Lie-algebraic classical simulations for variational quantum computing

Matthew L. Goh[1,2], Martin Larocca[1], Lukasz Cincio[1], Marco Cerezo[1], and Frédéric Sauvage[1]

[1]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

[2]*Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*

(LA-UR-24-25024)

The barren plateau problem [1] is the single greatest hindrance to quantum machine learning (QML) with variational quantum algorithms (VQAs). Much progress has been made in understanding the origins of barren plateaus, and the conditions under which a VQA is provably barren-plateau-free [1, 2, 3], however quantum advantage with VQAs remains elusive. A growing connection is emerging between trainability of VQAs and classical simulability [4], and presently all known VQAs that provably lack barren plateaus can be efficiently simulated on a classical computer (with caveats). Achieving a genuine quantum advantage therefore requires an appreciation of what is and is not classically simulable.

Most barren plateaus can be connected to the Lie algebra $\mathfrak{g}$ associated with the ansatz circuit [2, 3, 5], which characterizes the set of possible rotations the ansatz can induce in Hilbert space. In particular, when the dimension of $\mathfrak{g}$ scales polynomially in the system size n as $\dim(\mathfrak{g}) \in \mathrm{poly}(n)$, the circuit's expressivity is sufficiently low that the gradients are only polynomially small, and efficient training is possible. Such polynomial scaling is typically a consequence of underyling structure or symmetry, such as free-fermionic structure [6] or permutation symmetry [7]. In this work, we extend upon the tools of Somma [8, 9] to develop efficient classical simulations of this class of circuits. We present and implement a framework for Lie-algebraic classical simulations of VQAs, which we dub '$\mathfrak{g}$-sim'. For a poly-dimensional circuit Lie algebra, our framework allows for weak simulation of *any* initial state, so long as a polynomial number of observable expectation values are known and the desired observables are supported by $\mathfrak{g}$, or for strong simulation of initial states supported by $\mathfrak{g}$. Since the Lie algebra dimension of a variational circuit encapsulates most barren plateau phenomena [2, 3], our work represents the central pillar of the crucial connection between trainability and simulability [4].

We present efficient implementations of $\mathfrak{g}$-sim, demonstrating our methods on a transverse field XY model up to 200 qubits. In particular, we simulate the evolution of an initial magic state that is intractable in other classical simulation methods (such as fermionic linear optics), simulating otherwise classically intractable dynamics in minutes on a single CPU core. Furthermore, we show that *gradients* of loss functions with respect to variational circuit parameters can be computed efficiently on a classical computer for circuits with $\dim(\mathfrak{g}) \in \mathrm{poly}(n)$, and employ this to extend the framework to *optimization* of variational circuits. This enables a diverse range of further applications, which we detail below.

Although our methods themselves are efficient on a classical computer, they may nonetheless have utility for QML and quantum computing more broadly: by serving as a companion model for scalable study, providing a means

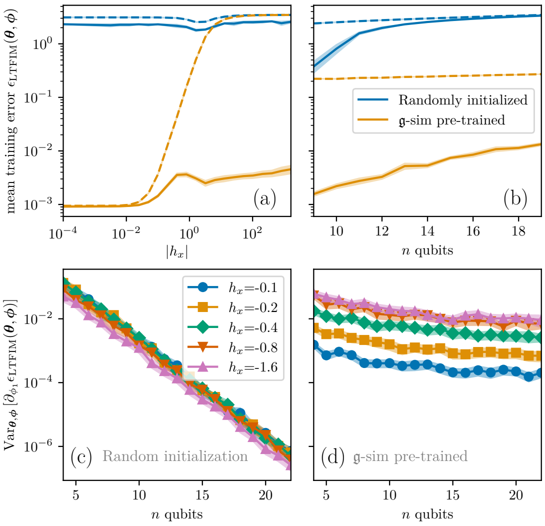

Figure 1: **Mitigating barren plateaus via $\mathfrak{g}$-sim pre-training.** **(a, b)** Mean training error for the variational quantum eigensolver on the longitudinal and transverse field Ising model with random initialization (blue) and pre-trained with $\mathfrak{g}$-sim (orange), (a) at $n = 12$ qubits and varying longitudinal field strengths h_x and (b) at field strength $h_x = -1.0$ and varying system sizes n. Dashed lines represent initial ansatz configurations, while solid lines represent trained ansätze. In all cases seen here, the pre-trained solution significantly outperforms random initialization. **(c, d)** Comparison of gradient variances at varying system size n and longitudinal field strengths h_x for (c) uniform random parameter initialization and (d) $\mathfrak{g}$-sim pre-training. The use of pre-training greatly improves initial parameter gradients, mitigating the barren plateau effect.

of classical pre-training, enabling unitary compilation for important algorithm subroutines, or classically training supervised QML models that can later be used on quantum hardware to perform inference. We thus proceed to explore in depth a range of applications highlighting these potential areas of utility:

- **Benchmarking variational circuits:** The success of QML with variational circuits depends not only on the ability to compute gradients, but other features of the optimization landscape such as poor local minima. Analytic study of such phenomena can be extremely challenging, and numerical experiments are typically limited to small system sizes that can be exactly represented in state-vector form on a classical computer, leaving ambiguity about the scalability of many approaches. Classical simulation methods like $\mathfrak{g}$-sim can be of great utility in studying trainability phenomena at scale. To this end, we simulate the variational quantum eigensolver (VQE) [10] on a spin chain problem at system sizes intractable with state-vector simulation methods. We focus on the phenomenon of overparameterization [11], where poor local minima can become optimizable saddle points past a critical parameter threshold. In its original demonstration, this phenomenon was only demonstrated numerically up to 10 qubits - we demonstrate the overparametrization phenomenon at up to 50 qubits.

