جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

BARREN PLATEAUS & VARIATIONAL QUANTUM ALGORITHMS

Tackling scaling challenges in near-term quantum computing

Variational quantum algorithms (VQAs) have emerged as a central framework for near-term quantum computation, combining parametrized quantum circuits with classical optimisation loops. A key challenge to their scalability is the barren plateau phenomenon, where gradients vanish exponentially with system size, circuit depth or random initialisation.

This seminar will illustrate barren plateaus, review state-of-the-art mitigation techniques—including problem-inspired and symmetry-preserving ansätze, shallow or adaptive circuit designs, layerwise training protocols, improved initialisation schemes and advanced optimisers—and highlight open questions at the frontier of VQA research.

Dr. Mohammad Haidar

Lead & Founder of OpenVQA Hub, Senior Quantum Scientist

Creator OpenVQA, founder OpenVQA Hub, founder Wyw, Paris France, developing open-source tools for variational quantum algorithms and their applications in quantum chemistry and beyond.

Date: December 3, 2025
Time: 4:00 PM KSA
Format: Hybrid (Spine Auditorium & Zoom)

Hosted by the Quantum Computing Reading Group (QCRG) at KAUST

BARREN PLATEAUS

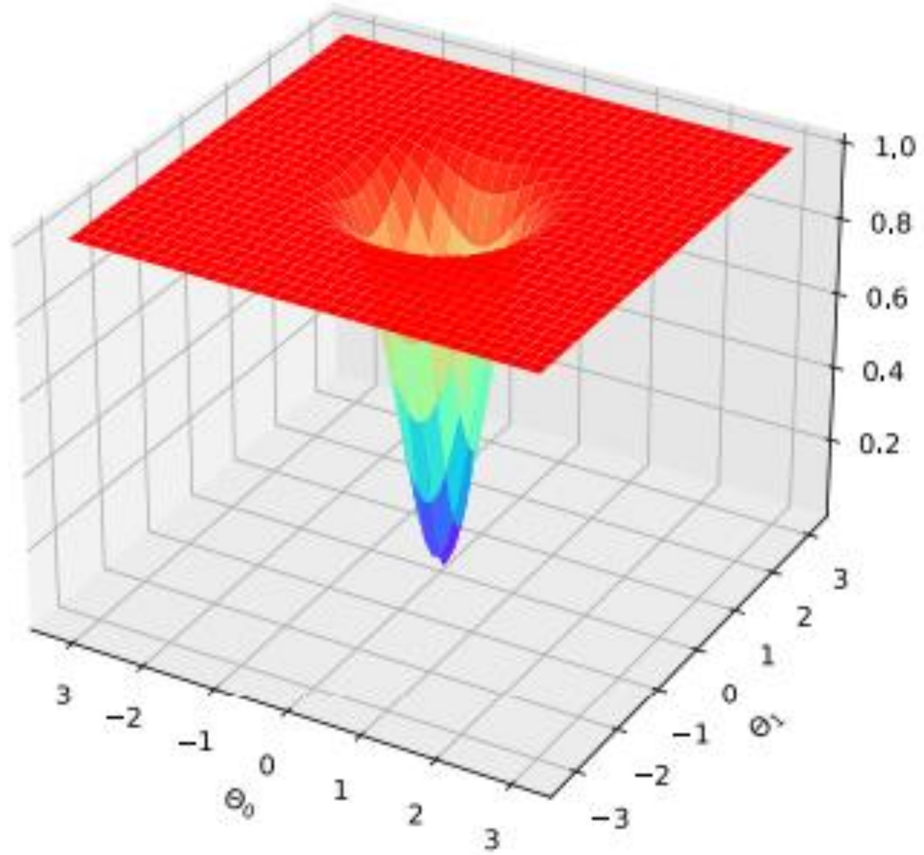
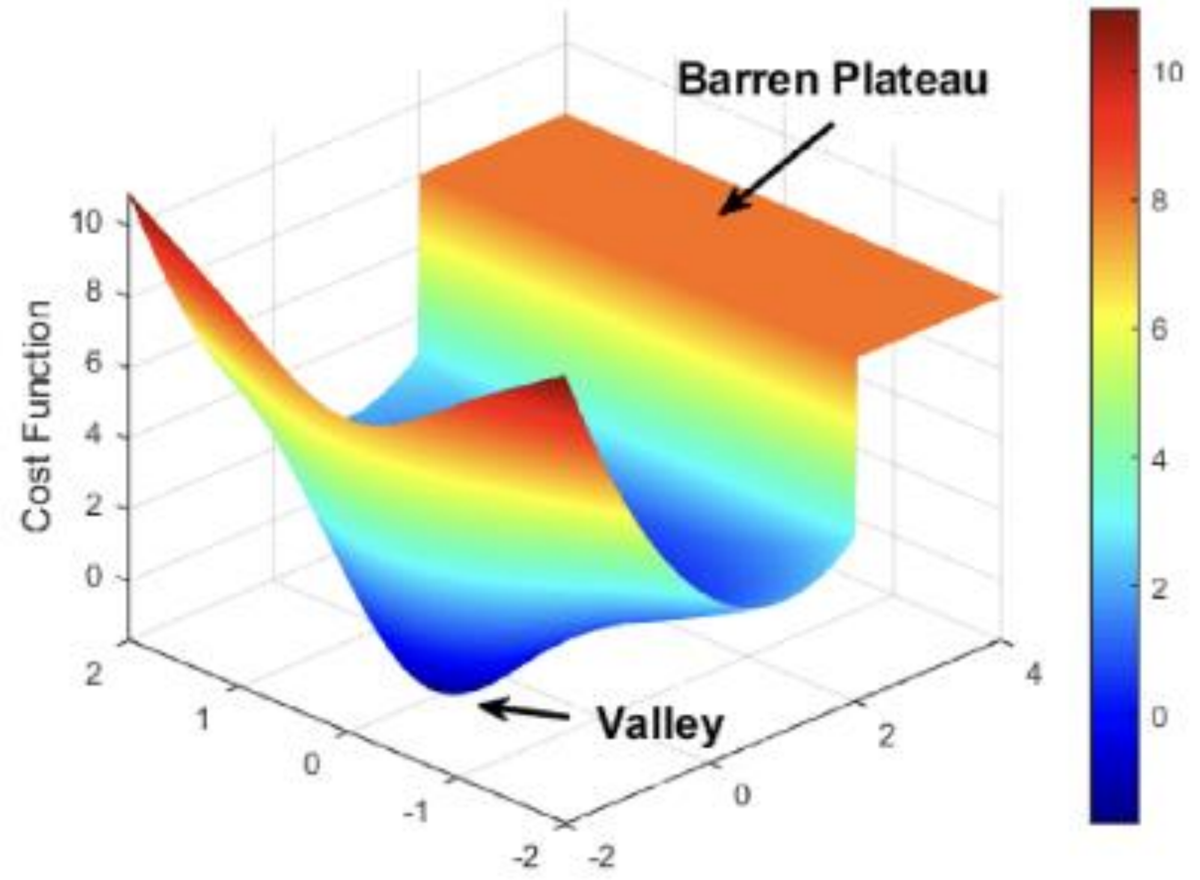


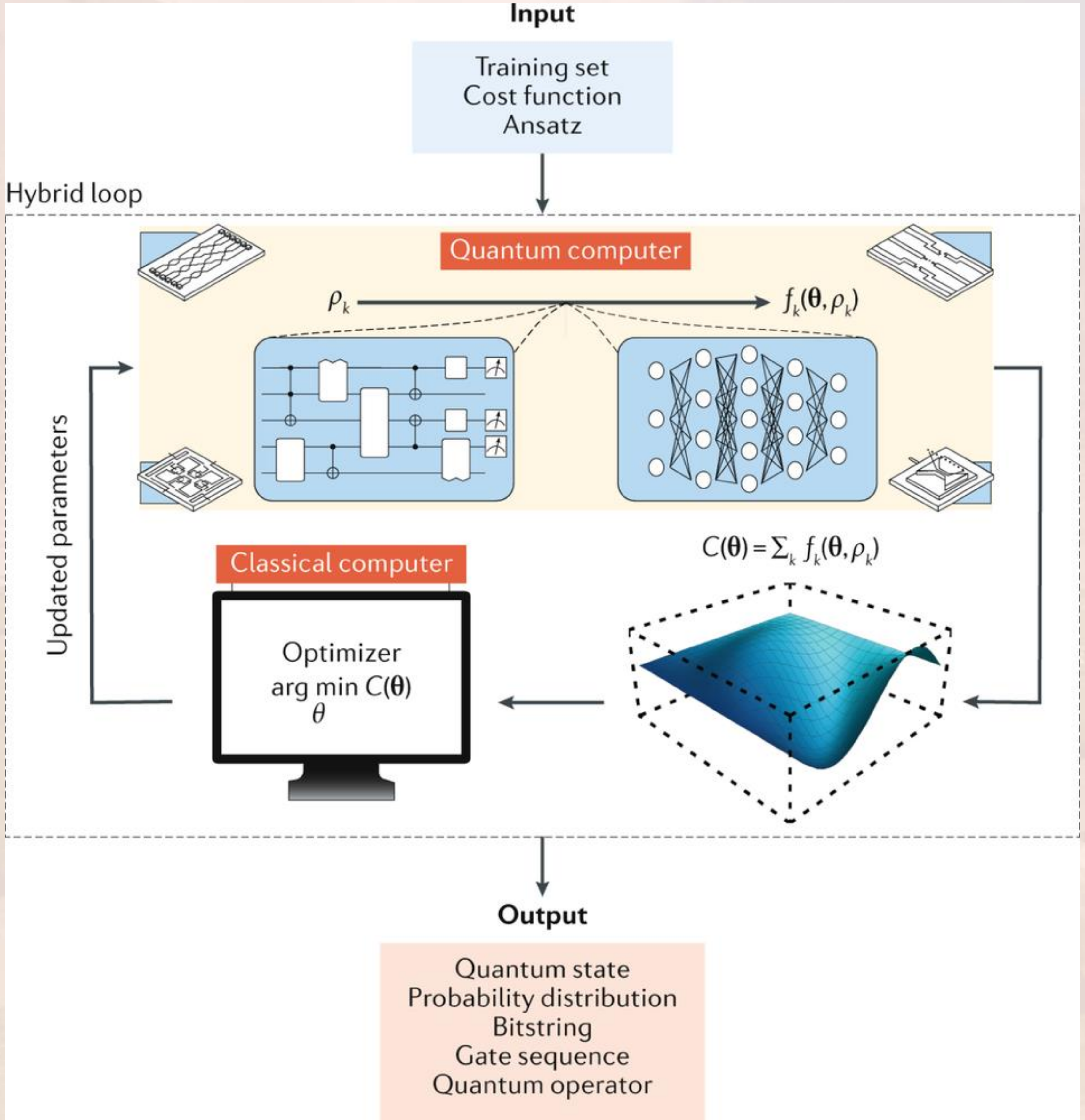
Figure 1. A cost landscape when BP exists. Parameters in the BP region have vanishing gradients. Only parameters in the exponentially suppressed central part can be optimized efficiently.



