

Exponentially many initializations to avoid barren plateaus: too much choice?

Ankit Kulshrestha*, Ricard Puig*, Diego García-Martín, Lukasz Cincio, Ilya Safro, Zoë Holmes, M. Cerezo

This work is based on Ref. [1].

Introduction Over the past few years, a significant amount of effort has been put towards understanding if and how parametrized quantum circuits (PQCs) can be trained to accurately and efficiently find the global minima of loss functions for variational quantum algorithms [2–8]. A central challenge in this task is the barren plateau phenomenon, where loss function values and their gradients concentrate exponentially [9, 10]. When a landscape exhibits a barren plateau, an exponential amount of measurements (in the number of qubits) are required just to determine the direction of steepest descent, rendering the optimization practically impossible for large system sizes.

Strategies to avoid barren plateaus generally fall into two broad categories. The first restricts the relevant dynamics to some smaller effective subspaces, either by employing shallow local architectures [11–18], incorporating symmetries [19–29], or ensuring the circuit spans small dynamical Lie algebras [30–41]. These methods fundamentally restrict the expressivity of the model, taming the underlying curse of dimensionality. However, this restriction also makes the model potentially more amenable to classical simulation [17, 27, 42–52], potentially undermining the quantum advantage. The second route attempts to avoid the bad average-case scenario without fundamentally restricting the ansatz, doing so by initializing the parameters in a smart way [53–64]. By steering optimization toward special patches of the landscape where gradients remain large, the model preserves its full expressivity. While promising, this introduces a new challenge: because the landscape is highly non-convex, the choice of “where to star” becomes critical due to the existence of exponentially many local minima [65, 66].

In this work, we focus on this second route and demonstrate that avoiding barren plateaus via initialization is profoundly non-unique. We develop an analytical framework, based on the first-moment operator, to diagnose when a parameter initialization may avoid the fully concentrated barren plateau regime, and we show that there exist *exponentially many* such initializations that are nonequivalent. As illustrated in Figure 1a), once exponentially small gradients are avoided, the relevant question becomes which initialization among many possible choices leads to a sufficiently favorable region of the landscape.

General Framework We consider a generic loss function that we aim to minimize $\min_{\alpha} \mathcal{L}(\alpha)$, where $\mathcal{L}$ takes the form

$$\mathcal{L}(\alpha) = \text{Tr}[U(\alpha)\rho U^\dagger(\alpha)O], \quad U(\alpha) = \prod_{l=1}^L e^{-i\alpha_l \sigma_l/2} V_l, \quad (1)$$

for some traceless Hermitian observable O such that $\|O\|_2 \in \mathcal{O}(2^n)$. Where σ_l is a Pauli operator taken from a set of generators $\mathcal{G}$, $\alpha = \{\theta_l\}_{l=1}^L$ a set of trainable parameters, and V_l a non-parametrized gate. In what follows, we will assume that the PQC is universal. However, the techniques can be extended to further symmetries or subgroups. We can define the moment operator of the circuit $\hat{\tau}_{\gamma}^{(t)} = \prod_{l=1}^L \hat{\tau}_{\gamma_l}^{(t)}$, where $\hat{\tau}_{\gamma_l}^{(t)}$ is the t -moment operator for the parametrized unitary matrix in the vectorized picture. γ_l denotes the distribution used to initialize the parameters at that gate. This can be related to the t -moment of the loss function $\mathbb{E}_{\gamma}[\mathcal{L}(\alpha)^t] = \langle\langle \rho^{\otimes t} | \hat{\tau}_{\gamma}^{(t)} | O^{\otimes t} \rangle\rangle$. With these tools and definitions we are ready to start presenting the results.

First-Moment Framework for Parameterized Quantum Circuits Our analytical framework starts by building upon the analysis of first moment operators. We realize that a way to probe for BPs is to determine whether the average between two distributions is the same. Specifically, if an initialization distribution leads the loss function to concentrate around a value that is substantially different from the fixed point of the fully concentrated barren plateau landscape, then the initialization successfully avoids the barren plateau. Then, a necessary condition to avoid BPs is that the distance between the two first-moment twirls applied to the observable is at worst polynomially vanishing $\|\hat{\tau}_{\text{BP}}|O\rangle\rangle - \hat{\tau}_{\gamma}|O\rangle\rangle\|_1 \in \Omega(1/\text{poly}(n))$. Thus, what follows is that if we can show this distance is non-exponentially vanishing, we can probe distributions for the absence of BPs. More generally, we introduce the following definition to compare initialization strategies.

