Quantum Machine Learning Applications in High-Energy Physics

(Invited Paper)

Andrea Delgado
Physics Division, Oak Ridge National Laboratory
Oak Ridge, Tennessee, USA
delgadoa@ornl.gov

Kathleen E. Hamilton
Computer Science and Engineering Division, Oak Ridge National Laboratory
Oak Ridge, Tennessee, USA
hamiltonke@ornl.gov

ABSTRACT

Some of the most significant achievements of the modern era of particle physics, such as the discovery of the Higgs boson, have been made possible by the tremendous effort in building and operating large-scale experiments like the Large Hadron Collider or the Tevatron. In these facilities, the ultimate theory to describe matter at the most fundamental level is constantly probed and verified. These experiments often produce large amounts of data that require storing, processing, and analysis techniques that continually push the limits of traditional information processing schemes. Thus, the High-Energy Physics (HEP) field has benefited from advancements in information processing and the development of algorithms and tools for large datasets. More recently, quantum computing applications have been investigated to understand how the community can benefit from the advantages of quantum information science. Nonetheless, to unleash the full potential of quantum computing, there is a need to understand the quantum behavior and, thus, scale up current algorithms beyond what can be simulated in classical processors. In this work, we explore potential applications of quantum machine learning to data analysis tasks in HEP and how to overcome the limitations of algorithms targeted for Noisy Intermediate-Scale Quantum (NISQ) devices.

CCS CONCEPTS

• **Hardware → Quantum computation**; • **Applied computing → Physics**; • **Computing methodologies → Unsupervised learning**.

KEYWORDS

quantum computing, quantum machine learning, physics, unsupervised learning

ACM Reference Format:
Andrea Delgado and Kathleen E. Hamilton. 2023. Quantum Machine Learning Applications in High-Energy Physics: (Invited Paper). In *Proceedings of (ICCAD '22)*. ACM, San Diego, CA, USA, 5 pages.

ICCAD '22, October 30 - November 03, 2022, San Diego, CA

1 INTRODUCTION

High-energy physics (HEP) is the branch of physics that deals with understanding matter at the most fundamental level. Our current understanding of the universe is encapsulated in the so-called standard model (SM) of particle physics. Although the SM has been verified extensively through validating theoretical predictions with experimental results, many questions remain unanswered, such as the nature of dark matter and the neutrino mass.

Large-scale experiments have been built to probe or extend the SM. An example of these experiments is the Large Hadron Collider (LHC) and the upcoming Deep Underground Neutrino Experiment (DUNE). At these large-scale facilities, high-performance data storage and processing schemes are needed to store, access, retrieve, distribute, and process experimental data, posing a challenge to conventional information processing techniques.

HEP research programs could benefit tremendously from current and future QIS technologies and applications. Both fields are woven into the fabric of reality at the deepest level, and it should not be surprising that QIS represents powerful enabling technological advances in HEP. Some of the applications that have been explored include:

- Simulation of quantum systems [3, 13],
- Calculation of HEP-relevant nuclear physics calculations, such as neutrino-nucleus scattering cross sections [7],
- Understanding quantum gravity [15],
- Quantum sensors for the detection of beyond-the-SM physics and particles [1, 2], and
- Data analysis with quantum computers [11].

This manuscript focuses on the latter, analyzing experimental data from HEP experiments, assisted by quantum-enhanced algorithms, and particularly quantum machine learning (QML) techniques.

2 QUANTUM MACHINE LEARNING

QIS is a rapidly developing field focused on understanding the analysis, processing, and transmission of information using quantum mechanical principles and computational techniques. QIS could address the conventional computing gap associated with HEP-related problems, specifically those computational tasks that challenge CPUs and GPUs, such as efficient and accurate event generators and classifiers. Another potential groundbreaking application of QIS technologies is the quantum-enhanced search for beyond-the-SM physics.

QML lies at the intersection of QIS and machine learning. The hope is to leverage QIS and speed up some subroutines in classical

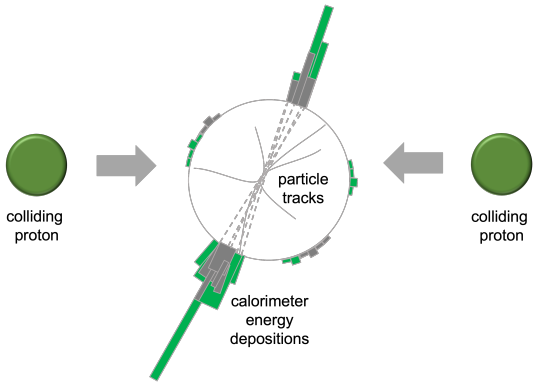

Figure 1: Diagram of a typical interaction of a proton-proton collision at the LHC, where two protons collide at a speed close to that of light. The collision debris is studied by multi-purpose detectors such as the Compact Muon Solenoid (CMS) and ATLAS, which employ different detection mechanisms, including calorimetry and silicon pixel trackers, to reconstruct the initial interaction as accurately as possible. The combined information from these sub-detector systems is integrated in an optimized way to construct higher-level objects such as jets and particle tracks, amenable to statistical analysis for signal extraction. Thus, a typical HEP analysis of experimental data involves reliable simulation, clustering, and classification tools.