https://www.quair.group/software/pq/tutorials/qnn_research/barren plateaus_en

Huan-Yu Liu et al 2023 New J. Phys. 25 013039

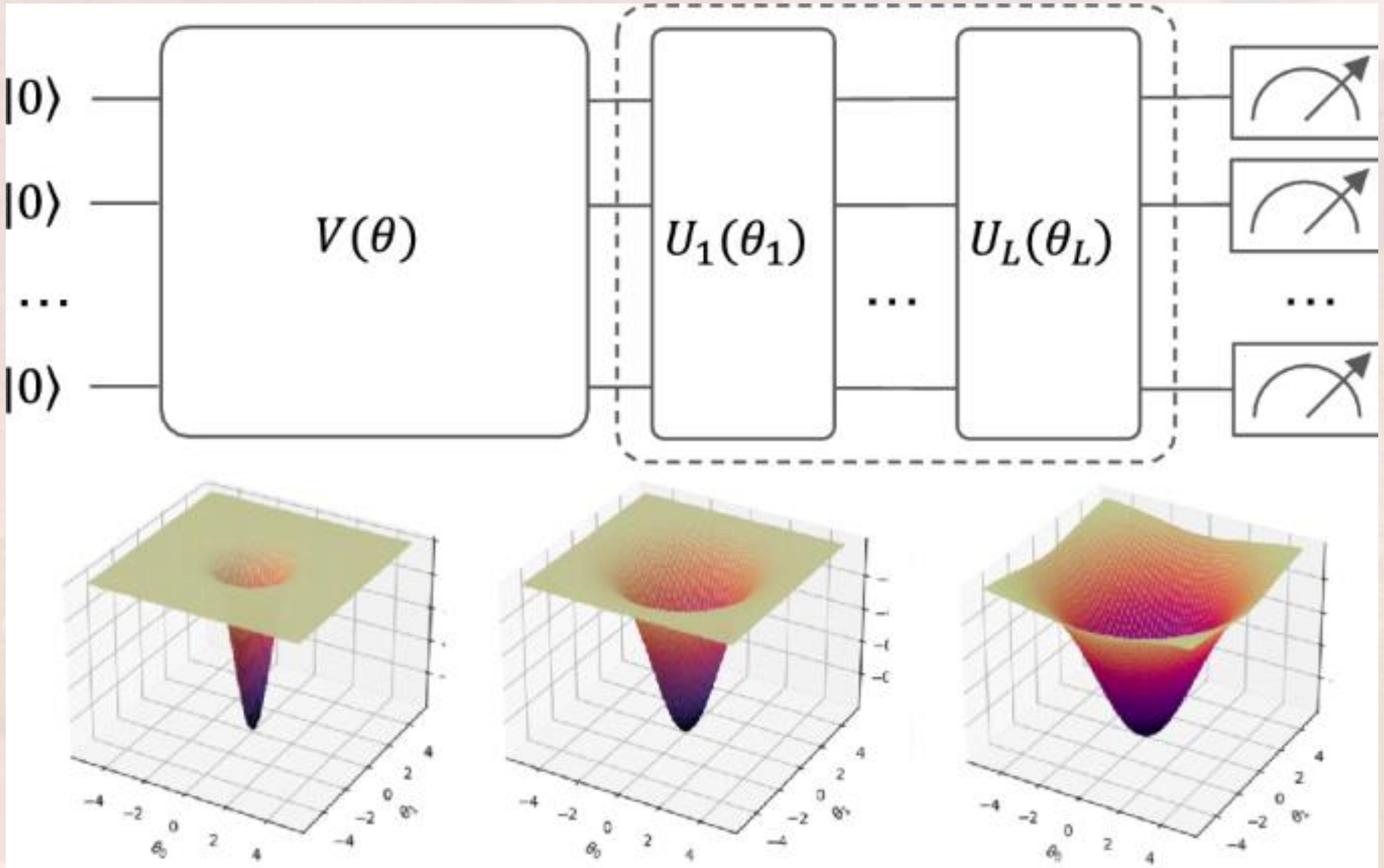
Variational Quantum Algorithms



Cerezo, M., Arrasmith, A., Babbush, R. *et al.* Variational quantum algorithms. *Nat Rev Phys* **3**, 625–644 (2021). https://doi.org/10.1038/s42254-021-00348-9

Reasons Make BARREN PLATEAUS exist in Variational Quantum Algorithms

- Random circuits approximate unitary 2-designs → gradients vanish.
- Use of global cost functions → exponential gradient suppression.
- Poor ansatz architecture or random initialization.
- Noise drives states toward maximally mixed → flat landscapes.



A general structure of VQCs where the optimization-based strategies are usually applied in the training of $U(\theta)$ (dotted box), which is decomposed into L layer from $U_1(\theta_1)$ to $U_L(\theta_L)$ (upper side). On the lower side, we present the process for addressing BPs. The left-most cost-function landscape represents BPs, a flat landscape with no discernible slope toward the minimum. In this case, many optimization approaches may be initially trapped in the flat landscape, leading to a failure of training. By employing optimization-based strategies, the cost-function landscape could be gradually recovered, as shown in the middle and right-most landscapes

Cunningham, J., Zhuang, J. Investigating and mitigating barren plateaus in variational quantum circuits: a survey. *Quantum Inf Process* **24**, 48 (2025). https://doi.org/10.1007/s11128-025-04665-1

THE RANDOMNESS (FIRST)

A barren plateau is a phenomenon in variational quantum algorithms where the gradient of the cost function vanishes exponentially with system size:

$$\mathrm{Var}\left[\frac{\partial C}{\partial \theta_i}\right] \sim \frac{1}{2^n} \quad \text{for n qubits.}$$

$$\mathrm{Var}\left[\partial_k C\right] = G(n)$$

$$G(n) \in \mathcal{O}(1/2^{2n})$$

This makes training variational quantum circuits (VQAs) very difficult because the optimization landscape is almost flat.

Paper: *McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 4812.*

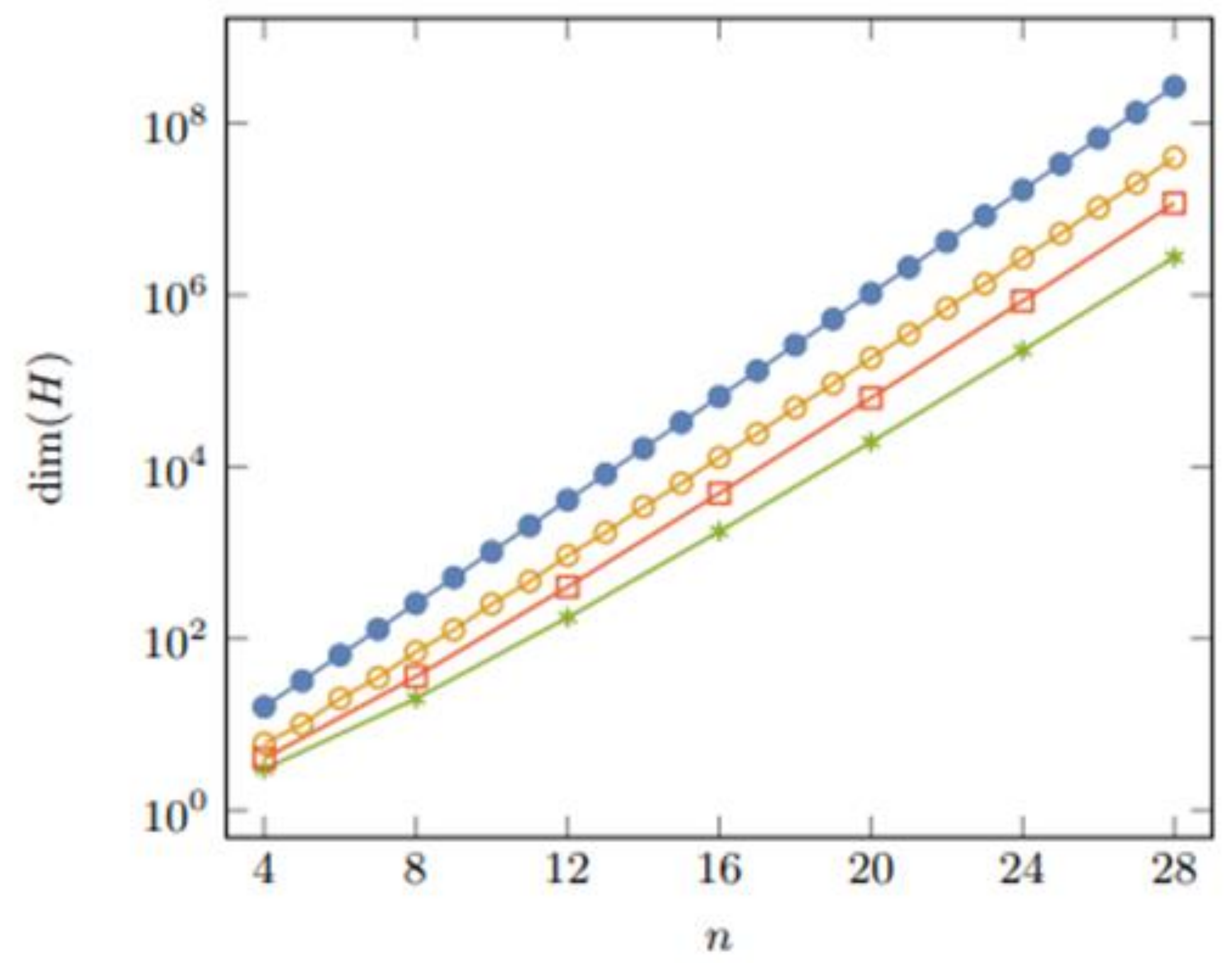


FIG. 7. (Color Online) Hilbert space dimension as a function of number of qubits, n, when relevant symmetries are enforced. We show the dimension of the full Hilbert space (blue), the largest particle-number subspace with $m = n/2$ (yellow), the subspace with $m = n/2$ and $s_z = 0$ (red), and the subspace with $m = n/2$ and $s = 0$ (green). Lines are included only as a guide. All cases remain an exponential scaling Hilbert space, where use of symmetries reduces the exponential factor.

Gard, B.T., Zhu, L., Barron, G.S. *et al.*. *npj Quantum Inf* **6**, 10 (2020).