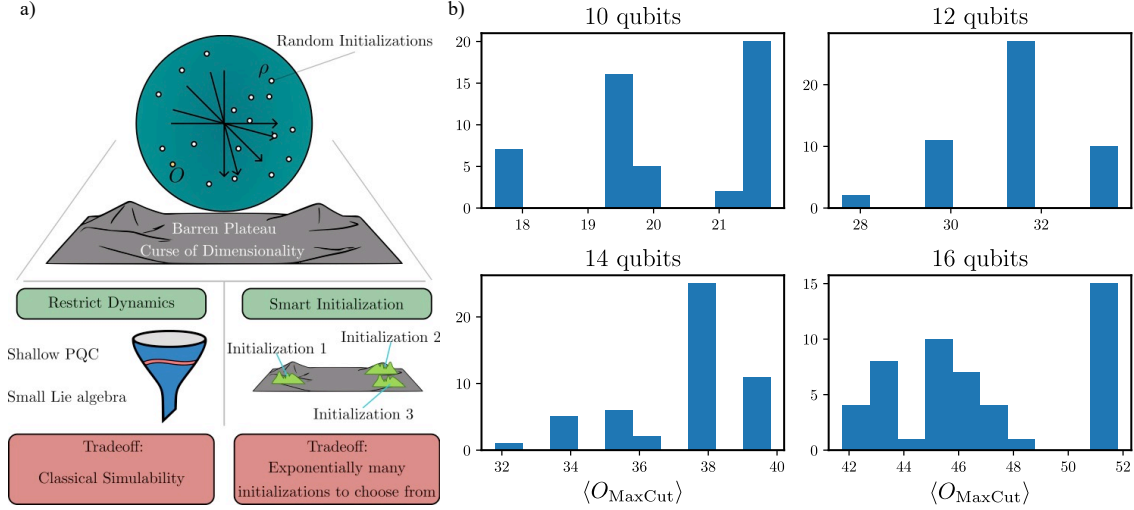


Figure 1: **a) Schematic representation.** Avoiding the barren plateau regime (gray) by selecting a non-concentrated initialization forces a choice among many inequivalent “pockets” (colored regions) with different local minima. **b) Distribution of local minima after training for a Max-Cut problem.** For each $n \in \{10, 12, 14, 16\}$, we initialize the PQC with L layers, where each layer consists of three single-qubit parametrized rotations on every qubit (R_z , R_y , R_z) together with fixed nearest-neighbor CZ gates. We take $L = n$, initialize the R_z angles uniformly in $[0, 2\pi)$, and the R_y angles from a Gaussian distribution $\mathcal{N}(\theta^*, \zeta^2)$ with $\zeta^2 = 1/n$ and $\theta^* \in \{0, \pi\}$ chosen uniformly at random for each R_y gate. The optimization is run for 500 iterations, or until the gradient norm falls below the BP.

Definition 1 (Expected Heisenberg-evolved operator gap). *Let $U(\theta)$ be a parametrized quantum circuit, let O be an observable, and let $\hat{\tau}_\alpha$ and $\hat{\tau}_\gamma$ be two first-moment operators associated with probability distributions $P_\alpha(\theta)$ and $P_\gamma(\theta)$. We define the expected Heisenberg-evolved operator gap as $\xi_{\alpha,\gamma} = \|\hat{\tau}_\alpha|O\rangle\rangle - \hat{\tau}_\gamma|O\rangle\rangle\|_1$.*

For now, let us understand what each of these different strategies does at a fundamental level to the observable. To do so, we present Theorem 1.

Theorem 1 (Informal). *Let $U = e^{-i\theta\sigma_l/2}$, where θ is sampled according to $P_{\gamma_l}(\theta)$. Then, $\hat{\tau}_{\gamma_l} = (W_l \otimes W_l^*) \eta_l$, where η_l is a contraction matrix, and $W_l = e^{-i\phi_l\sigma_l/2}$, $\phi_l = \arccos\left(\frac{a_l}{\sqrt{a_l^2 + b_l^2}}\right)$. Where we used*

$$a_l = \int P_{\gamma_l}(\alpha) \cos(\alpha) d\alpha, \quad b_l = \int P_{\gamma_l}(\alpha) \sin(\alpha) d\alpha. \quad (2)$$

In particular, the action of the matrix η_l on Pauli operators is given by $\eta_l|\sigma_i\rangle\rangle = \kappa_l|\sigma_i\rangle\rangle$ if $\{\sigma_l, \sigma_i\} = 0$, and $\eta_l|\sigma_i\rangle\rangle = |\sigma_i\rangle\rangle$ if $[\sigma_l, \sigma_i] = 0$, where the damping factor is $\kappa_l = \sqrt{a_l^2 + b_l^2}$.