machine learning algorithms to speed up training times, enhance model expressibility and generalization, find correlations that are not trivial for classical neural network-based models, and reduce the training sample size.

In HEP, several algorithms have been explored to analyze experimental data, including track reconstruction [19], jet clustering [10] and tagging [12], and generative modeling [9]. See Ref. [11] for an extensive overview of QML applications to HEP data analysis.

3 VARIATIONAL CIRCUIT TRAINING.

Currently, most QML models are trained in a hybrid setting in the universal quantum computing model, incorporating both classical and quantum resources. In this context, a parameterized quantum circuit (PQC) replaces the classical neural network. Its parameters are optimized through classical means, such as gradient and non-gradient based optimization on a classical device.

A parameterized quantum circuit is typically structured as a composition of repeated layers

$$U(\theta) = \prod_{d=1}^{D} U_d(\theta^{(d)}) \tag{1}$$

where U_d is a parameterized unitary with trainable parameters $\theta^{(d)}$ and that is repeated D times. Each *unit-cell* or *circuit block* consists of a rotation and an entangling component.

Thus, PQCs can be seen as components of a model designed for a data-driven task, such as classification or generative modeling.

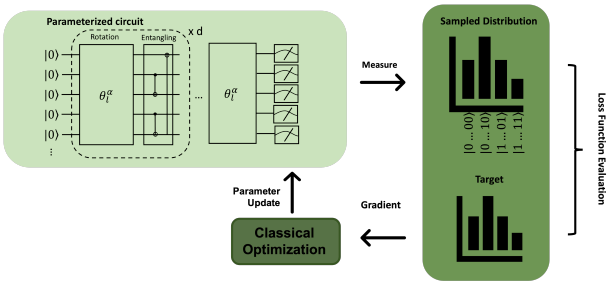

Figure 2: Diagram of the differentiable QCBM training scheme.

Circuit learning then involves iteratively updating the parameters according to a pre-defined cost function towards a goal.

4 UNSUPERVISED QUANTUM CIRCUIT LEARNING

Quantum generative models are expected to exhibit an advantage over their classical counterparts in terms of runtime and the number of parameters needed to learn data distributions due to their strong expressive power. Furthermore, the ability of quantum information processors to represent vectors in $N-$dimensional spaces using $\log(N)$ qubits and to perform manipulations of sparse and low-rank matrices in time $O(\text{poly}(\log(N)))$ [16] motivates the exploration of quantum generative models as an alternative to classical generative models in HEP.

A quantum circuit Born machine (QCBM) is an example of an implicit model for generative learning [18] that generates samples by performing measurements in a given quantum register, as a Born machine[4, 8]. To train a QCBM, we start by constructing a parameterized unitary $\langle 0|\theta|0\rangle$ that prepares N- qubits in the state $|\Psi_\Theta\rangle = U(\vec{\theta})\,|\Phi_0\rangle$. Then, a classical distribution can be obtained over 2^N computational basis states by measuring $|\Psi_\Theta\rangle$ in a fixed basis. The parameter optimization is then performed by minimizing a loss function $\mathcal{L}(P_{target}, P_{QCBM})$ that computes the similarity between the target distribution (Q) and the distribution sampled from the QCBM(P).

In Ref. [9], we use non-adversarial gradient-based training of 8- and 12-qubit QCBMs to generate joint distributions over 2 and 3 variables to generate synthetic data of a typical HEP process. We used several circuit ansatzes found in the quantum computing literature and tested the trainability of QCBM initialized with different quantum states. These ansatzes were chosen because of their proven trainability for QML applications and their ability to be deployed on hardware.

Although both circuit designs (Figure 3) were able to reproduce the joint target distribution and associated marginals with high fidelity, only Ansatz 1 was able to fit the correlations between 2- or 3-variables. Furthermore, we observed a high impact in the ansatz choice in the model capacity and trainability, suggesting a need for an efficient way to design circuit ansatz tailored to the task at hand.