Example 1

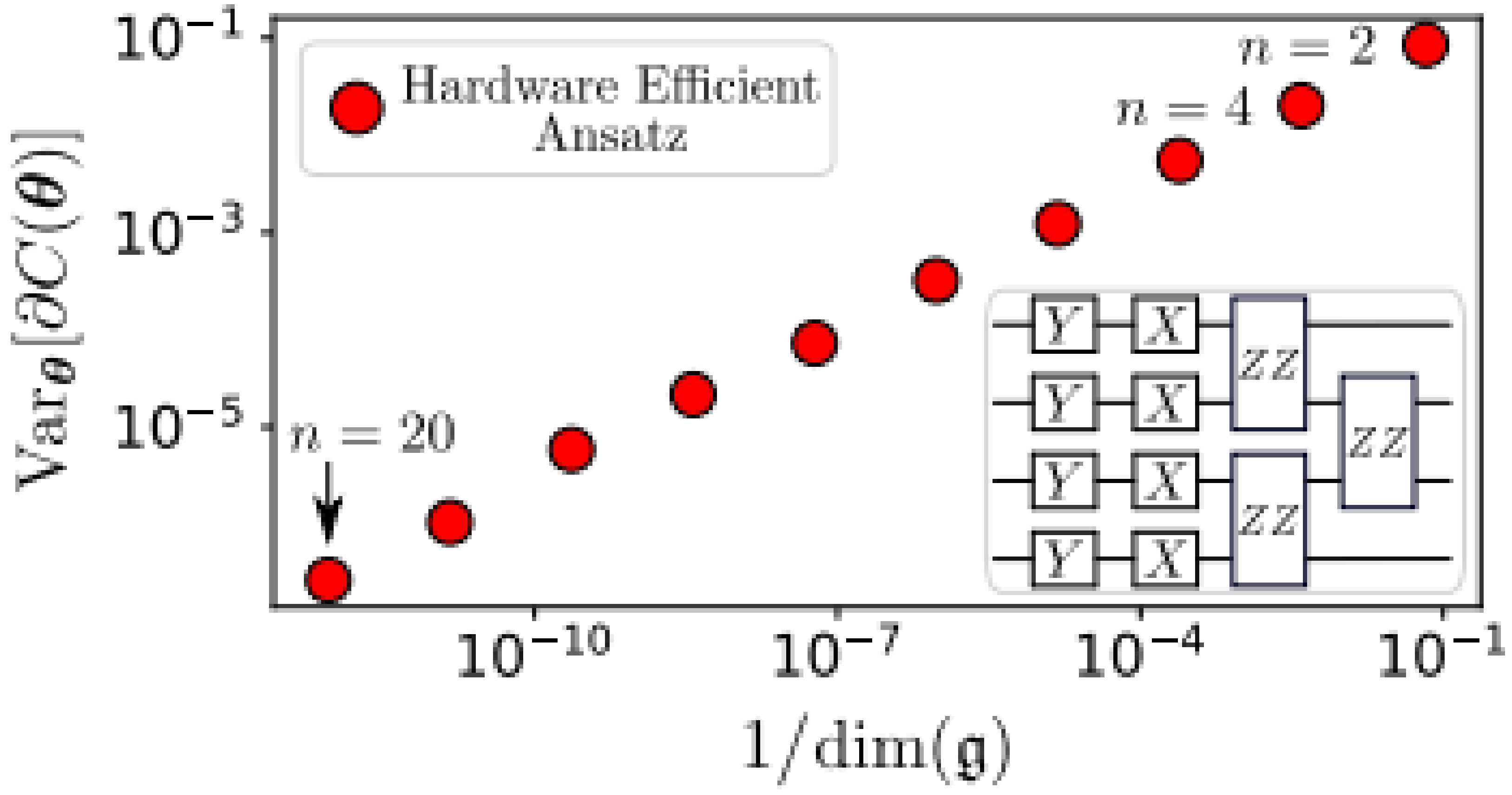


Figure 4: **Variance of cost function partial derivatives versus inverse of the DLA dimension for a controllable system.** The layered Hardware Efficient Ansatz (shown in the inset for $n = 4$) is a controllable system with generators given in Proposition 2. Then, as shown in Proposition 1, the cost function of Eq. (23) exhibits a barren plateau and hence $\mathrm{Var}_{\boldsymbol{\theta}}[\partial_\mu C(\boldsymbol{\theta})] \in \mathcal{O}(1/2^n)$. Moreover, since the system is controllable one finds that $\dim(\mathfrak{g}) = 4^n - 1$. Hence, as shown in the plot, Conjecture 1 holds for controllable systems, since the dependence of $\mathrm{Var}_{\boldsymbol{\theta}}[\partial_\mu C(\boldsymbol{\theta})]$ versus $1/\dim(\mathfrak{g})$ is linear on a log-log scale.

https://quantum-journal.org/papers/q-2022-09-29-824/

Example 2:

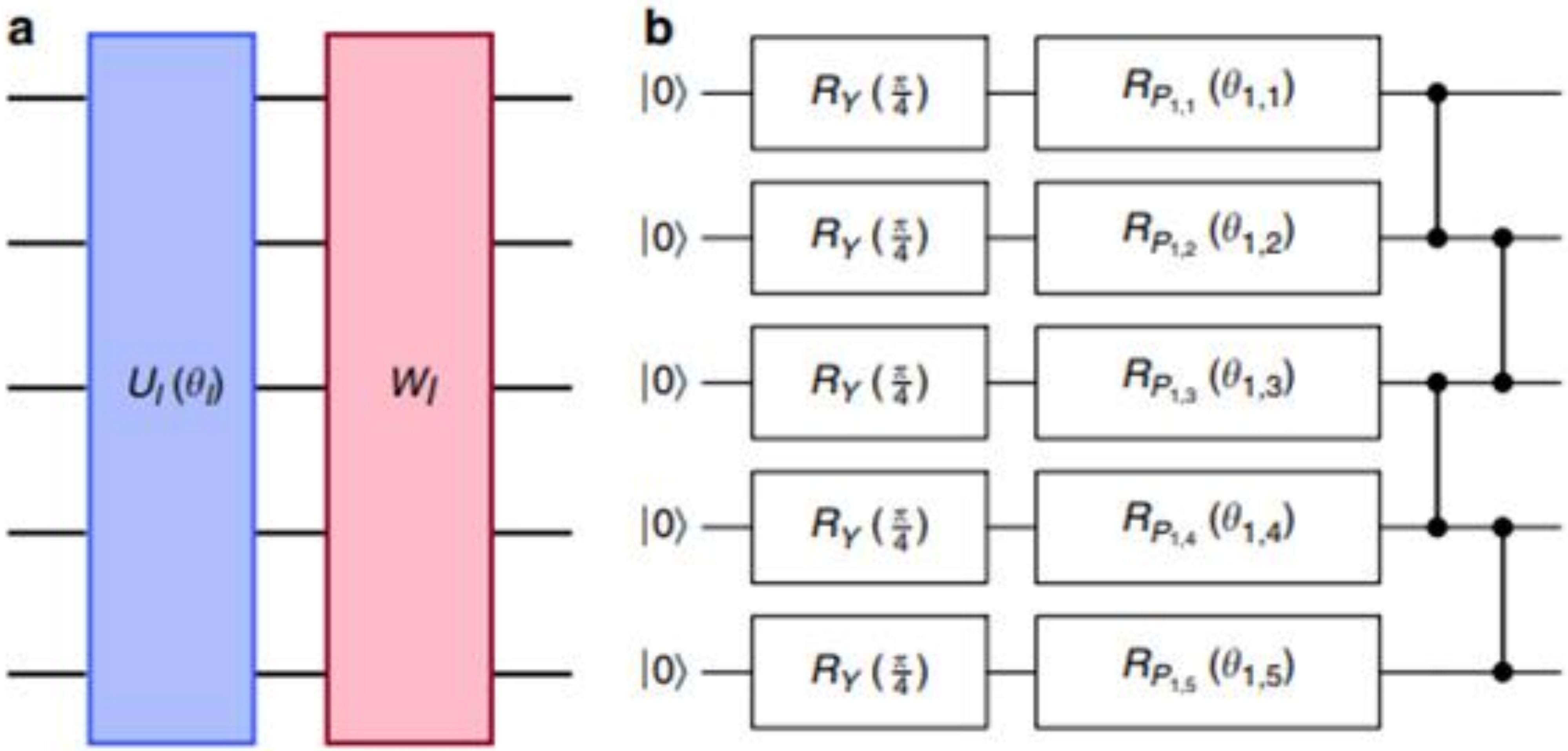


Fig. 2 Structure of quantum circuits. **a** The generic subunit of circuits we study in this work, with a parameterized component $U_l(\theta_l)$ and non-parameterized unit W_l for each layer l. **b** Example schematic of the 1D random circuits used in our numerical experiments. The circuit begins with $R_Y(\frac{\pi}{4})$ gates applied to all qubits followed by a specified number of layers of randomly chosen Pauli rotations applied to each qubit and then a 1D ladder of controlled Z gates. The initial $R_Y(\frac{\pi}{4})$ gates are not repeated in each layer. The indices i and j in $\theta_{i,j}$ index the layer and qubit, respectively. For each layer and qub it $P_{i,j} \in \{X, Y, Z\}$ and $\theta_{i,j} \in [0,2\pi)$ are sampled independently

: *McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 4812*

Example 2 Follow-up

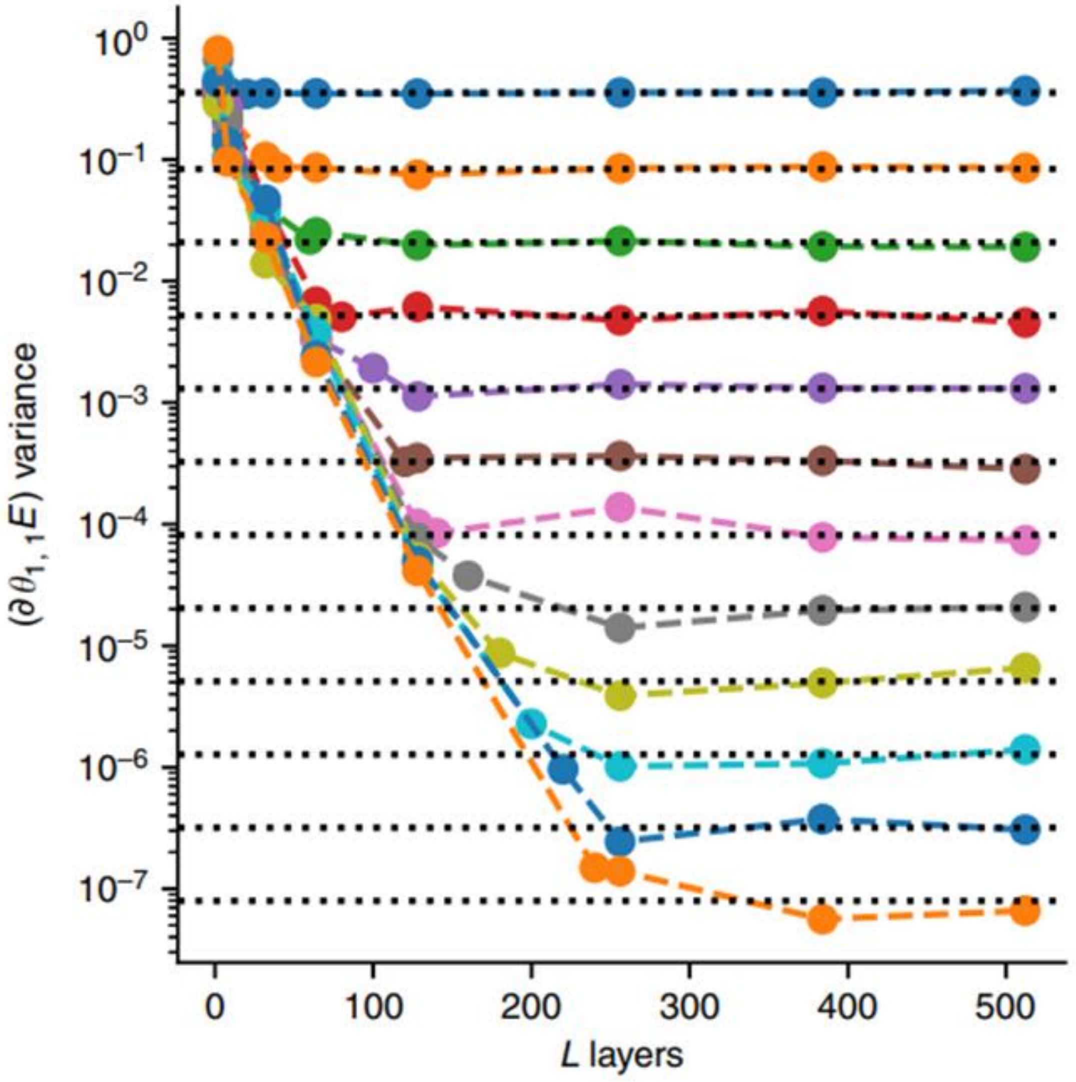


Fig. 4 Convergence to 2-design limit. Here we show the sample variance of the gradient of the energy for the first circuit component of a two-local Pauli term $\left(\partial_{\theta_{1,1}}E\right)$ plotted as a function of the number of layers, *L*, in a 1D quantum circuit. The different lines correspond to all even numbers of qubits between 2 and 24, with 2 qubits being the top line, and the rest being ordered by qubit number. The dotted black lines depict the 2-design asymptotes for this Hamiltonian as determined by our analytic results. This shows the convergence of the second moment as a function of the number of layers to a fixed value determined by the number of qubits