Theorem 1 demonstrates that the first moment operator admits a simple operational decomposition: it acts as a positive contraction followed by a rotation in operator space. The initialization distribution plays two distinct roles: the damping factor κ_l determines how strongly averaging damps operator components, while the rotation W_l determines the direction in operator space toward which those components are biased.

Crucially, the contraction factor controls the lower bound for avoiding barren plateaus. We define the Heisenberg-evolved operator gap $\xi_{\text{BP},\gamma}$ as a measure of distance from the fully concentrated state. In the manuscript we present Theorem 2 that proves $\xi_{\text{BP},\gamma} \in \Omega\left(\prod_{l=1}^L \kappa_l\right) \in \Omega(\kappa)$. Theorem 2 is quite general and can be used to assess how a given parameter-initialization strategy affects the gap $\xi_{\text{BP},\gamma}$ via the damping factors κ_l . When this lower bound is inverse-polynomial, the initialization passes our first-moment escape criterion.

Using this framework, we introduce a concrete example that shows that this approach successfully identifies initializations maintaining a polynomial separation from the barren plateau ensemble. Furthermore, the same operator-distance methods can be used to directly compare distinct initializations. To showcase this, we present a simplified version of Example 9 in the manuscript.

Example 9 (Exponentially many initializations for MaxCut, Informal). Consider a parameterized quantum circuit applied to a Max-Cut problem on a graph $G = (V, E)$. Then, we consider an L -layered PQC composed of local R_y rotations interleaved with diagonal entangling gates. Let $\gamma_{\mathbf{c}}$ be a combinatorial family of initializations governed by an $L \times n$ binary matrix $\mathbf{c} \in \{0, 1\}^{L \times n}$. For a given $\mathbf{c}$, the initialization $\gamma_{\mathbf{c}}$ draws the parameter for the R_y gate on qubit j at layer l from a Gaussian distribution with mean $c_{l,j}\pi$, and variance ς^2 . Then there are exponentially many initializations, in the sense that there are exponentially many pairs of $\mathbf{c}, \mathbf{c}'$ in n such that $\xi_{\gamma_{\mathbf{c}}, \gamma_{\mathbf{c}'}}$ is non-exponentially vanishing and avoids BPs (in the same sense).

This example highlights the existence of many different initializations that avoid Barren Plateaus. However, we further formalize this by presenting two extra Theorems.

Exponentially Many Initializations Moving beyond the first moment analysis, we formalize the following Theorems proving that there are exponentially many initializations that avoid BPs. In our Theorem 3 in the manuscript we extend the previous example to more general PQC, and show that these initializations are barren plateau-free. In the following theorem we reveal richer structures for specific problem Hamiltonians, such as Max-Cut.

Theorem 4. Let $U(\theta)$ be a PQC composed of R_y and R_z rotations interleaved with diagonal entangling gates for $L \in \mathcal{O}(\text{poly}(n))$. Let $O = \sum_{(w,w') \in E} \sigma_z^{(w)} \sigma_z^{(w')}$ be the Max-Cut Hamiltonian on a constant-odd-degree graph $G = (E, V)$, with $S \subseteq V$ a subset of vertices. For each $j \in S$, let one $\theta_{i,j}^y$ angle be initialized in a **constant-size** region (non-shrinking volume). While the rest of the parameters are initialized following $\mathcal{N}(\theta^*, \varsigma^2)$ with $\theta^* \in \{0, \pi\}$ and $\varsigma \in \Theta(1/\text{poly}(n))$, the gradients are guaranteed to be in $\Omega(1/\text{poly}(n))$ with high probability.

Theorem 4 demonstrates that for specific problem structures, one can pick a subset of $\Theta(n)$ angles and freely initialize them in a region whose volume does *not* shrink with growing system size. Having made this choice, there are then exponentially many choices for how to initialize the remaining parameters.

Numerical Results and Implications While our theoretical results rigorously confirm the existence of exponentially many barren plateau-free initializations, they do not prescribe how these initializations impact the actual training. To investigate this, we conducted numerical simulations for a Max-Cut problem. Concretely, we draw a random Erdős–Rényi graph with edge probability $p = 0.5$. Then we follow the standard variational quantum eigensolver pipeline to minimize the expectation value of the corresponding Max-Cut Hamiltonian. We perform training using vanilla gradient descent such that the optimization itself is deterministic and all variation between solutions on different runs comes from the initialization.

