omega(\epsilon \log n)$.

Noisy quantum circuits

Short-term quantum devices suffer from noises³². As a result, a question of both theoretical and experimental interest is whether the quantum-classical separation discussed above persists for noisy quantum circuits. It is shown that the existence of noise leads to the ineffectiveness of deep quantum circuits caused by the barren plateau phenomenon^{49–52}. However, the impact of noises on shallow circuits in the average case remains largely unexplored. To address this important question, we focus on the depolarization noise model⁵³ $\rho \rightarrow (1 - p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z)$ for the shallow quantum circuit as an example. We prove that the variational quantum circuits in Theorem 1 suffering from noise rate p can only output correctly with probability $\exp(-\Theta(pn))$. Therefore, we conclude that the quantum-classical separation in Theorem 1 remains when $p = O(n^{-\epsilon})$. Although both the classical and quantum success probabilities are exponentially decaying within this regime, these two probabilities decay with different rates of $\exp(-n^\epsilon)$ and $\exp(-pn)$, which yields a super-polynomial separation. However, the noise rate on most near-term quantum devices without any error mitigation or correction is typically assumed to be $p = O(1)$ ^{32,54–56}. In such a scenario, we rigorously prove that the separation in Theorem 1 vanishes. In the above discussion, we have proposed a classical algorithm (see Supplementary Fig. 2 for illustration) using logarithmic depth to output a correct label of a shallow quantum circuit with probability at least $\exp(-O(n/\sqrt{\log n}))$ when proving the near-optimality of the noiseless quantum-classical separation. By comparing this result and the performance bound for noisy shallow quantum circuits, the largest noise threshold for this quantum-classical separation will not exceed $p = O(1/\sqrt{\log n})$, which decreases as the system size increases. An illustrative figure for the noise regimes where the quantum-classical separation exists or disappears for shallow-circuit-based learning is provided as Fig. 1(b). We leave the details for the derivation of the upper and lower bound on the noise threshold for demonstrating Theorem 1 on noisy devices in Supplementary Note 3.

We further find that the nonexistence of quantum-classical separation in the context of supervised learning using shallow circuits can be extended

to more general cases. In particular, we derive the following theorem showing that there is no super-polynomial quantum-classical separation between noisy shallow variational quantum circuits and classical learning algorithms using shallow neural networks or circuits.

Theorem 2. For any classification task corresponding to a relation $R_C: \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}$ that can be defined by a (classically-controlled) Clifford circuit of any depth, no quantum circuit undergoing the depolarization noise $\rho \rightarrow (1 - p)\rho + p(X\rho X + Y\rho Y + Z\rho Z)/3$ with constant noise rate p can achieve a loss less than

$$1 - (1 - \mathcal{L}_p(R_C, R_r))^{\log_2 \frac{1}{1-3p}}, \quad (3)$$

where R_r is the relation represented by a classical neural network that outputs uniformly randomly over $y \in \{0, 1\}^m$.

Proof Sketch. We sketch the major steps here and leave the technical details to Supplementary Note 4. We adapt a well known lemma^{57,58} and prove that the output of the quantum circuit defining R_C uniformly distributes over an affine subspace of \mathbb{Z}_2^m . Therefore, there exists a subset y_S of output bits such that when the input bits and the output bits outside of y_S are fixed, only one assignment of the bits in y_S can satisfy R_C , meaning that the probability of success for a classical random guessing algorithm is $(\frac{1}{2})^{|y_S|}$. Similarly, for noisy quantum circuits, we consider the noise in the layer exactly before measuring the bits in y_S conditioned on all other noises. We conclude that the output label for a noisy quantum circuit for R_C is correct with probability at most $(1 - \frac{2}{3}p)^{|y_S|}$. This completes the proof for this theorem. \square

We recall that the shallow variational quantum circuit in Theorem 1 contains non-Clifford gates. However, it can be proved that this classification task can also be represented by a classically-controlled Clifford quantum circuit³⁴. Thus, the result in Theorem 2 also works for the classification task in Theorem 1. As p is a constant, $\log_2(1/(1 - 2p/3))$ is a constant. Therefore, Theorem 2 shows that the success probability for a noisy shallow quantum circuit to output a correct label for any classification task defined by a noiseless shallow Clifford circuit scale at most polynomially as that of a classical algorithm that randomly guesses the labels. Since random guesses can be accomplished by a classical neural network of constant depth, the separation in Theorem 1 does not exist here. Theorem 2 further indicates that we cannot expect a near-term experimental demonstration of quantum advantage only using classification tasks characterized by a Clifford circuit.