Motivated by the success of gradient-based training for 12 qubits QCBM, we seek to explore the scalability of quantum generative models trained in an unsupervised setting beyond 12 qubits and

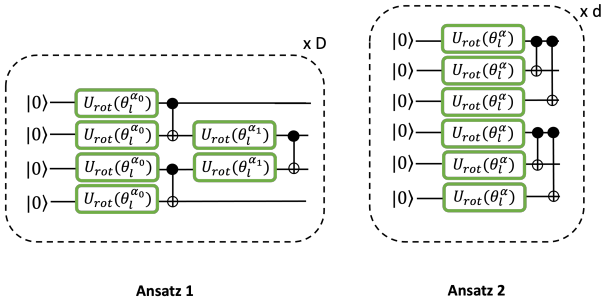

Figure 3: Diagram for one layer of the circuit templates used to construct and train the QCBM in Ref. [9]. For both ansatz, a final layer of rotation gates is added before measurement. The rotational gates used in the unitary are the universal rotation gates, with three trainable parameters.

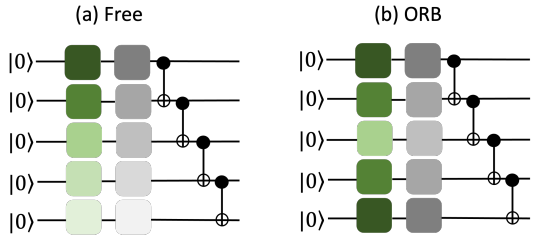

Figure 4: Diagram of the circuit designs considered in this study for n=5 qubits.

to develop strategies to overcome the limitations of training large-scale QML models.

5 BUILDING SYMMETRIES INTO QUANTUM CIRCUIT LEARNING

The PQC design plays an essential role in the performance of any QML model by defining the hypothesis class. Highly parameterized circuits designed for noisy intermediate-scale quantum (NISQ) device applications offer great flexibility by providing generalization capabilities for a wide range of potential solutions in shallow depths. However, scalability is often an issue. Enhanced expressibility comes at the expense of trainability issues limiting the extent to which these circuits can be optimized at a scale [6, 14, 17]. Furthermore, parameter optimization over deep circuits can lead to noise-induced trainability issues [21]. This problem affects QML applications specifically, where it is hard to find a "problem-inspired" ansatz, as in the case of quantum simulations. To this end, we study an informed ansatz design with sufficiently independent parameters per layer to be realized in a shallow depth but maintain the expressibility needed for the QML task at hand.

In Ref. [20], the concept of *ORB* circuits that incorporate spatial symmetries of the system under study was introduced. The method is based on the correlation of gate parameters belonging to the same "orbit" according to the symmetries of the problem. It was also shown that the *ORB* method allows maximizing the number of free parameters per circuit layer without impeding the trainability of the circuits. The general idea behind the construction of ORB-like circuits for unsupervised circuit learning tasks is the possibility of reducing the space of states that need to be explored during training.

We employ the hardware-efficient SU(2) 2-local circuit to explore this effect. The circuit consists of layers of single-qubit operations spanned by *SU(2)* and *CX* entanglements. $SU(2)$ stands for *special unitary group of degree 2*. The group elements are 2×2 unitary matrices with determinant 1, such as the Pauli rotation gates. Here, we use the Pauli-Y and Z single qubit gates, along with *CX* gates for entanglement, applied in a "linear" configuration, among neighboring pairs of qubits

$$U(\theta^{(d)}) = U_{ent} \times \prod_{i=1}^{n} R_Y(\theta^d_{i,1}) R_Z(\theta^d_{i,2}) \tag{2}$$

for rotations acting on the qubits $i = 1, ..., n$. All the rotational gates used to construct the circuit are independent and updated in every single iteration of the training workflow (Figure 4(a)).

An *ORB*-inspired circuit is also considered, where, assuming a vertical symmetry on reflection over the middle qubit in the circuit, rotational gates are grouped into "orbits". Gates in the same orbit share parameters, effectively reducing the number of trainable parameters (Figure 4(b)).

6 RESULTS

To understand the impact of reducing the number of trainable parameters in our circuit design, we study their performance, focusing on their scaling for increased problem size.

We start by training a parameterized Efficient SU(2) circuit to reproduce a gaussian probability distribution through gradient-based optimization. The cost function is the Jensen-Shannon loss evaluated between the target and the QCBM-sampled distributions, and optimized using the Adam optimizer with a learning rate optimized for every circuit. We obtain the optimal number of layers ($L_C(\epsilon)$) by trating $L_C(\epsilon)$ as a hyperparameter during training. By starting from circuits with two layers, we increase the number of layers by one, until the JS value reaches a plateau. Circuits are initialized in the all-zero state and random parameters. A numerical simulation is performed in a noiseless setting in PennyLane [5].