As depicted in Figure 1b), for each system size the final objective values are broadly distributed and multi-modal. Hence, even within this single family of barren-plateau-free initializations, training does not repeatedly converge to one effectively unique minimum. Rather, the optimizer terminates in a collection of distinct local minima of different quality. This provides direct numerical evidence that not all the exponentially many initialization choices sampled here are practically interchangeable: they bias the optimizer toward different basins of attraction. As such, these results indicate that avoiding the fully concentrated barren-plateau regime does not collapse the optimization landscape into a single well-behaved region. Instead, it can reveal a landscape containing many trainable pockets, not all of which are equally favorable for the task at hand.

Conclusions By analyzing the first-moment operator of parameterized quantum circuits and the different Theorems, we have demonstrated that a careful initialization scheme does not result in a single valid starting point, but rather exponentially many non-equivalent initializations that successfully maintain large gradients. Furthermore, these initializations are fundamentally inequivalent. They prepare the circuit in completely distinct sub-regions (or “pockets”) of the Hilbert space. Consequently, avoiding barren plateaus via smart initializations does not guarantee a successful optimization outcome. Instead, it merely trades the exponential concentration problem (where training cannot even begin) for an equally daunting problem: selecting the correct trainable pocket among exponentially many options to actually find the global minimum. This underscores the immense subtlety and complexity of variational quantum optimization in realistic landscapes.

References

- [1] Ankit Kulshrestha, Ricard Puig, Diego García-Martín, Lukasz Cincio, Ilya Safro, Zoë Holmes, and M Cerezo. Exponentially many initializations to avoid barren plateaus. *arXiv preprint arXiv:2606.18515*, 2026. URL: <https://arxiv.org/abs/2606.18515>, doi:10.48550/arXiv.2606.18515.
- [2] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(1):625–644, 2021. URL: <https://www.nature.com/articles/s42254-021-00348-9>, doi:10.1038/s42254-021-00348-9.
- [3] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, 94(1):015004, 2022. URL: <https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.94.015004>, doi:10.1103/RevModPhys.94.015004.
- [4] Suguru Endo, Zhenyu Cai, Simon C Benjamin, and Xiao Yuan. Hybrid quantum-classical algorithms and quantum error mitigation. *Journal of the Physical Society of Japan*, 90(3):032001, 2021. doi:10.7566/JPSJ.90.032001.
- [5] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, 2015. URL: <https://www.tandfonline.com/doi/full/10.1080/00107514.2014.964942>, doi:10.1080/00107514.2014.964942.
- [6] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017. URL: <https://www.nature.com/articles/nature23474>, doi:10.1038/nature23474.
- [7] M Cerezo, Guillaume Verdon, Hsin-Yuan Huang, Lukasz Cincio, and Patrick J Coles. Challenges and opportunities in quantum machine learning. *Nature Computational Science*, 2022. URL: <https://www.nature.com/articles/s43588-022-00311-3>, doi:10.1038/s43588-022-00311-3.
- [8] Alberto Di Meglio, Karl Jansen, Ivano Tavernelli, Constantia Alexandrou, Srinivasan Arunachalam, Christian W Bauer, Kerstin Borrás, Stefano Carrazza, Arianna Crippa, Vincent Croft, et al. Quantum computing for high-energy physics: state of the art and challenges. summary of the qc4hep working group. *Prx quantum*, 5(3):037001, 2024. URL: <https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.5.037001>, doi:10.1103/PRXQuantum.5.037001.
- [9] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):1–6, 2018. doi:10.1038/s41467-018-07090-4.
- [10] Martin Larocca, Supanut Thanasilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J Coles, Lukasz Cincio, Jarrod R McClean, Zoë Holmes, and M Cerezo. Barren plateaus in variational quantum computing. *Nature Reviews Physics*, 7(4):174–189, 2025. URL: <https://www.nature.com/articles/s42254-025-00813-9>, doi:10.1038/s42254-025-00813-9.
- [11] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1):1–12, 2021. doi:10.1038/s41467-021-21728-w.
- [12] Arthur Pesah, M. Cerezo, Samson Wang, Tyler Volkoff, Andrew T Sornborger, and Patrick J Coles. Absence of barren plateaus in quantum convolutional neural networks. *Physical Review X*, 11(4):041011, 2021. URL: <https://journals.aps.org/prx/abstract/10.1103/PhysRevX.11.041011>, doi:10.1103/PhysRevX.11.041011.