We note that Theorem 2 is a strong no-go theorem that applies to any relations defined by Clifford circuits of arbitrary depth. The noise model that applies also holds for not only the standard independent and identically distributed depolarization noise at each layer (as shown in Fig. 1(b)), but also holds for even if the depolarizing noise happens only at the last layer (see Supplementary Note 4 for details). It is worthwhile to remark that the relation considered in Theorem 2 can be defined by a noiseless Clifford quantum circuit, which is different from the relations considered in works by Bravyi et al. and Grier et al.^{38,39}. There, an implicit error correction technique is introduced to guarantee that the quantum-classical separation persists for noisy circuits. Such a technique takes a problem with a separation between classical circuits and noiseless quantum circuits and encodes the problem in error-correcting codes. Then the error syndromes are measured without carrying out the error correction and the classical circuit is required to do the same. This yields a separation between shallow classical circuits and noisy quantum circuits of constant noise rate. Therefore, the resulting relation can be represented but not defined by a shallow quantum circuit as the quantum circuit is allowed to make a mistake with a moderate probability. Whereas, in our scenario, the output distribution of the relation is assumed to be exactly the same with some quantum circuits, yielding a more restricted class of relations. We conclude that this kind of restriction is vital for the nonexistence of the separation in works by Bravyi et al. and Grier et al.^{38,39}.

We also clarify the differences between Theorem 2 and the recent results by Aharonov et al.⁵⁵ in the absence of quantum advantage on constant-rate noisy random circuit sampling. Aharonov et al.⁵⁵ considered the task aiming at approximating the output distribution of a random quantum circuit within a small total variational difference. It is widely believed that classically simulating the output distributions of a (noisy) random circuit of logarithmic depth exactly is difficult^{17,59-64}, but Aharonov et al.⁵⁵ explicitly propose an efficient algorithm to approximate the outputs distribution of noisy random circuits with polynomial complexity. While this result gives evidence against the quantum advantage of random circuit sampling with a constant noise rate, Theorem 2 focus on the classification task represented by quantum circuits of depth $O(1)$. The random circuits are assumed to satisfy the anti-concentration condition, which requires a circuit depth of $\Omega(\log n)$, while we consider shallow circuits of constant depth. It is hard to classically simulate the result of random circuits, but the output distribution of the shallow quantum circuit considered in this paper can be efficiently simulated by classical algorithms^{65,66}. To construct a quantum-classical separation in this regime, we compare the representation power of shallow quantum circuits to that of shallow classical neural networks and show the quantum-classical separation. We prove that classification tasks defined by shallow Clifford circuits can not demonstrate any super-polynomial quantum advantage on near-term devices with constant noise rates.

Discussion

Many questions remain and are worth further investigation. For instance, our discussion in this paper focused on the relation exactly defined by a shallow quantum circuit. For quantum devices with a constant noisy rate, it is still possible to find a quantum-classical separation in the success probability for relations not exactly defined by a quantum circuit. How to explicitly construct such a hard classification task to rigorously obtain a noise-tolerant quantum-classical separation without relying on quantum error correction is crucial for demonstrating quantum advantages with near-term quantum devices. In practice, many classical neural networks bear unbounded connectivity. How to extend our results to this scenario demands further exploration. One may also explore the potential representation power trade-off between width and depth for classical neural networks with unbounded connectivity. In addition, how to extend our results to unsupervised learning and reinforcement learning scenarios remains unclear. Whether we can transfer our separation in representation power to a separation in sample complexity or computational complexity is an important future direction as well. Finally, it would be of both fundamental and practical importance to carry out an experiment to demonstrate the unconditional quantum-classical separation proved in Theorem 1. This would be a far-reaching step towards practical quantum learning supremacy in the future.