- **Pre-training / warm-starting:** Although barren plateaus render an optimization landscape exponentially flat *on average*, preventing efficient training from random parameter initialization, it is hoped that judicious selection of initial parameters (pre-training or warm-starting) may mitigate this effect. Although models that are classically simulable can obviously yield no quantum advantage, they may be closely related to models which are genuinely intractable on a classical device. We focus on two high-value problems: the disordered Heisenberg chain [12], and the quantum approximate optimization algorithm (QAOA) [13]. In both cases, we begin with a related, but simpler form of the model that can be efficiently simulated and optimized in $\mathfrak{g}$-sim, and then transfer these pre-trained parameters to the full model (which must be post-trained on a quantum device or classical emulator). In essence, this is a form of transfer learning between a barren-plateau-free classically simulable model, and a barren-plateau-afflicted target model. Due to the overparametrization phenomenon [11], the simplified classical model can *always* be trained. We show that this method yields far superior solutions for both VQE and deep-circuit QAOA on average (compared to random initialization), and mitigates the barren plateau problem by yielding greatly increased loss function gradients. Thus, with such methods one may be able to efficiently train a useful model in spite of its barren plateau. We provide a sample of our pre-training results in Figure 1.

- **Circuit synthesis:** we show that $\mathfrak{g}$-sim can be used not only to optimize observable expectation values, but also to optimize unitary fidelities. Combined with the phenomenon of overparametrization [11], we thus present a method by which any unitary in the corresponding Lie group can be approximately compiled to a polynomial-depth parametrized circuit. We numerically demonstrate favourable scaling of trainability up to 50 qubits, and apply the scheme to dynamical simulation of a spin-chain problem. Related compilation methods, such as the use of Givens rotations to implement transformations with free-fermion generators in quantum chemistry, have been of enormous utility [14, 15]. We reiterate that $\mathfrak{g}$-sim is tractable for any circuit and cost function supported by $\mathfrak{g}$, allowing classically-efficient circuit compilation for a wider class of subroutines in quantum algorithms.

- **Supervised quantum machine learning:** we use our scheme to train and simulate a quantum neural network (QNN) for classifying ground state phases across the phase transition of a 'disguised' spin chain model. Our model learns an observable that distinguishes between phases and can be measured with a polynomial number of shots on a quantum computer, or in some cases even approximated in a logarithmic number of shots with classical shadows [16]. Even though the classifier is classically simulable, it was recently remarked in Ref. [4] that there can be utility in executing models like this directly on quantum hardware, since performing inference directly on hardware with the trained classifier can be cheaper than the data acquisition required to initialize a classical simulation. Such an approach is useful for quantum sensing applications, where classical trainability may be useful, but inference on test states necessarily must be performed on real hardware.

In summary, the $\mathfrak{g}$-sim framework is a central pillar in the emerging connection between trainability and simulability of variational quantum circuits. Expressivity (measured by $\dim(\mathfrak{g})$) is a leading cause of barren plateaus, and the $\mathfrak{g}$-sim framework constructively turns the very same tools used to diagnose such barren plateaus into a classical simulation technique. These classical simulations then enable high-value applications: by exploiting efficient implementations of $\mathfrak{g}$-sim, we are able to study trainability at scale, pre-train models that would otherwise suffer barren plateaus, compile unitaries to compact circuits, and train supervised QML models.

[1] M. Larocca et al., arXiv preprint arXiv:2405.00781 (2024).

[2] E. Fontana et al., arXiv preprint arXiv:2309.07902 (2023).

[3] M. Ragone et al., arXiv preprint arXiv:2309.09342 (2023).

[4] M. Cerezo et al., arXiv preprint arXiv:2312.09121 (2023).

[5] M. Larocca et al., Quantum **6**, 824 (2022).

[6] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, Physical Review X **10**, 031064 (2020).

[7] E. R. Anschuetz, A. Bauer, B. T. Kiani, and S. Lloyd, Quantum **7**, 1189 (2023).

[8] R. D. Somma, arXiv preprint quant-ph/0512209 (2005).

[9] R. Somma, H. Barnum, G. Ortiz, and E. Knill, Physical Review Letters **97**, 190501 (2006).

[10] A. Peruzzo et al., Nature Communications **5**, 1 (2014).

[11] M. Larocca, N. Ju, D. García-Martín, P. J. Coles, and M. Cerezo, Nature Computational Science **3**, 542 (2023).

[12] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Proceedings of the National Academy of Sciences **115**, 9456 (2018).

[13] E. Farhi, J. Goldstone, and S. Gutmann, arXiv preprint arXiv:1411.4028 (2014).

[14] I. D. Kivlichan et al., Physical review letters **120**, 110501 (2018).

[15] F. Arute et al., Science **369**, 1084 (2020).

[16] H.-Y. Huang, R. Kueng, and J. Preskill, Nature Physics **16**, 1050 (2020).