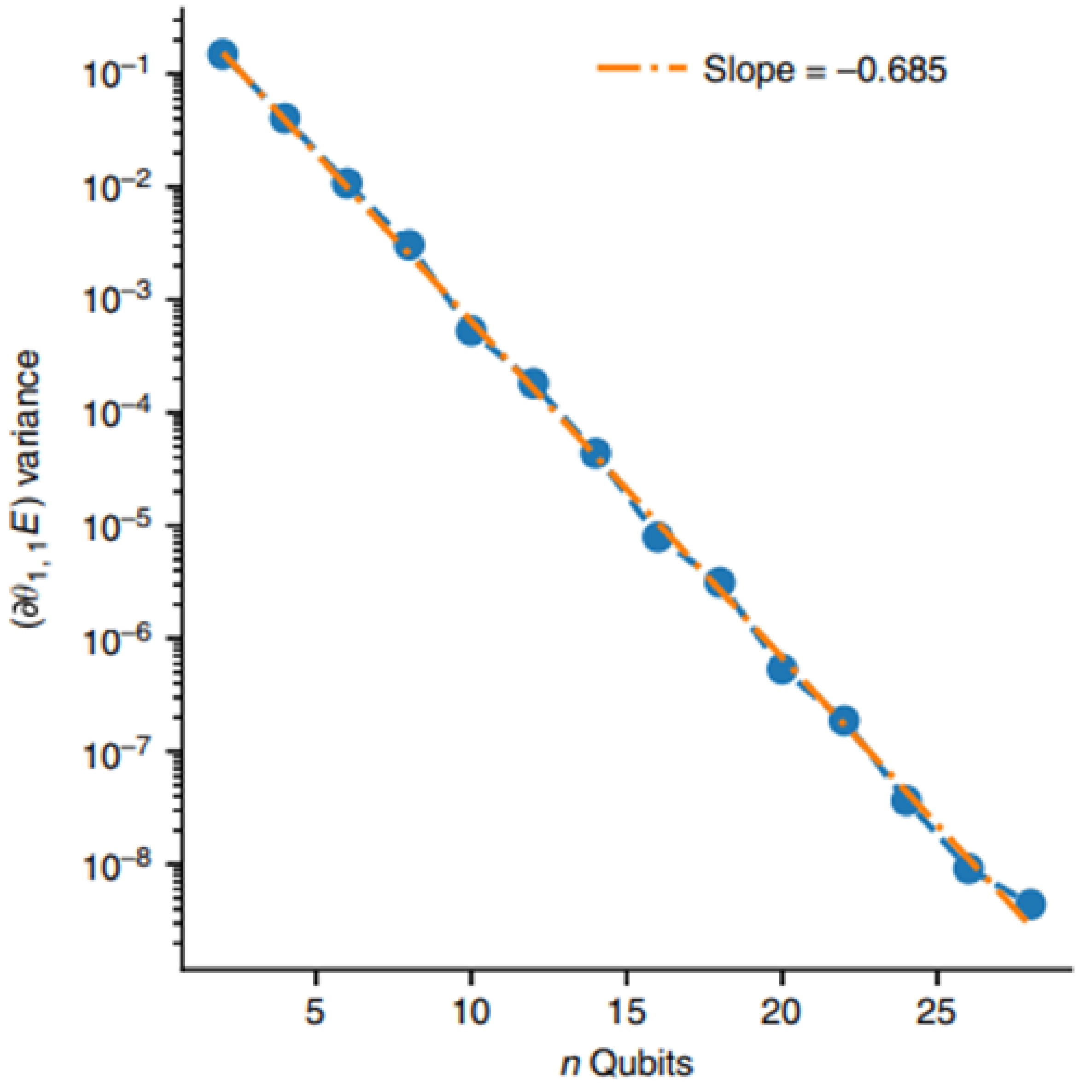


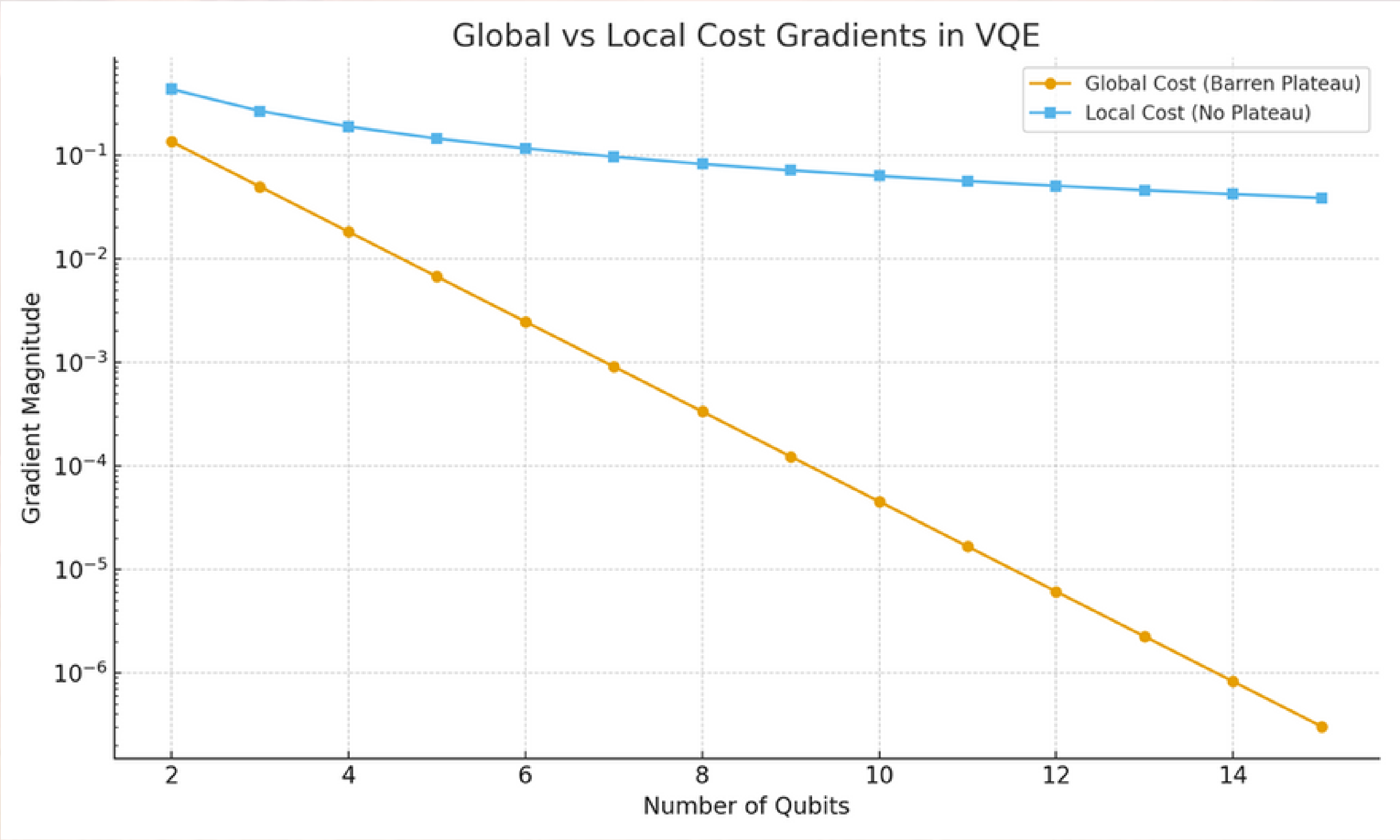
Fig. 3 Exponential decay of variance. The sample variance of the gradient of the energy for the first circuit component of a two-local Pauli term $\left(\partial_{\theta_{1,1}}E\right)$ plotted as a function of the number of qubits on a semi-log plot. As predicted, an exponential decay is observed as a function of the number of qubits, *n*, for both the expected value and its spread. The slope of the fit line is indicative of the rate of exponential decay as determined by the operator

: *McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 4812*

THE GLOBAL COST FUNCTION (SECOND)

The Global Cost function

For global cost functions, variance of gradient decreases exponentially with system size → hard to train



Intuition:

- A random circuit "scrambles" the state across the entire Hilbert space.
- A global operator is like taking the "average" over all qubit configurations.
- Gradients tend to cancel out, leaving nearly zero signal → Barren Plateau.
- Local cost functions only depend on a subset of qubits, so the gradient variance decreases polynomially instead of exponentially.

Example

$\mathrm{Var}[\partial_\nu C]$ exponentially vanishes for

$$L \in \mathcal{O}(\mathrm{poly}(\log(n)))$$

$\mathrm{Var}[\partial_\nu C]$ at worst polynomially decreases for

$$L \in \mathcal{O}(\log(n))$$

Example Figures

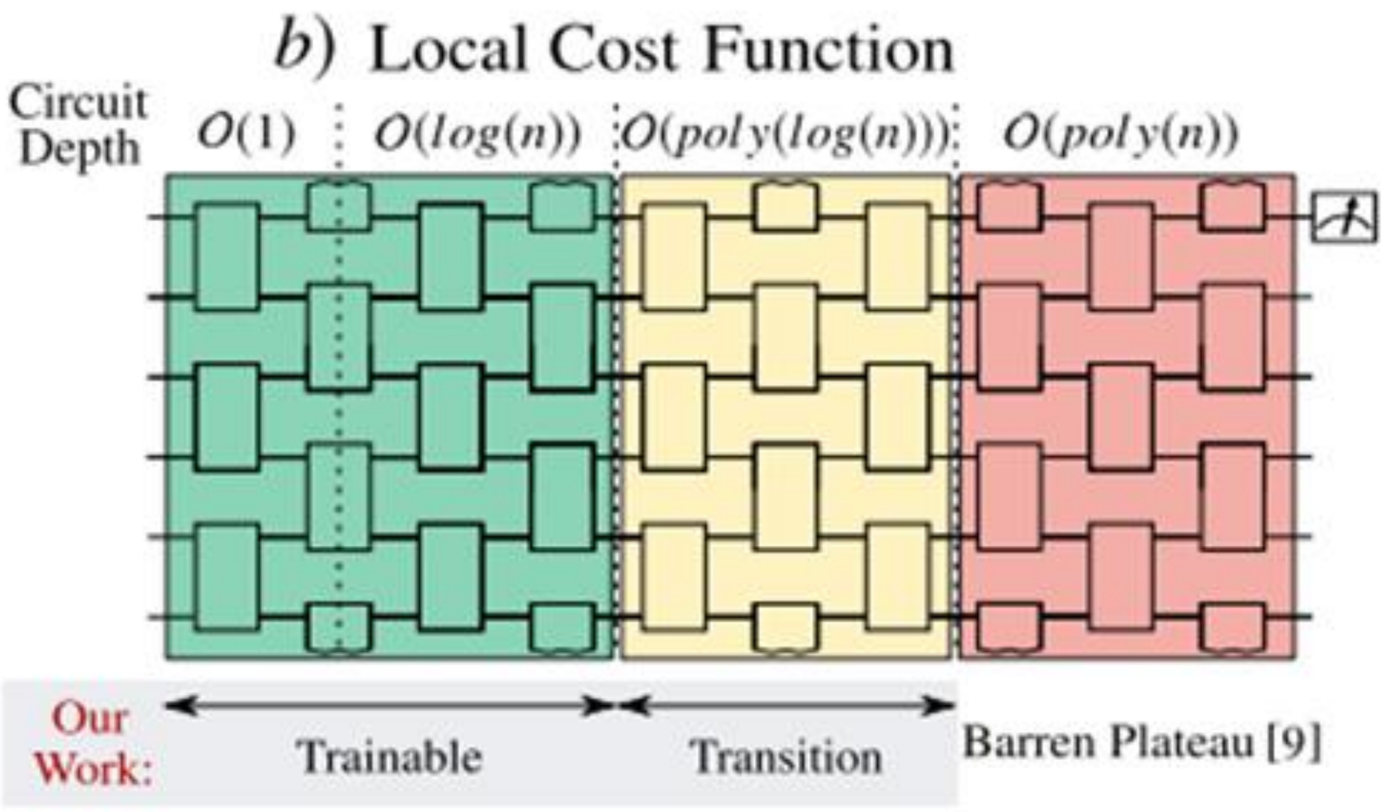


FIG. 1. Summary of our main results. McClean *et al.* [9]

[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, 4812 (2018)

Cerezo, M., Sone, A., Volkoff, T. *et al.* Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat Commun* **12**, 1791 (2021).