Received: 26 July 2024; Accepted: 19 August 2024;

Published online: 29 August 2024

References

1. Biamonte, J. et al. Quantum machine learning. *Nature* **549**, 195 (2017).
2. Sarma, S. D., Deng, D.-L. & Duan, L.-M. Machine learning meets quantum physics. *Phys. Today* **72**, 48 (2019).
3. Arunachalam, S. & de Wolf, R. Guest column: A survey of quantum learning theory. *ACM Sigact N.* **48**, 41–67 (2017).
4. Dunjko, V. & Briegel, H. J. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep. Prog. Phys.* **81**, 074001 (2018).
5. Ciliberto, C. et al. Quantum machine learning: a classical perspective. *Proc. R. Soc. A* **474**, 20170551 (2017).
6. Carleo, G. et al. Machine learning and the physical sciences. *Rev. Mod. Phys.* **91**, 045002 (2019).
7. Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.* **103**, 150502 (2009).

8. Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. *Nat. Phys.* **10**, 631–633 (2014).
9. Lloyd, S. & Weedbrook, C. Quantum generative adversarial learning. *Phys. Rev. Lett.* **121**, 040502 (2018).
10. Rebentrost, P., Mohseni, M. & Lloyd, S. Quantum support vector machine for big data classification. *Phys. Rev. Lett.* **113**, 130503 (2014).
11. Dunjko, V., Taylor, J. M. & Briegel, H. J. Quantum-enhanced machine learning. *Phys. Rev. Lett.* **117**, 130501 (2016).
12. Amin, M. H., Andriyash, E., Rolfe, J., Kulchitsky, B. & Melko, R. Quantum Boltzmann machine. *Phys. Rev. X* **8**, 021050 (2018).
13. Gao, X., Zhang, Z.-Y. & Duan, L.-M. A quantum machine learning algorithm based on generative models. *Sci. Adv.* **4**, eaat9004 (2018).
14. Hu, L. et al. Quantum generative adversarial learning in a superconducting quantum circuit. *Sci. Adv.* **5**, eaav2761 (2019).
15. Liu, Y., Arunachalam, S. & Temme, K. A rigorous and robust quantum speed-up in supervised machine learning. *Nat. Phys.* **17**, 1013–1017 (2021).
16. Huang, H.-Y., Kueng, R. & Preskill, J. Information-theoretic bounds on quantum advantage in machine learning. *Phys. Rev. Lett.* **126**, 190505 (2021).
17. Bouland, A., Fefferman, B., Landau, Z. & Liu, Y. Noise and the frontier of quantum supremacy. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, 1308–1317 <https://ieeexplore.ieee.org/abstract/document/9719705> (2022).
18. Gao, X., Anschuetz, E. R., Wang, S.-T., Cirac, J. I. & Lukin, M. D. Enhancing generative models via quantum correlations. *Phys. Rev. X* **12**, 021037 (2022).
19. Gyurik, C. & Dunjko, V. Exponential separations between classical and quantum learners. *arXiv:2306.16028* <https://arxiv.org/abs/2306.16028> (2023).
20. Gyurik, C. & Dunjko, V. On establishing learning separations between classical and quantum machine learning with classical data. *arXiv:2208.06339* <https://arxiv.org/pdf/2208.06339.pdf> (2022).
21. Molteni, R., Gyurik, C. & Dunjko, V. Exponential quantum advantages in learning quantum observables from classical data. *arXiv:2405.02027* <https://arxiv.org/abs/2405.02027> (2024).
22. Huang, H.-Y. et al. Power of data in quantum machine learning. *Nat. Commun.* **12**, 2631 (2021).
23. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021).
24. Havlíček, V. et al. Supervised learning with quantum-enhanced feature spaces. *Nature* **567**, 209–212 (2019).
25. Mitarai, K., Negoro, M., Kitagawa, M. & Fujii, K. Quantum circuit learning. *Phys. Rev. A* **98**, 032309 (2018).
26. Grant, E. et al. Hierarchical quantum classifiers. *npj Quant. Inf.* **4**, 65 (2018).
27. Farhi, E., Goldstone, J. & Gutmann, S. A quantum approximate optimization algorithm. *arXiv:1411.4028* <https://arxiv.org/abs/1411.4028> (2014).
28. Farhi, E. & Neven, H. Classification with quantum neural networks on near term processors. *arXiv:1802.06002* <https://arxiv.org/abs/1802.06002> (2018).
29. Li, W., Lu, Z. & Deng, D.-L. Quantum neural network classifiers: A tutorial. *SciPost Phys. Lect. Notes*, 61 (2022).