- [13] Sumeet Khatri, Ryan LaRose, Alexander Poremba, Lukasz Cincio, Andrew T Sornborger, and Patrick J Coles. Quantum-assisted quantum compiling. *Quantum*, 3:140, 2019. URL: <https://quantum-journal.org/papers/q-2019-05-13-140/>, doi:10.22331/q-2019-05-13-140.
- [14] Chen Zhao and Xiao-Shan Gao. Analyzing the barren plateau phenomenon in training quantum neural networks with the ZX-calculus. *Quantum*, 5:466, June 2021. doi:10.22331/q-2021-06-04-466.
- [15] Zidu Liu, Li-Wei Yu, L-M Duan, and Dong-Ling Deng. The presence and absence of barren plateaus in tensor-network based machine learning. *Physical Review Letters*, 129(27):270501, 2022. URL: <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.129.270501>, doi:10.1103/PhysRevLett.129.270501.
- [16] Qiang Miao and Thomas Barthel. Isometric tensor network optimization for extensive hamiltonians is free of barren plateaus. *Physical Review A*, 109(5):L050402, 2024. URL: <https://journals.aps.org/pra/abstract/10.1103/PhysRevA.109.L050402>, doi:10.1103/PhysRevA.109.L050402.
- [17] Pablo Bermejo, Paolo Braccia, Manuel S. Rudolph, Zoë Holmes, Lukasz Cincio, and M. Cerezo. Quantum convolutional neural networks are effectively classically simulable. *PRX Quantum*, 7:020304, Apr 2026. URL: <https://link.aps.org/doi/10.1103/8qt9-72ts>, doi:10.1103/8qt9-72ts.
- [18] Bao Bach, Jose Falla, and Ilya Safro. MLQAOA: Graph learning accelerated hybrid quantum-classical multilevel QAOA. In *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, volume 1, pages 1–12. IEEE, 2024.
- [19] Martín Larocca, Frédéric Sauvage, Faris M. Sbahi, Guillaume Verdon, Patrick J. Coles, and M. Cerezo. Group-invariant quantum machine learning. *PRX Quantum*, 3:030341, Sep 2022. URL: <https://link.aps.org/doi/10.1103/PRXQuantum.3.030341>, doi:10.1103/PRXQuantum.3.030341.
- [20] Johannes Jakob Meyer, Marian Mularski, Elies Gil-Fuster, Antonio Anna Mele, Francesco Arzani, Alissa Wilms, and Jens Eisert. Exploiting symmetry in variational quantum machine learning. *PRX Quantum*, 4(1):010328, 2023. URL: <https://link.aps.org/doi/10.1103/PRXQuantum.4.010328>, doi:10.1103/PRXQuantum.4.010328.
- [21] Andrea Skolik, Michele Cattelan, Sheir Yarkoni, Thomas Bäck, and Vedran Dunjko. Equivariant quantum circuits for learning on weighted graphs. *npj Quantum Information*, 9(1):47, 2023. URL: <https://www.nature.com/articles/s41534-023-00710-y>, doi:10.1038/s41534-023-00710-y.
- [22] Michael Ragone, Quynh T. Nguyen, Louis Schatzki, Paolo Braccia, Martin Larocca, Frederic Sauvage, Patrick J. Coles, and M. Cerezo. Representation theory for geometric quantum machine learning. *arXiv preprint arXiv:2210.07980*, 2022. URL: <https://arxiv.org/abs/2210.07980>, doi:10.48550/arXiv.2210.07980.
- [23] Quynh T. Nguyen, Louis Schatzki, Paolo Braccia, Michael Ragone, Patrick J. Coles, Frédéric Sauvage, Martín Larocca, and M. Cerezo. Theory for equivariant quantum neural networks. *PRX Quantum*, 5:020328, May 2024. URL: <https://link.aps.org/doi/10.1103/PRXQuantum.5.020328>, doi:10.1103/PRXQuantum.5.020328.
- [24] Louis Schatzki, Martin Larocca, Quynh T Nguyen, Frederic Sauvage, and Marco Cerezo. Theoretical guarantees for permutation-equivariant quantum neural networks. *npj Quantum Information*, 10(1):12, 2024. URL: <https://www.nature.com/articles/s41534-024-00804-1>, doi:10.1038/s41534-024-00804-1.
- [25] Han Zheng, Zimu Li, Junyu Liu, Sergii Strelchuk, and Risi Kondor. Speeding up learning quantum states through group equivariant convolutional quantum ansätze. *PRX Quantum*, 4(2):020327, 2023. URL: <https://link.aps.org/doi/10.1103/PRXQuantum.4.020327>, doi:10.1103/PRXQuantum.4.020327.
- [26] Richard D P East, Guillermo Alonso-Linaje, and Chae-Yeun Park. All you need is spin: $Su(2)$ equivariant variational quantum circuits based on spin networks. *Quantum Science and Technology*, 11(2):025025, mar 2026. doi:10.1088/2058-9565/ae4cff.