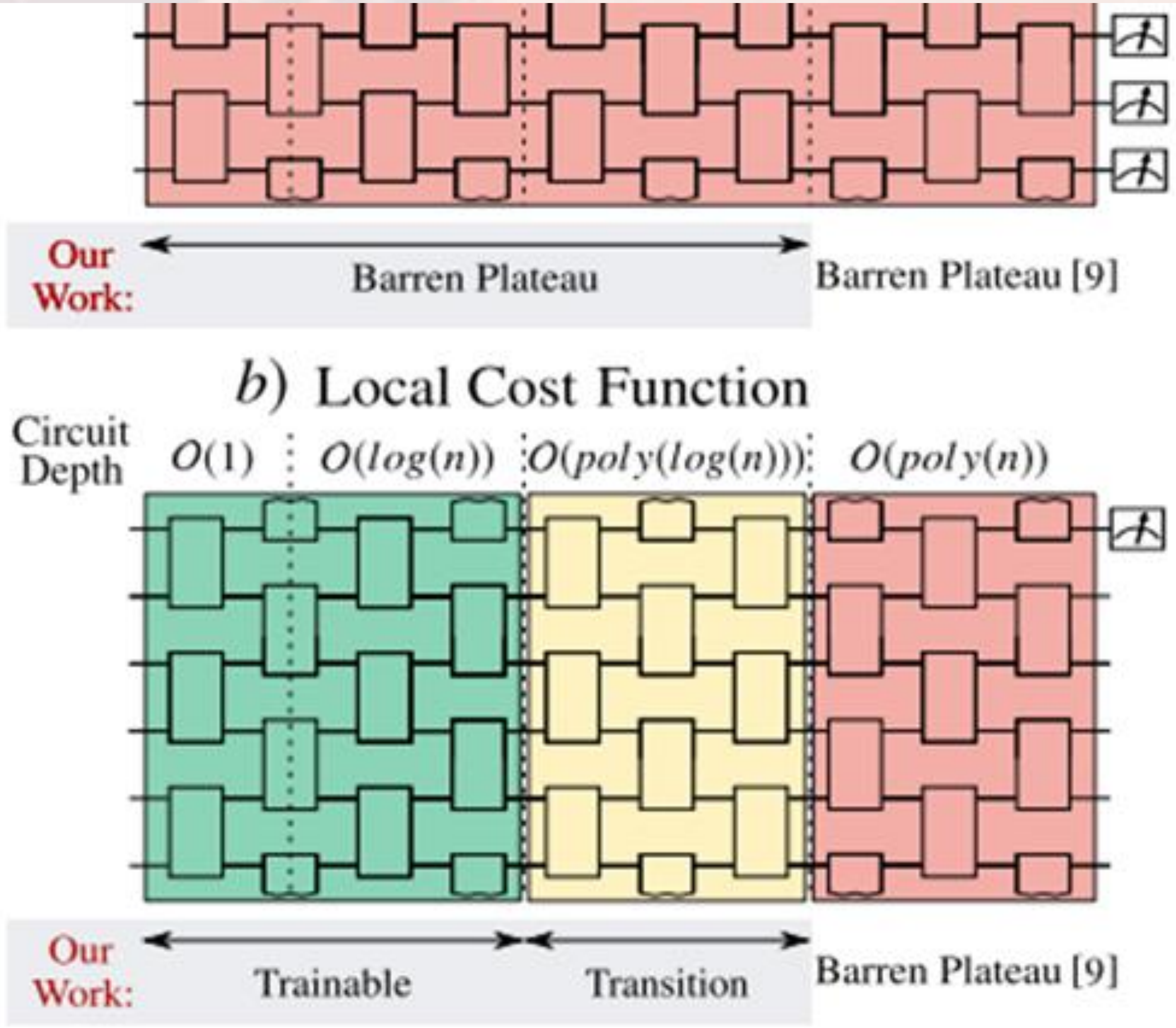
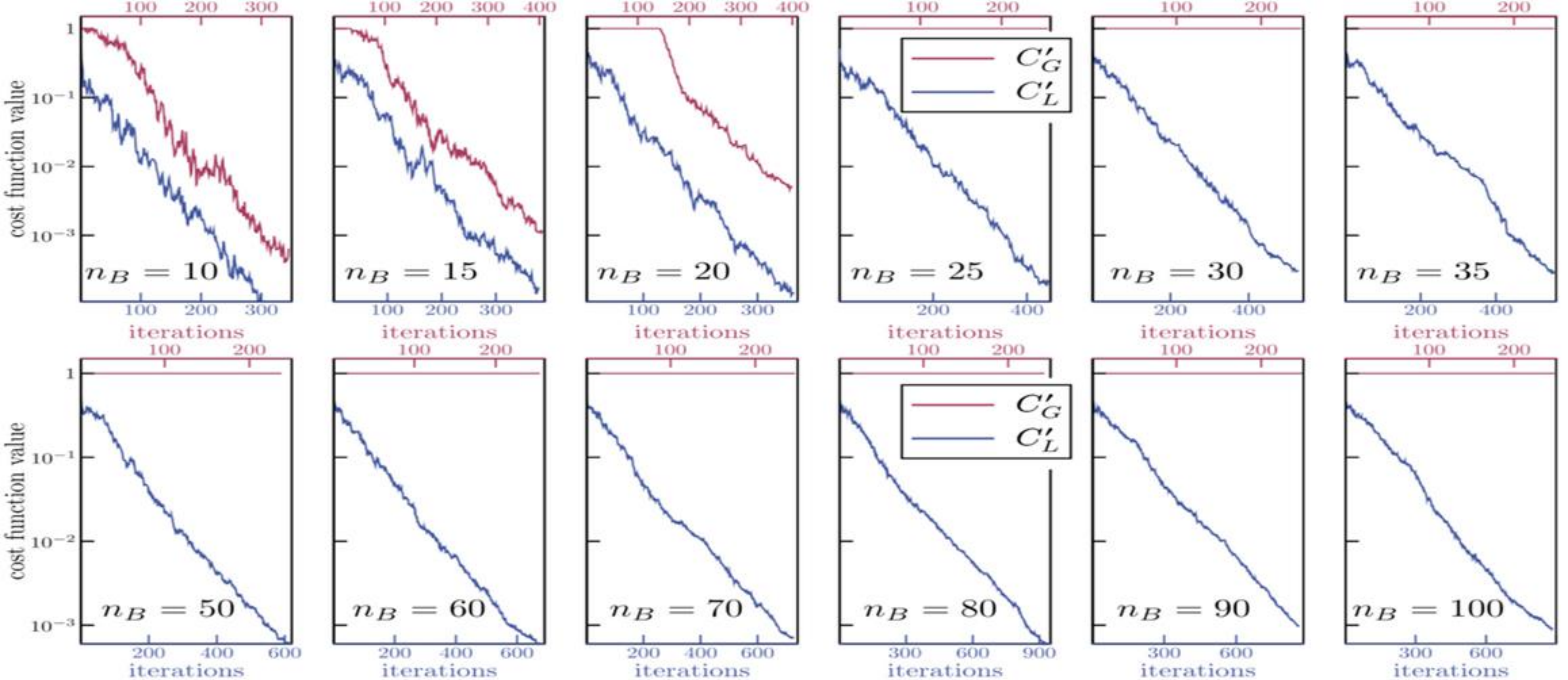


FIG. 1. Summary of our main results. McClean *et al.* [9] proved that a barren plateau can occur when the depth D of a hardware-efficient ansatz is $D \in \mathcal{O}(\mathrm{poly}(n))$. Here we extend these results by providing bounds for the variance of the gradient of global and local cost functions as a function of D. In particular, we find that the barren plateau phenomenon is



Numerical simualtions: Quantum autoencoders for efficient compression of quantum data Romero, J., Olson, J. P., & Aspuru-Guzik, A. (2017). Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, *2*(4), 045001.

LOCALITY BUT NO GOOD INITIALIZATION!

Locality but no good initialization

Results

Park, C. Y., & Killoran, N. (2024). Hamiltonian variational ansatz without barren plateaus. *Quantum*, *8*, 1239.

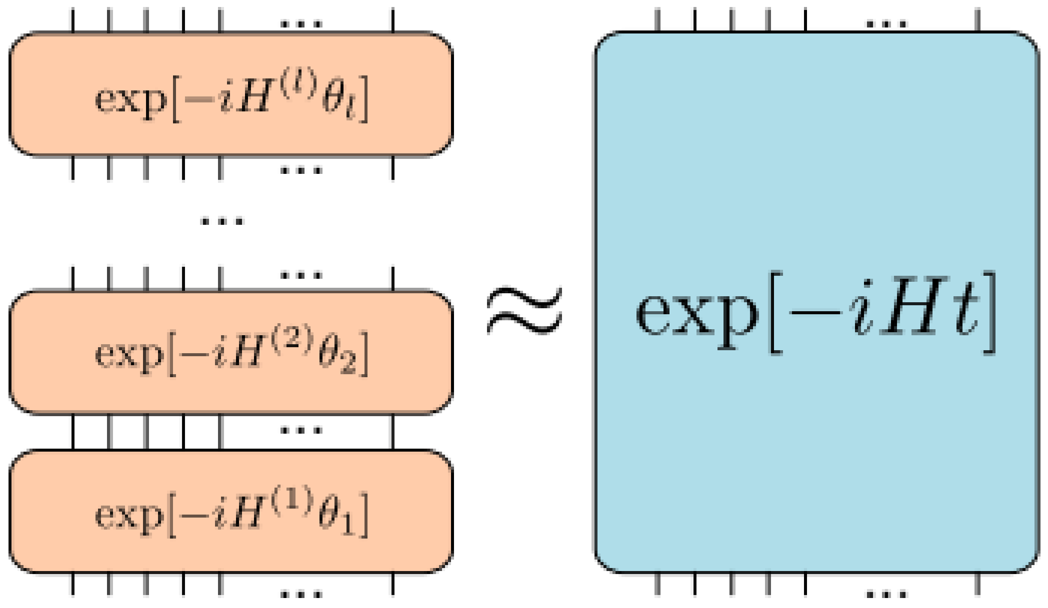


Figure 1: We find a parameter constraint such that layers of Hamiltonian evolution in the HVA (left) approximate to the time evolution under a single local Hamiltonian (right). Using the dynamical properties of local Hamiltonians, we argue that the HVA has large gradients.

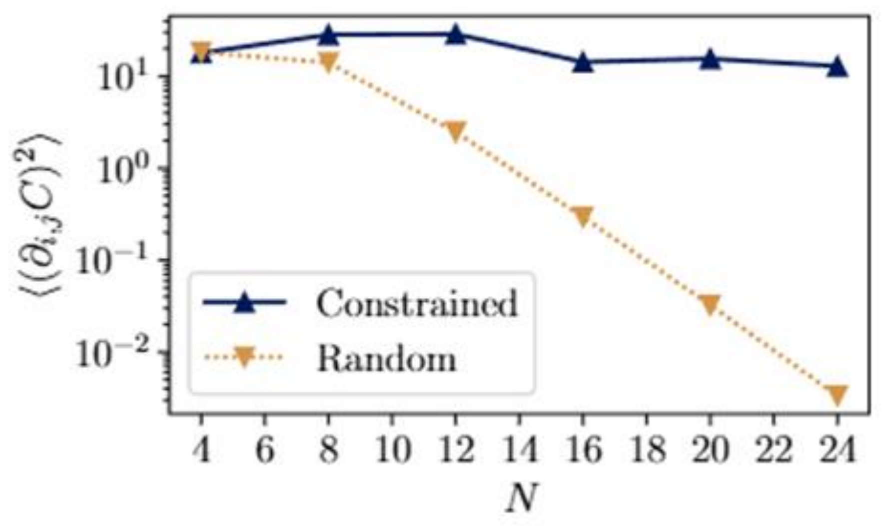
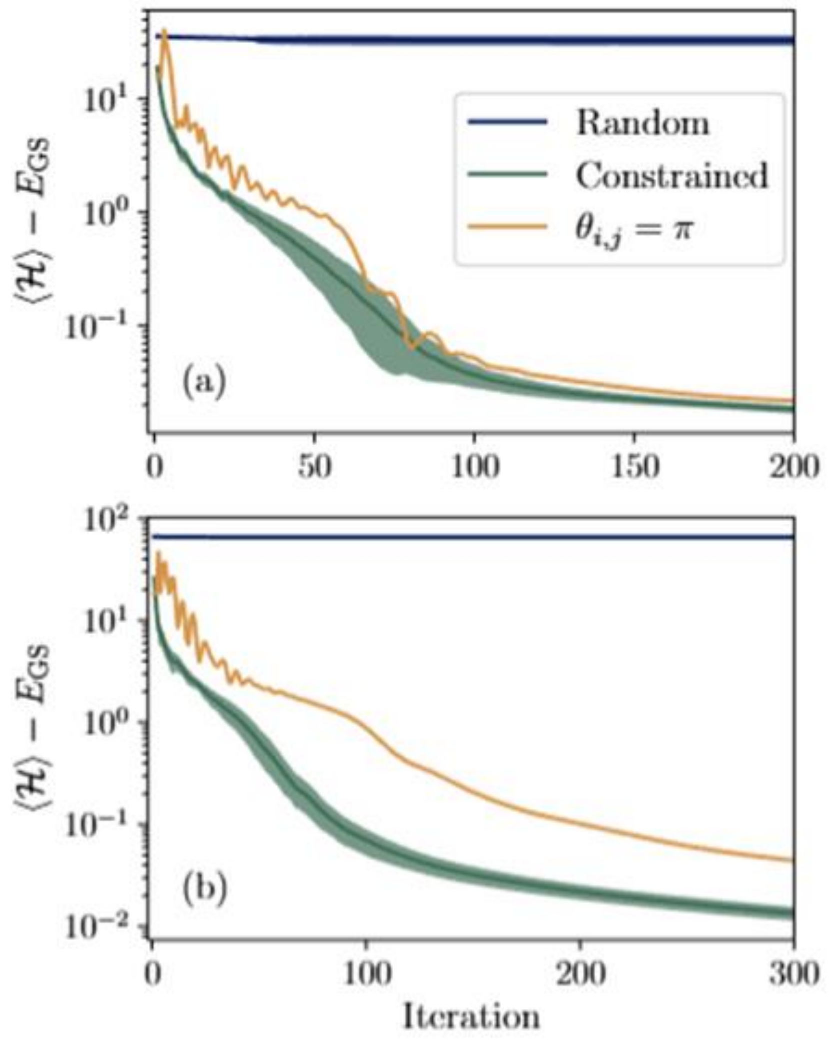