30. Li, Q. et al. Ensemble-learning error mitigation for variational quantum shallow-circuit classifiers. *Phys. Rev. Res.* **6**, 013027 (2024).
31. Ren, W. et al. Experimental quantum adversarial learning with programmable superconducting qubits. *Nat. Comput. Sci.* **2**, 711–717 (2022).
32. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
33. Bravyi, S., Gosset, D. & König, R. Quantum advantage with shallow circuits. *Science* **362**, 308–311 (2018).
34. Le Gall, F. Average-case quantum advantage with shallow circuits. In *Proc. of the 34th Computational Complexity Conference (CCC)*, 1–20 <https://dl.acm.org/doi/abs/10.4230/LIPLcs.CCC.2019.21> (2019).
35. Watts, A. B., Kothari, R., Schaeffer, L. & Tal, A. Exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits. In *Proc. of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 515–526 <https://dl.acm.org/doi/abs/10.1145/3313276.3316404> (2019).
36. Coudron, M., Stark, J. & Vidick, T. Trading locality for time: certifiable randomness from low-depth circuits. *Commun. Math. Phys.* **382**, 49–86 (2021).
37. Aasnaess, S. Comparing two cohomological obstructions for contextuality, and a generalised construction of quantum advantage with shallow circuits. *arXiv:2212.09382* <https://arxiv.org/abs/2212.09382> (2022).
38. Bravyi, S., Gosset, D., König, R. & Tomamichel, M. Quantum advantage with noisy shallow circuits. *Nat. Phys.* **16**, 1040–1045 (2020).
39. Grier, D., Ju, N. & Schaeffer, L. Interactive quantum advantage with noisy, shallow Clifford circuits. *arXiv:2102.06833* <https://arxiv.org/abs/2102.06833> (2021).
40. Watts, A. B. & Parham, N. Unconditional quantum advantage for sampling with shallow circuits. *arXiv:2301.00995* <https://arxiv.org/abs/2301.00995> (2023).
41. Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. & Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **86**, 419–478 (2014).
42. Kirillov, A. et al. Segment anything. In *Proc. of the IEEE/CVF International Conference on Computer Vision*, 4015–4026 http://openaccess.thecvf.com/content/ICCV2023/html/Kirillov_Segment_Anything_ICCV_2023_paper.html (2023).
43. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18*, 234–241 https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28 (Springer, 2015).
44. Sutskever, I., Vinyals, O. & Le, Q. V. Sequence to sequence learning with neural networks. In *Proc. of 27th Advances in Neural Information Processing Systems*, **27** <https://proceedings.neurips.cc/paper/2014/file/a14ac55a4f27472c5d894ec1c3c743d2-Paper.pdf> (2014).
45. Vaswani, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems* **30** (2017).
46. Barrett, J., Caves, C. M., Eastin, B., Elliott, M. B. & Pironio, S. Modeling Pauli measurements on graph states with nearest-neighbor classical communication. *Phys. Rev. A* **75**, 012103 (2007).
47. Hein, M., Eisert, J. & Briegel, H. J. Multiparty entanglement in graph states. *Phys. Rev. A* **69**, 062311 (2004).
48. Kullback, S. & Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951).
49. Wang, S. et al. Can error mitigation improve trainability of noisy variational quantum algorithms? *Quantum* **8**, 1287 (2024).
50. Cerezo, M., Sone, A., Volkoff, T., Cincio, L. & Coles, P. J. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat. Commun.* **12**, 1–12 (2021).
51. Wang, S. et al. Noise-induced barren plateaus in variational quantum algorithms. *Nat. Commun.* **12**, 1–11 (2021).
52. Anschuetz, E. R. & Kiani, B. T. Quantum variational algorithms are swamped with traps. *Nat. Commun.* **13**, 7760 (2022).
53. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
54. Chen, S., Cotler, J., Huang, H.-Y. & Li, J. The complexity of NISQ. *Nat. Commun.* **14**, 6001 (2023).
55. Aharonov, D., Gao, X., Landau, Z., Liu, Y. & Vazirani, U. A polynomial-time classical algorithm for noisy random circuit sampling. In *Proc. of*