- [27] Su Yeon Chang, Martin Larocca, and Marco Cerezo. Practical framework for simulating permutation-equivariant quantum circuits. *Manuscript in preparation*, 2026.
- [28] Boris Tselikhovskiy, Ilya Safro, and Yuri Alexeev. Equivariant quantum approximate optimization algorithm. *IEEE Transactions on Quantum Engineering*, 2026.
- [29] Ruslan Shaydulin, Stuart Hadfield, Tad Hogg, and Ilya Safro. Classical symmetries and the quantum approximate optimization algorithm. *Quantum Information Processing*, 20:1–28, 2021. URL: <https://doi.org/10.1007%2Fs11128-021-03298-4>, doi:10.1007/s11128-021-03298-4.
- [30] Martin Larocca, Piotr Czarnik, Kunal Sharma, Gopikrishnan Muraleedharan, Patrick J. Coles, and M. Cerezo. Diagnosing Barren Plateaus with Tools from Quantum Optimal Control. *Quantum*, 6:824, September 2022. doi:10.22331/q-2022-09-29-824.
- [31] Léo Monbroussou, Jonas Landman, Alex B Grilo, Romain Kukla, and Elham Kashefi. Trainability and expressivity of hamming-weight preserving quantum circuits for machine learning. *Quantum*, 9:1745, 2025. URL: <https://quantum-journal.org/papers/q-2025-05-15-1745/>, doi:10.22331/q-2025-05-15-1745.
- [32] Snehal Raj, Iordanis Kerenidis, Abhishek Shekhar, Ben Wood, Jon Dee, Shouvanik Chakrabarti, Richard Chen, Dylan Herman, Shaohan Hu, Pierre Minssen, et al. Quantum deep hedging. *Quantum*, 7:1191, 2023. URL: <https://quantum-journal.org/papers/q-2023-11-29-1191/>, doi:10.22331/q-2023-11-29-1191.
- [33] Enrico Fontana, Dylan Herman, Shouvanik Chakrabarti, Niraj Kumar, Romina Yalovetzky, Jamie Heredge, Shree Hari Sureshbabu, and Marco Pistoia. Characterizing barren plateaus in quantum ansätze with the adjoint representation. *Nature Communications*, 15(1):7171, 2024. URL: <https://www.nature.com/articles/s41467-024-49910-w>, doi:10.1038/s41467-024-49910-w.
- [34] Michael Ragone, Bojko N Bakalov, Frédéric Sauvage, Alexander F Kemper, Carlos Ortiz Marrero, Martín Larocca, and M Cerezo. A lie algebraic theory of barren plateaus for deep parameterized quantum circuits. *Nature Communications*, 15(1):7172, 2024. URL: <https://www.nature.com/articles/s41467-024-49909-3#citeas>, doi:10.1038/s41467-024-49909-3.
- [35] N L Diaz, Diego García-Martín, Sujay Kazi, Martin Larocca, and M Cerezo. Showcasing a barren plateau theory beyond the dynamical lie algebra. *arXiv preprint arXiv:2310.11505*, 2023. URL: <https://arxiv.org/abs/2310.11505>, doi:10.48550/arXiv.2310.11505.
- [36] Maxwell T West, Jamie Heredge, Martin Sevier, and Muhammad Usman. Provably trainable rotationally equivariant quantum machine learning. *PRX Quantum*, 5(3):030320, 2024. URL: <https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.5.030320>, doi:10.1103/PRXQuantum.5.030320.
- [37] Jamie Heredge, Niraj Kumar, Dylan Herman, Shouvanik Chakrabarti, Romina Yalovetzky, Shree Hari Sureshbabu, Changhao Li, and Marco Pistoia. Prospects of privacy advantage in quantum machine learning. *arXiv preprint arXiv:2405.08801*, 2024. URL: <https://arxiv.org/abs/2405.08801>.
- [38] Gerard Aguilar, Simon Cichy, Jens Eisert, and Lennart Bittel. Full classification of pauli lie algebras. *arXiv preprint arXiv:2408.00081*, 2024. URL: <https://arxiv.org/abs/2408.00081>, doi:10.48550/arXiv.2408.00081.
- [39] Efehan Kökcü, Roeland Wiersema, Alexander F. Kemper, and Bojko N. Bakalov. Classification of dynamical lie algebras generated by spin interactions on undirected graphs. *arXiv preprint arXiv:2409.19797*, 2024. URL: <https://arxiv.org/abs/2409.19797>, doi:10.48550/arXiv.2409.19797.
- [40] Sujay Kazi, Martín Larocca, Marco Farinati, Patrick J Coles, M Cerezo, and Robert Zeier. Analyzing the quantum approximate optimization algorithm: ansätze, symmetries, and lie algebras. *PRX Quantum*, 6(4):040345, 2025. URL: <https://journals.aps.org/prxquantum/abstract/10.1103/yfwq-yqmk>, doi:10.1103/yfwq-yqmk.