Figure 7: Scaling of the averaged squared gradients $(\partial_{i,j}C)^2$ from the repeated ansatz Eq. (41) with (solid) and without (dotted) a parameter constraint. For the parameter-constrained ansatz, we sample parameters under the constraint $\theta_{i,1} + \theta_{i,2} + \theta_{i,3} = \pi/(2N)$. In contrast, $\theta_{i,j} \sim \mathcal{U}_{[0,2\pi]}$ is used for the ansatz without the constraint. The HVA is for the 1D XYZ model with $O = Y_0Y_1$ with $\tilde{p} = 16$ and $r = N^2/4$. The results are averaged over 2^{10} random parameters.



NOISE CONDITION?

28 Jul 2020

Noise-Induced Barren Plateaus in Variational Quantum Algorithms

Samson Wang,[1,2] Enrico Fontana,[1,3,4] M. Cerezo,[1,5] Kunal Sharma,[1,6]
Akira Sone,[1,5] Lukasz Cincio,[1] and Patrick J. Coles[1]

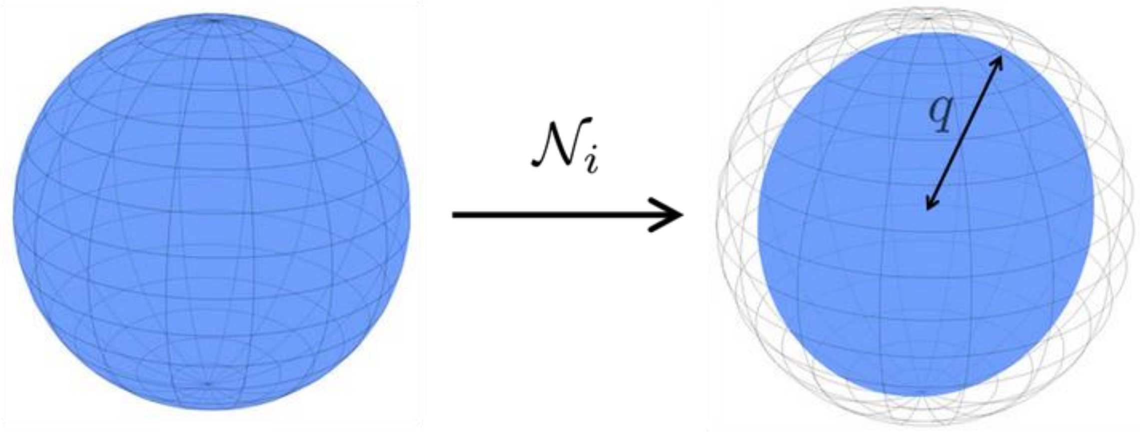
[1] *Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2] *Imperial College London, London, UK*
[3] *University of Strathclyde, Glasgow, UK*
[4] *National Physical Laboratory, Teddington, UK*
[5] *Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM, USA*
[6] *Hearne Institute for Theoretical Physics and Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA USA*

Variational Quantum Algorithms (VQAs) may be a path to quantum advantage on Noisy Intermediate-Scale Quantum (NISQ) computers. A natural question is whether the noise on NISQ devices places any fundamental limitations on the performance of VQAs. In this work, we rigorously prove a serious limitation for noisy VQAs, in that the noise causes the training landscape to have a barren plateau (i.e., vanishing gradient). Specifically, for the local Pauli noise considered, we prove that the gradient vanishes exponentially in the number of layers L. This implies exponential decay in the number of qubits n when L scales as $\text{poly}(n)$, for sufficiently large coefficients in the polynomial. These noise-induced barren plateaus (NIBPs) are conceptually different from noise-free barren plateaus, which are linked to random parameter initialization. Our result is formulated for an abstract ansatz that includes as special cases the Quantum Alternating Operator Ansatz (QAOA) and the Unitary Coupled Cluster Ansatz, among others. In the case of the QAOA, we implement numerical heuristics that confirm the NIBP phenomenon for a realistic hardware noise model.

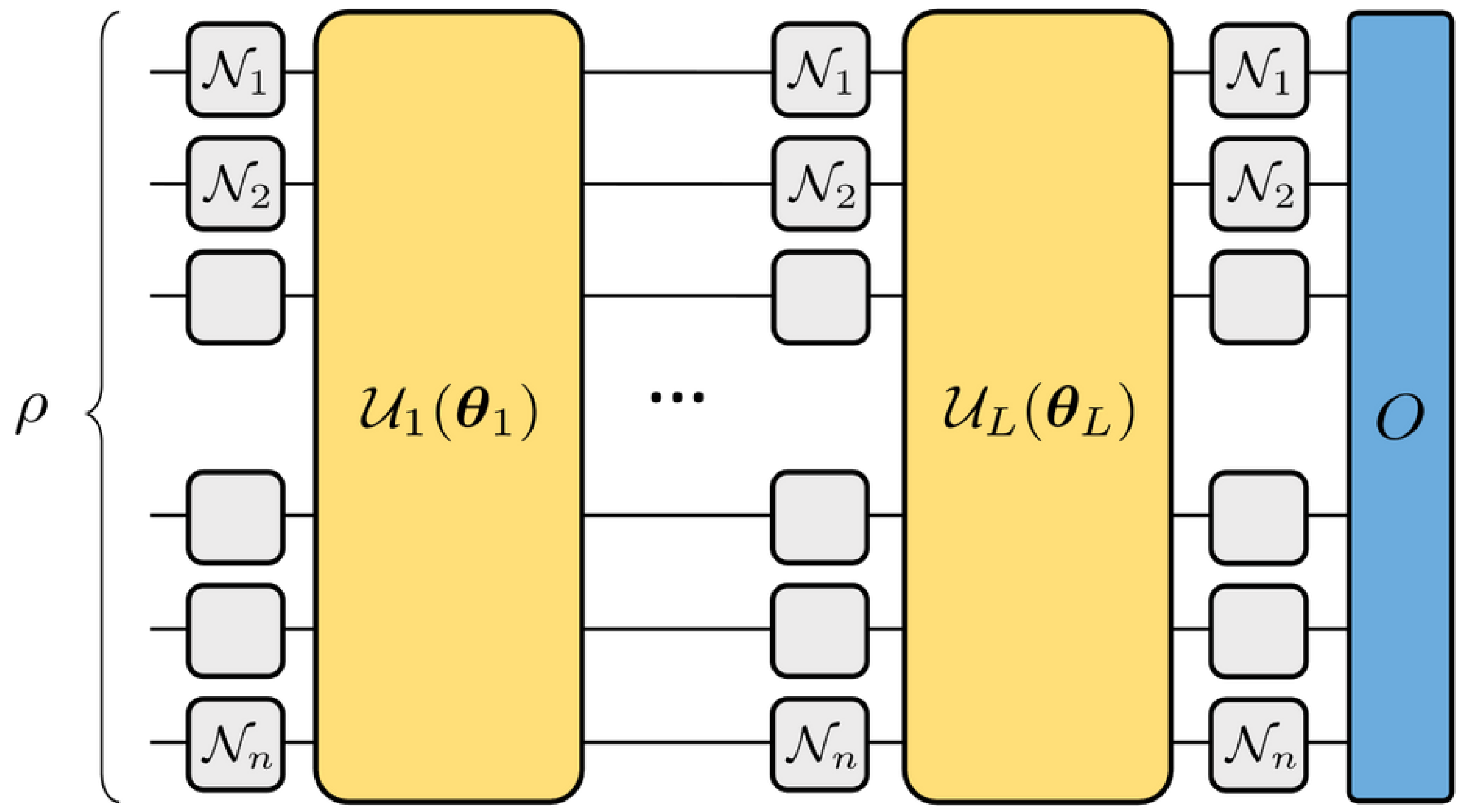
Noise model:

Noise-Induced Barren Plateaus

- local noise $\mathcal{N}_i$ acts on qubit i with action on Bloch sphere:



- refer to q as "noise parameter"



Ansatz: $U_l\left(\boldsymbol{\theta}_l\right)=\prod_m e^{-i \theta_{lm} H_{lm}} W_{lm}$

Noisy cost function: $\widetilde{C}=\operatorname{Tr}\left[O\left(\mathcal{N} \circ \mathcal{U}_L \circ \cdots \circ \mathcal{N} \circ \mathcal{U}_1 \circ \mathcal{N}\right)(\rho)\right]$

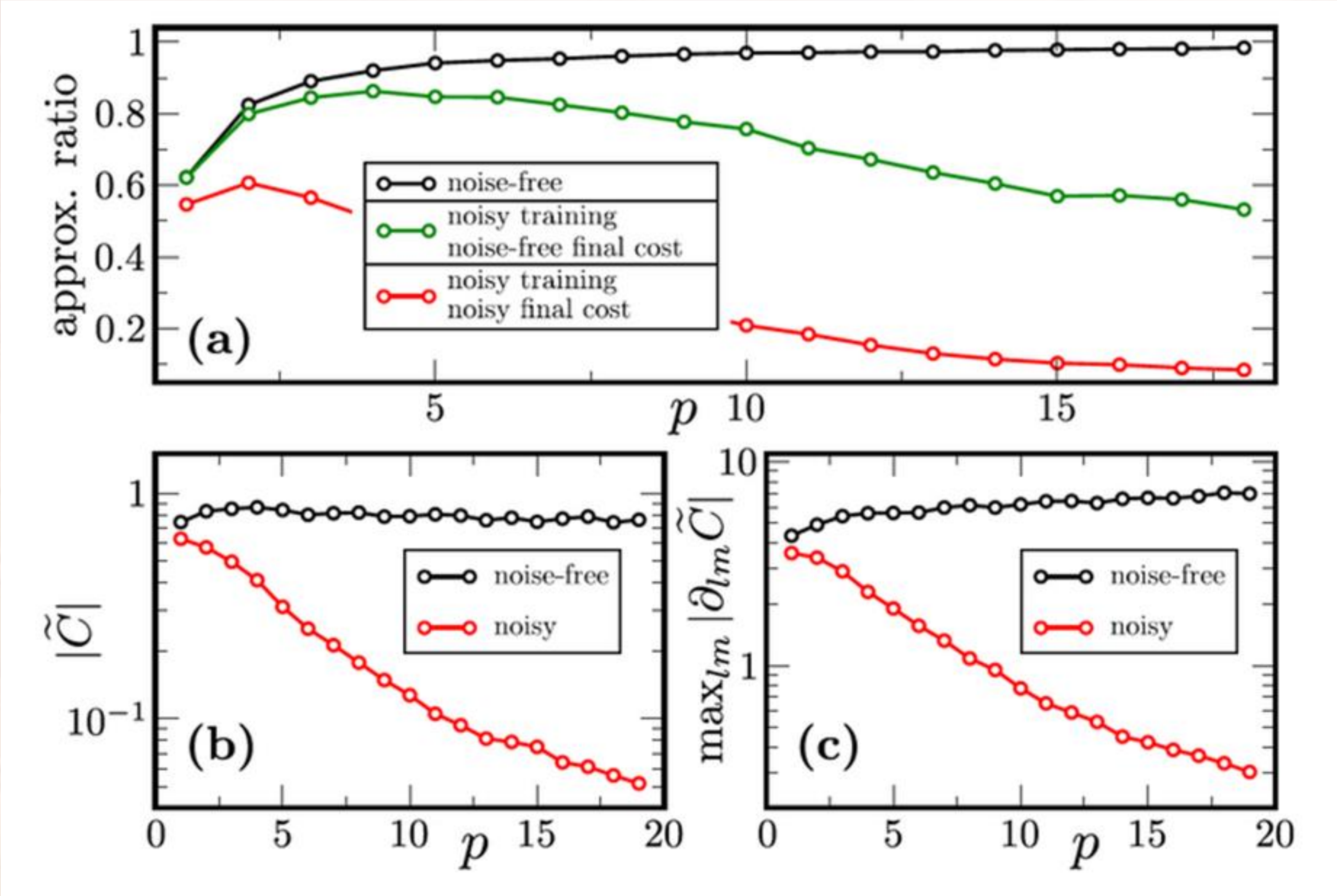
If condition is satisfied, cost function exponentially concentrates and the gradient exponentially vanishes