- the 55th Annual ACM Symposium on Theory of Computing, 945–957 <https://dl.acm.org/doi/abs/10.1145/3564246.3585234> (2023).
56. Liu, Y., Otten, M., Bassirianjahromi, R., Jiang, L. & Fefferman, B. Benchmarking near-term quantum computers via random circuit sampling. *arXiv:2105.05232* <https://arxiv.org/abs/2105.05232> (2021).
 57. Dehaene, J. & De Moor, B. Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Phys. Rev. A* **68**, 042318 (2003).
 58. Van Den Nes, M. Classical simulation of quantum computation, the Gottesman-Knill theorem, and slightly beyond. *Quantum Inf. Comput.* **10**, 258–271 (2010).
 59. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. In *Proc. of the 43rd annual ACM Symposium on Theory of Computing (STOC)*, 333–342 <https://dl.acm.org/doi/abs/10.1145/1993636.1993682> (2011).
 60. Boixo, S. et al. Characterizing quantum supremacy in near-term devices. *Nat. Phys.* **14**, 595–600 (2018).
 61. Bouland, A., Fefferman, B., Nirkhe, C. & Vazirani, U. On the complexity and verification of quantum random circuit sampling. *Nat. Phys.* **15**, 159–163 (2019).
 62. Movassagh, R. Quantum supremacy and random circuits. *arXiv:1909.06210* <https://arxiv.org/abs/1909.06210> (2019).
 63. Aaronson, S. & Chen, L. Complexity-theoretic foundations of quantum supremacy experiments. In *Proc. of the 32nd Computational Complexity Conference (CCC)*, 1–67 <https://dl.acm.org/doi/abs/10.5555/3135595.3135617> (2017).
 64. Aaronson, S. & Gunn, S. On the classical hardness of spoofing linear cross-entropy benchmarking. *Theory Comput.* **16**, 1–8 (2020).
 65. Napp, J. C., La Placa, R. L., Dalzell, A. M., Brandao, F. G. & Harrow, A. W. Efficient classical simulation of random shallow 2D quantum circuits. *Phys. Rev. X* **12**, 021021 (2022).
 66. Wang, D. Possibilistic simulation of quantum circuits by classical circuits. *Phys. Rev. A* **106**, 062430 (2022).

Acknowledgements

We thank Sitan Chen, Tongyang Li, Dong Yuan, Zidu Liu, Qi Ye, Xingjian Li, Zhide Lu, and Xun Gao for helpful discussions. This work is supported by the National Natural Science Foundation of China (Grants No. 12075128 and No. T2225008), the Innovation Program for Quantum Science and Technology (No. 2021ZD0302203), Tsinghua University Dushi Program, and the Shanghai Qi Zhi Institute.

Author contributions

Z.Z., W.G., W.L., and D.-L.D. developed the theory and contributed to writing the manuscript. D.-L.D. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42005-024-01783-7>.

Correspondence and requests for materials should be addressed to Dong-Ling Deng.

Peer review information This manuscript has been previously reviewed at another Nature Portfolio journal. The manuscript was considered suitable for publication without further review at *Communications Physics*.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024