- [41] Boris Tselikhovski, Bao Bach, Jose Falla, and Ilya Safro. Reductions of qaoa induced by classical symmetries: Theoretical insights and practical implications. *arXiv preprint arXiv:2602.16141*, 2026.
- [42] M Cerezo, Martin Larocca, Diego García-Martín, N L Diaz, Paolo Braccia, Enrico Fontana, Manuel S Rudolph, Pablo Bermejo, Aroosa Ijaz, Supanut Thanasilp, et al. Does provable absence of barren plateaus imply classical simulability? *Nature Communications*, 16(1):7907, 2025. URL: <https://www.nature.com/articles/s41467-025-63099-6>, doi:10.1038/s41467-025-63099-6.
- [43] Armando Angrisani, Alexander Schmidhuber, Manuel S. Rudolph, M. Cerezo, Zoë Holmes, and Hsin-Yuan Huang. Classically estimating observables of noiseless quantum circuits. *Phys. Rev. Lett.*, 135:170602, Oct 2025. URL: <https://link.aps.org/doi/10.1103/lh6x-7rc3>, doi:10.1103/lh6x-7rc3.
- [44] Sacha Lerch, Ricard Puig, Manuel Rudolph, Armando Angrisani, Tyson Jones, M. Cerezo, Supanut Thanasilp, and Zoë Holmes. Efficient quantum-enhanced classical simulation for patches of quantum landscapes. *arXiv preprint arXiv:2411.19896*, 2024. URL: <https://arxiv.org/abs/2411.19896>, doi:10.48550/arXiv.2411.19896.
- [45] Eric R Anschuetz and Xun Gao. Arbitrary polynomial separations in trainable quantum machine learning. *arXiv preprint arXiv:2402.08606*, 2024. URL: <https://arxiv.org/abs/2402.08606>, doi:10.48550/arXiv.2402.08606.
- [46] Antonio Anna Mele, Armando Angrisani, Soumik Ghosh, Sumeet Khatri, Jens Eisert, Daniel Stilck França, and Yihui Quek. Noise-induced shallow circuits and the absence of barren plateaus. *Nature Physics*, pages 1–6, 2026. URL: <https://www.nature.com/articles/s41567-026-03245-z>, doi:10.1038/s41567-026-03245-z.
- [47] Seongwook Shin, Yong Siah Teo, and Hyunseok Jeong. Dequantizing quantum machine learning models using tensor networks. *Phys. Rev. Res.*, 6:023218, May 2024. URL: <https://link.aps.org/doi/10.1103/PhysRevResearch.6.023218>, doi:10.1103/PhysRevResearch.6.023218.
- [48] Igor Ermakov, Oleg Lychkovskiy, and Tim Byrnes. Unified framework for efficiently computable quantum circuits. *arXiv preprint arXiv:2401.08187*, 2024. URL: <https://arxiv.org/abs/2401.08187>, doi:10.48550/arXiv.2401.08187.
- [49] Aaron Miller, Joachim Favre, Zoë Holmes, Özlem Salehi, Rahul Chakraborty, Anton Nykänen, Zoltan Zimboras, Adam Glos, and Guillermo García-Pérez. Simulation of fermionic circuits using majorana propagation. *arXiv preprint arXiv:2503.18939*, 2025. doi:10.48550/arXiv.2503.18939.
- [50] Manuel S Rudolph, Armando Angrisani, Andrew Wright, Iwo Sanderski, Ricard Puig, and Zoë Holmes. Thermal state simulation with pauli and majorana propagation. *arXiv preprint arXiv:2602.04878*, 2026. URL: <https://www.arxiv.org/abs/2602.04878>, doi:10.48550/arXiv.2602.04878.
- [51] Adelina Bärligea, Matthew L Sims-Goh, and Jakob S Kottmann. Enabling lie-algebraic classical simulation beyond free fermions. *arXiv preprint arXiv:2604.16701*, 2026. URL: <https://arxiv.org/abs/2604.16701>, doi:10.48550/arXiv.2604.16701.
- [52] Matthew L Goh, Martin Larocca, Lukasz Cincio, M Cerezo, and Frédéric Sauvage. Lie-algebraic classical simulations for quantum computing. *Physical Review Research*, 7(3):033266, 2025. URL: <https://journals.aps.org/prresearch/abstract/10.1103/3y65-f5w6>, doi:10.1103/3y65-f5w6.
- [53] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti. An initialization strategy for addressing barren plateaus in parametrized quantum circuits. *Quantum*, 3:214, 2019. URL: <https://quantum-journal.org/papers/q-2019-12-09-214/>, doi:10.22331/q-2019-12-09-214.
- [54] Ankit Kulshrestha and Ilya Safro. Beinit: Avoiding barren plateaus in variational quantum algorithms. In *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 197–203. IEEE, 2022. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9951195>, doi:10.1109/QCE53715.2022.00039.