$$\left|\widetilde{C}-\frac{1}{2^n} \operatorname{Tr}[O]\right| \leqslant G(n)$$

Numerics Example: QAOA

MaxCut problems Noise model based on gate-set tomography on IBM Ourense

QAOA heuristics in the presence of realistic hardware noise: increasing number of rounds for fixed problem size



Wang, S., Fontana, E., Cerezo, M., Sharma, K., Sone, A., Cincio, L., & Coles, P. J. (2021). Noise-induced barren plateaus in variational quantum algorithms. *Nature communications*, *12*(1), 6961.

A UNIFIED THEORY THAT HELPS TO DETECT BARREN PLATEAUS ?

Article

https://doi.org/10.1038/s41467-024-49909-3

A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits

Received: 11 October 2023

Accepted: 25 June 2024

Published online: 22 August 2024

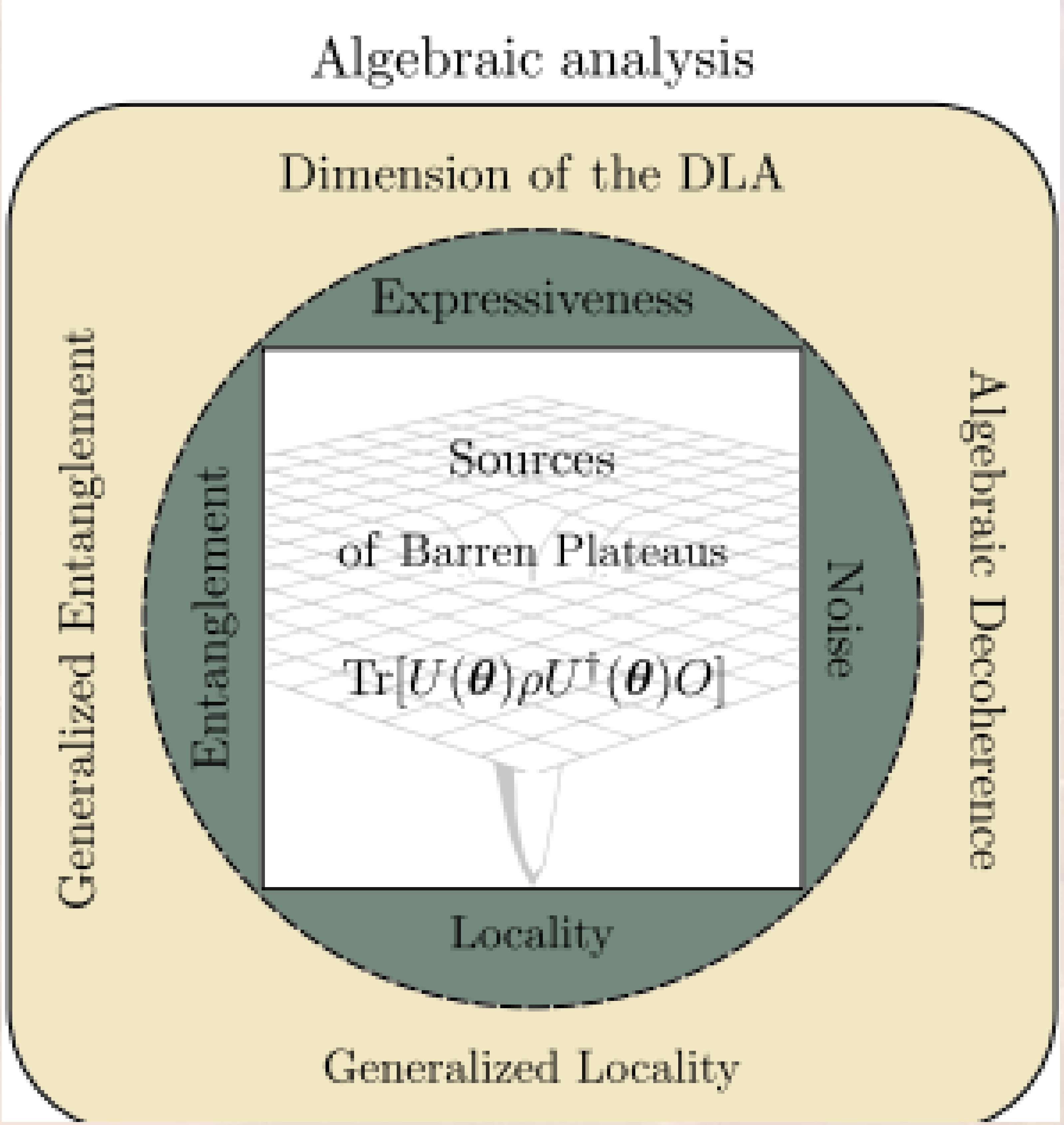
Check for updates

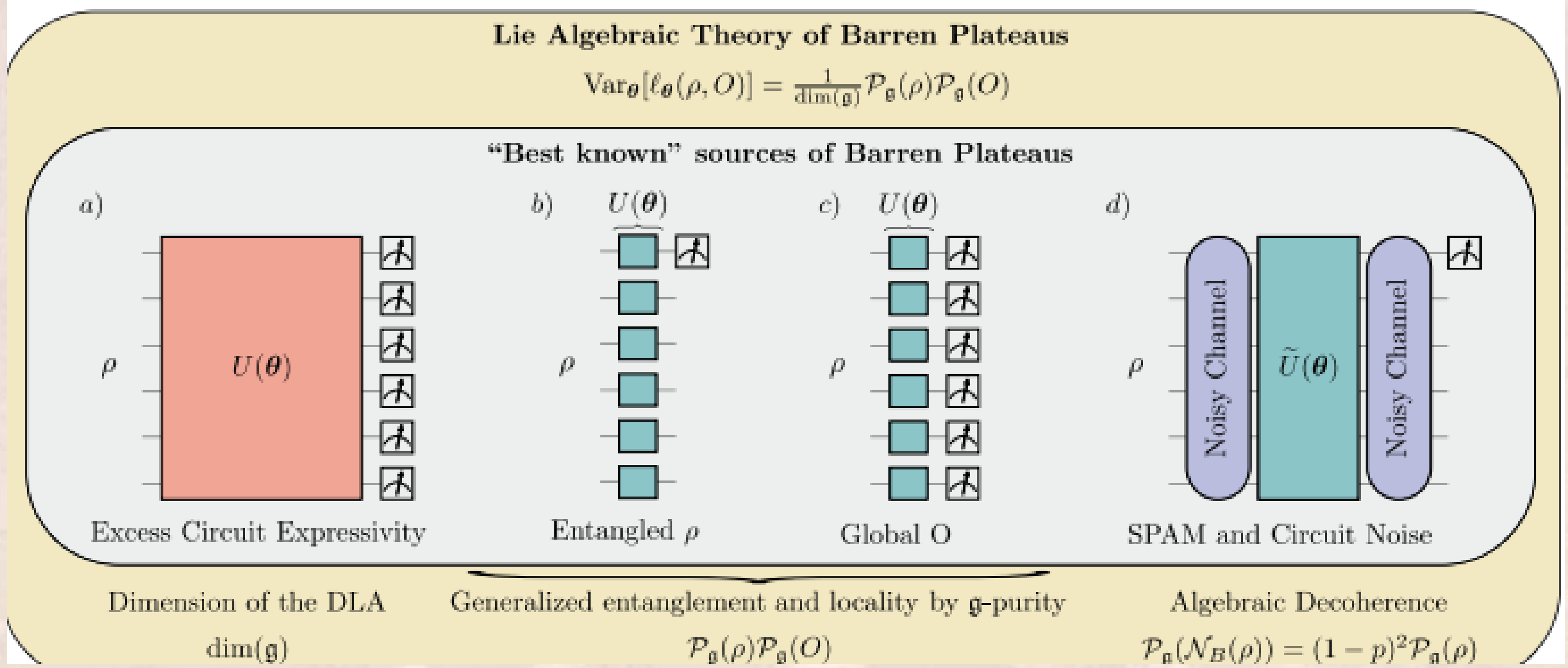
Michael Ragone[1,9], Bojko N. Bakalov[2,9], Frédéric Sauvage[3], Alexander F. Kemper[4], Carlos Ortiz Marrero[5,6], Martín Larocca[3,7] & M. Cerezo[8]

Variational quantum computing schemes train a loss function by sending an initial state through a parametrized quantum circuit, and measuring the expectation value of some operator. Despite their promise, the trainability of these algorithms is hindered by barren plateaus (BPs) induced by the expressiveness of the circuit, the entanglement of the input data, the locality of the observable, or the presence of noise. Up to this point, these sources of BPs have been regarded as independent. In this work, we present a general Lie algebraic theory that provides an exact expression for the variance of the loss function of sufficiently deep parametrized quantum circuits, even in the presence of certain noise models. Our results allow us to understand under one framework all aforementioned sources of BPs. This theoretical leap resolves a standing conjecture about a connection between loss concentration and the dimension of the Lie algebra of the circuit's generators.

Yes! Thanks to

Ragone, M., Bakalov, B.N., Sauvage, F. *et al.* A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits. *Nat Commun* **15**, 7172 (2024).





Ragone, M., Bakalov, B.N., Sauvage, F. *et al.* A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits. *Nat Commun* **15**, 7172 (2024).

WHAT IS HAPPENING FOR BARREN PLATEAUS IN QUANTUM COMPUTING FOR QUANTUM CHEMISTRY?

Solutions

2️⃣ Gard et al. (2020–2021) Key Idea

Gard et al. suggested that when a variational ansatz aligns with the symmetries of the Hamiltonian, it limits the state space, preventing the circuit from fully exploring the entire Hilbert space.

- Full Hilbert space → results in barren plateaus due to 2-design behavior.
- Symmetry-restricted space → creates a smaller effective Hilbert space → gradients do not vanish as quickly.

In simpler terms:

Symmetry ⟹ Reduced effective dimension ⟹ No barren plateau (or mitigated).

The Cost Function for LiH is More Global

For H_2, the Hamiltonian is brief and primarily 2-local.

In contrast, for LiH, the Hamiltonian contains numerous Pauli terms, such as:

- $Z \otimes Z$
- $Y \otimes X \otimes Z$
- $X \otimes Y \otimes I \otimes Z \otimes Z$, among others.

A random ansatz typically mixes all qubits, resulting in a nearly uniform state. This causes all expectation values to approximate zero, and the gradients to do the same.

Methods using Local Cost in Chemistry

Here are real techniques used today:

✔ 1. ADAPT-VQE

Adds one operator at a time using local gradients → no BP.

✔ 2. Qubit-Excitation Pool (QEB)

Lie algebra–based excitations acting on few qubits only.

✔ 3. Fragmentation / DMET-VQE

Compute energy of fragments → local cost only.

✔ 4. Operator grouping

Use qubit-wise commuting (QWC) groups → local clusters.

✔ 5. Local Unitary Coupled Cluster (tUCC)

Restrict excitation operators to local orbital neighborhoods.

Symmetry Restriction and Barren Plateaus in Variational Quantum Algorithms

Introduction

Barren plateaus: regions with vanishing gradients in variational quantum algorithms Training becomes difficult due to exponentially small gradients

Symmetry in Hamiltonians

- Hamiltonian H may have a symmetry group G
- Example: particle number conservation, spin parity
- Mathematically: $[H, \hat{O}] = 0, \hat{O} \in G$

Symmetry-Preserving Ansatz

- Construct $U(\theta)$ such that $[U(\theta), \hat{O}] = 0$
- Ensures states remain in symmetric subspace
- Dimension $d_{sym} \ll 2^n$

Gradient Scaling

- Gradient variance scales as $1/d_{sym}$ instead of $1/2^n$
- $\mathrm{Var}[\partial C/\partial\theta_i] \sim 1/d_{sym} \gg 1/2^n$ for large n

Concrete Example: Particle-Number Symmetry

- Hamiltonian H conserves particle number.
- Symmetric ansatz: only allows excitations preserving particle number: [$\mathrm{U}(\theta) = \prod_{i,j} \exp\left(\theta_{ij}(a_i^\dagger a_j - a_j^\dagger a_i)\right)$]
- Result: VQE only explores states with the correct particle number.
- Barren plateaus are avoided because the effective Hilbert space is exponentially smaller.

Conclusion

- Symmetry restriction reduces risk of barren plateaus
- Enables more efficient training of variational quantum algorithms

SYMMETRY RESTRICTION AND BARREN PLATEAUS IN VARIATIONAL QUANTUM ALGORITHMS

Introduction

- Barren plateaus: regions with vanishing gradients in variational quantum algorithms
- Training becomes difficult due to exponentially small gradients

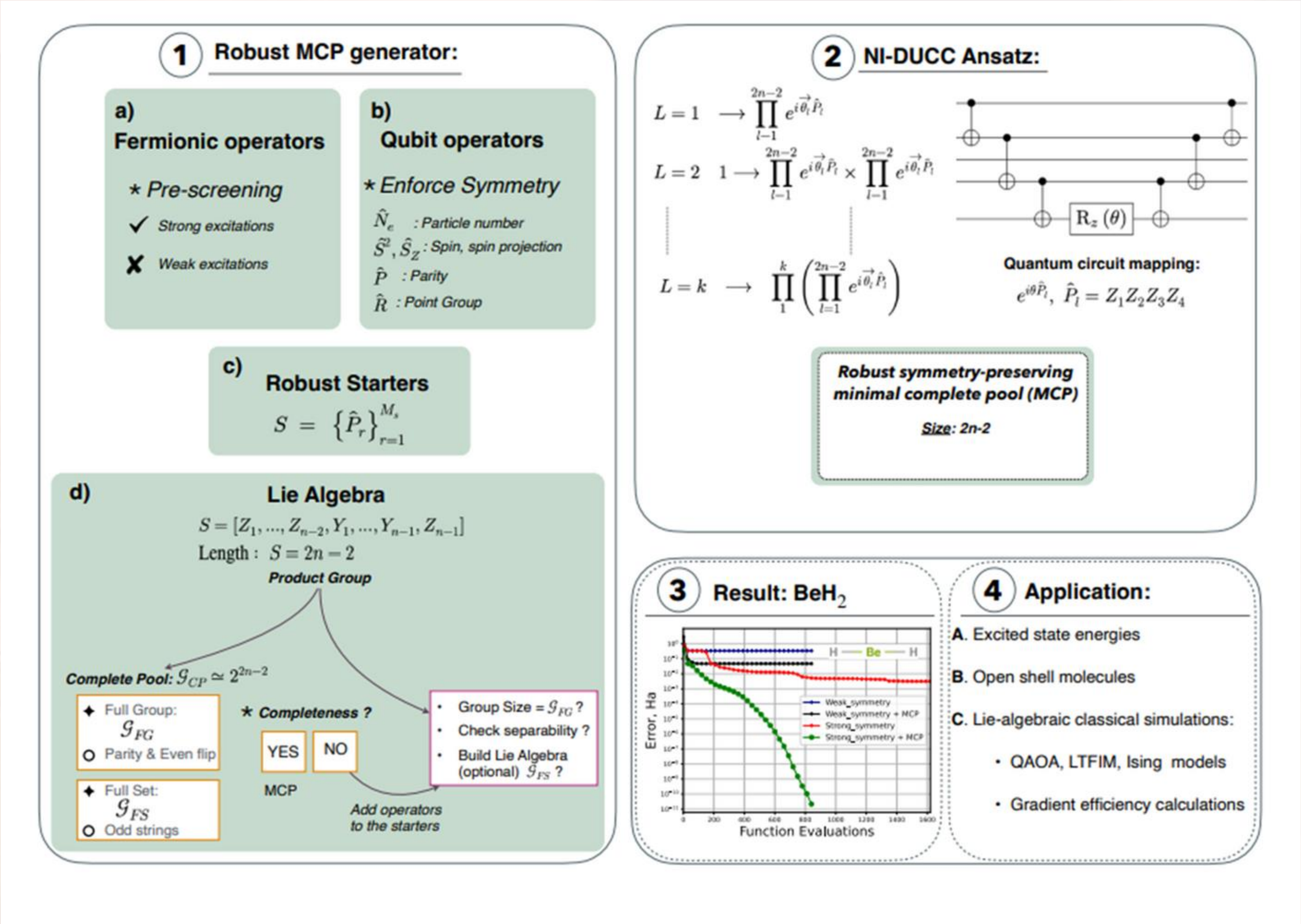
- **Efficient Symmetry-Preserving State Preparation Circuits for the VQE Algorithm (Gard et al., 2019)** Constructs ansätze preserving symmetries relevant for chemistry, avoiding irrelevant sectors. Tested on H_2 and LiH, showing improved performance and lower circuit depth. Link: https://arxiv.org/abs/1904.10910

- **Variational Quantum Circuits to Prepare Low Energy Symmetry States (Selvarajan, Sajjan, Kais, 2022)**Circuits map trial states into a symmetry subspace and optimize within it. Tested on spin-XXZ Hamiltonian and H_2 (Sz = 0). Converges well with low-depth circuits. Link: Efficient Symmetry-Preserving State Preparation Circuits for the VQE Algorithm (Gard et al., 2019) Constructs ansätze preserving symmetries relevant for chemistry, avoiding irrelevant sectors. Tested on H_2 and LiH, showing improved performance and lower circuit depth. Link: https://arxiv.org/abs/1904.10910

- **Efficient Particle-Conserving Symmetric Quantum Circuits (Ayeni et al., 2025) Method for building symmetric circuits respecting particle number and Z_2/Z_3 symmetries. Numerical analysis shows better performance with symmetry-preserving parametrizations.** Link: Variational Quantum Circuits to Prepare Low Energy Symmetry States (Selvarajan, Sajjan, Kais, 2022)Circuits map trial states into a symmetry subspace and optimize within it. Tested on spin-XXZ Hamiltonian and H_2 (Sz = 0). Converges well with low-depth circuits. Link: Efficient Symmetry-Preserving State Preparation Circuits for the VQE Algorithm (Gard et al., 2019) Constructs ansätze preserving symmetries relevant for chemistry, avoiding irrelevant sectors. Tested on H_2 and LiH, showing improved performance and lower circuit depth. Link: https://arxiv.org/abs/1904.10910

- **Preserving Symmetries for VQE in the Presence of Noise (Barron et al., 2021)** Encoding Hamiltonian symmetries reduces quantum/classical resource overhead and mitigates errors. Effective in noisy NISQ devices.Link: Efficient Symmetry-Preserving State Preparation Circuits for the VQE Algorithm (Gard et al., 2019) Constructs ansätze preserving symmetries relevant for chemistry, avoiding irrelevant sectors. Tested on H_2 and LiH, showing improved performance and lower circuit depth. Link: https://arxiv.org/abs/1904.10910

- **Exploiting Symmetry in Variational Quantum Machine Learning (Meyer et al., 2023)** Builds equivariant parametrized circuits respecting problem symmetries. Shows significantly better generalization performance than naive circuits. Link: Variational Quantum Circuits to Prepare Low Energy Symmetry States (Selvarajan, Sajjan, Kais, 2022)Circuits map trial states into a symmetry subspace and optimize within it. Tested on spin-XXZ Hamiltonian and H_2 (Sz = 0). Converges well with low-depth circuits. Link: Efficient Symmetry-Preserving State Preparation Circuits for the VQE Algorithm (Gard et al., 2019) Constructs ansätze preserving symmetries relevant for chemistry, avoiding irrelevant sectors. Tested on H_2 and LiH, showing improved performance and lower circuit depth. Link: https://arxiv.org/abs/1904.10910



Haidar, M., Adjoua, O., Badreddine, S., Peruzzo, A., & Piquemal, J. P. (2025). Non-iterative disentangled unitary coupled-cluster based on lie-algebraic structure. *Quantum Science and Technology*, *10*(2), 025031.

Quantum Computing Approach to Atomic and Molecular Three-Body Systems

Mohammad Haidar,[1] Hugo D. Nogueira,[2] and Jean-Philippe Karr[2,3]

[1] *OpenVQA Hub and WYW, 60 rue François Ier, F-75008 Paris, France*

[2] *Laboratoire Kastler Brossel, Sorbonne Université, CNRS, ENS-Université PSL, Collège de France, 4 place Jussieu, F-75005 Paris,France*

[3] *Université Evry Paris-Saclay, Boulevard François Mitterrand, F-91000 Evry, France*

We present high-precision quantum computing simulations of three-body atoms (He, H^-) and molecules (H_2^+, HD^+), the latter being studied beyond the Born–Oppenheimer approximation. The Non-Iterative Disentangled Unitary Coupled Cluster Variational Quantum Eigensolver (NI-DUCC-VQE) [M. Haidar *et al.*, Quantum Sci. Technol. **10**, 025031 (2025)] is used. By combining a first-quantized Hamiltonian with a Minimal Complete Pool (MCP) of Lie-algebraic excitations, we construct a compact ansatz with a gradient-independent construction, avoiding costly gradient evaluations and yielding efficient computational scaling with both basis size and electron number. It avoids barren plateaus and enables rapid convergence, achieving energy errors as low as 10^{-11} a.u. with state fidelities only limited by arithmetic precision in only a few thousand function evaluations in all four systems. These results make three-body atoms and molecules excellent candidates for benchmarking and testing on current Noisy Intermediate-Scale Quantum (NISQ) devices. Further, our approach can be extended to more complex systems with larger basis sets, taking advantage of the efficient scaling of qubit requirements to study electronic correlations and non-adiabatic effects with high precision. We also demonstrate the applicability of NI-DUCC-VQE for simulating higher-order effects such as relativistic corrections and hyperfine interactions.

First quantization has benefit!

=➔ Scale **ne · log2 N**, where ne is the number of electrons
and N is the number of basis functions. This favorable scaling enables the use of larger basis sets while keeping resource requirements within the range of NISQ devices. In fact, the promise of the first-quantization formalism has already been tested experimentally, in a recent helium calculation using a Quantinuum ion-trap quantum computer :

M. C. Per, N. Rhodes, M. Srikumar, and J. W. Dai, Chemically-accurate prediction of the ionisation potential of helium using a quantum processor, arXiv:2502.02023 (2025)

Conclusion

- A unified theory exist: A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits. What is beyond? Symmetry help? Are more sources of BP?
- Symmetry restriction reduces risk of barren plateaus
- Enables more efficient training of variational quantum algorithms

Acknowledgement

Hamza benKadour

PhD Student in Quantum Computing - Algeria