- [55] Samuel Duffield, Marcello Benedetti, and Matthias Rosenkranz. Bayesian learning of parameterised quantum circuits. *Machine Learning: Science and Technology*, 4(2):025007, 2023. URL: <https://iopscience.iop.org/article/10.1088/2632-2153/acc8b7>, doi:10.1088/2632-2153/acc8b7.
- [56] Valentin Heyraud, Zejian Li, Kaelan Donatella, Alexandre Le Boité, and Cristiano Ciuti. Efficient estimation of trainability for variational quantum circuits. *PRX Quantum*, 4(4):040335, 2023. URL: <https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.4.040335>, doi:10.1103/PRXQuantum.4.040335.
- [57] Kaining Zhang, Liu Liu, Min-Hsiu Hsieh, and Dacheng Tao. Escaping from the barren plateau via Gaussian initializations in deep variational quantum circuits. In *Advances in Neural Information Processing Systems*, 2022. URL: <https://openreview.net/forum?id=jXgbJdQ2YIy>.
- [58] Chae-Yeun Park and Nathan Killoran. Hamiltonian variational ansatz without barren plateaus. *Quantum*, 8:1239, 2024. URL: <https://quantum-journal.org/papers/q-2024-02-01-1239/>, doi:10.22331/q-2024-02-01-1239.
- [59] Yabo Wang, Bo Qi, Chris Ferrie, and Daoyi Dong. Trainability enhancement of parameterized quantum circuits via reduced-domain parameter initialization. *Physical Review Applied*, 22(5):054005, 2024. URL: <https://journals.aps.org/prapplied/abstract/10.1103/PhysRevApplied.22.054005>, doi:10.1103/PhysRevApplied.22.054005.
- [60] Chae-Yeun Park, Minhyeok Kang, and Joonsuk Huh. Hardware-efficient ansatz without barren plateaus in any depth. *arXiv preprint arXiv:2403.04844*, 2024. URL: <https://arxiv.org/abs/2403.04844>.
- [61] Xiao Shi and Yun Shang. Avoiding barren plateaus via Gaussian mixture model. *arXiv preprint arXiv:2402.13501*, 2024. URL: <https://arxiv.org/abs/2402.13501>.
- [62] Ricard Puig, Marc Drudis, Supanut Thanasilp, and Zoë Holmes. Variational quantum simulation: A case study for understanding warm starts. *PRX Quantum*, 6:010317, Jan 2025. URL: <https://link.aps.org/doi/10.1103/PRXQuantum.6.010317>, doi:10.1103/PRXQuantum.6.010317.
- [63] Ricard Puig, Berta Casas, Alba Cervera-Lierta, Zoë Holmes, and Adrián Pérez-Salinas. Warm starts, cold states: Exploiting adiabaticity for variational ground-states. *arXiv preprint arXiv:2602.06137*, 2026. URL: <https://arxiv.org/abs/2602.06137>, doi:10.48550/arXiv.2602.06137.
- [64] Sacha Lerch, Joseph Bowles, Ricard Puig, Erik Armengol, Zoë Holmes, and Supanut Thanasilp. Iqp born machines under data-dependent and agnostic initialization strategies. *arXiv preprint arXiv:2603.14576*, 2026. URL: <https://arxiv.org/abs/2603.14576>, doi:10.48550/arXiv.2603.14576.
- [65] Eric R Anschuetz and Bobak T Kiani. Quantum variational algorithms are swamped with traps. *Nature Communications*, 13(1):7760, 2022. doi:10.1038/s41467-022-35364-5.
- [66] Nikita A Nemkov, Evgeniy O Kiktenko, and Aleksey K Fedorov. Barren plateaus are swamped with traps. *arXiv preprint arXiv:2405.05332*, 2024. URL: <https://arxiv.org/abs/2405.05332>, doi:10.48550/arXiv.2405.05332.