In Figure 5, we observe the effect of increasing the number of layers in a circuit. For a small number of layers, the JS value is high, and performance is highly dependent on the choice of initial parameters.

To gauge the performance of *ORB* circuits when compared to the fully parameterized Efficient SU(2) circuit, we compare the JS loss value between circuits constructed with the Efficient SU(2) and the *ORB*-like ansatzes evaluated on the trained parameters. The experiment is repeated 25 times, and the mean JS value is plotted in Figure 6. The bars correspond to the standard deviation. We notice that for $n_{qubits} = 5$ (gray), the JS value converges at a similar number of layers for both Efficient SU(2) and *ORB* circuits. For $n_{qubits} = 7$, it takes a larger number of layers for the *ORB* circuit to reach a similar performance than the Efficient SU(2) circuit.

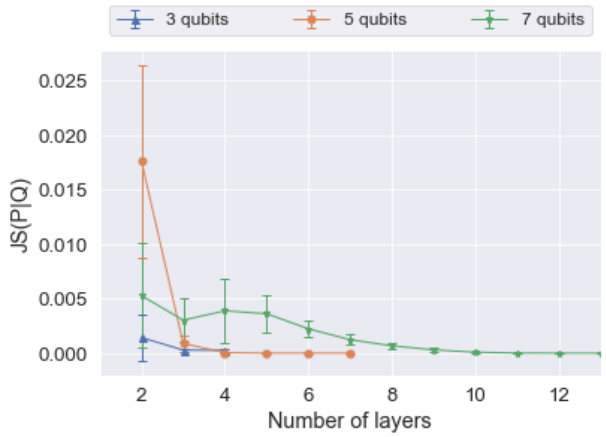

Figure 5: JS value between target and QCBM distributions as a function of the number of layers in the circuit.

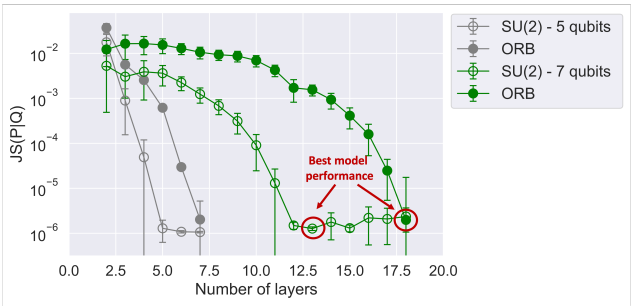

Figure 6: JS value between target and QCBM distributions as a function of the number of layers in circuits constructed with the Efficient SU(2) ansatz (open circles) and ORB-inspired ansatz (solid circles) for n_{qubits} = 5 (gray) and 7 (green).

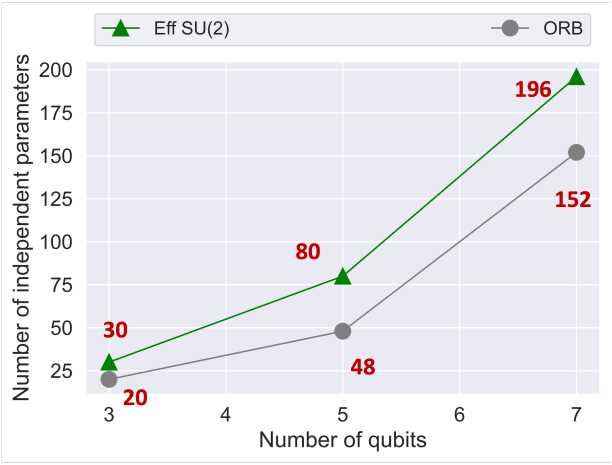

Figure 7: Number of independent parameters at $L_C(\epsilon)$ as a function of the number of qubits in circuits constructed with the Efficient SU(2) ansatz (green) and *ORB*-inspired ansatz (gray).

Additionally, we will report the number of independent parameters corresponding to $L_C(\epsilon)$ and where n_l denotes the number of parameters per layer. In Figure 7, the number of independent parameters at $L_C(\epsilon)$ is plotted as a function of the number of qubits in circuits constructed with the Efficient SU(2) ansatz (green) and *ORB*-inspired ansatz (gray). For the case of n_{qubits} = 3 and 5, $L_C(\epsilon)$ is the same. For n_{qubits} = 7, $L_C(\epsilon)$ = 13 and 18 for circuits constructed with the Efficient SU(2) ansatz and *ORB*-inspired ansatz, respectively. The general trend is a smaller number of independent parameters needed to train *ORB*–inspired circuits than the traditional Efficient SU(2) ansatz.

7 CONCLUSION

Quantum computing, specifically quantum machine learning, has found many applications in the field of HEP for analyzing experimental data. Although several algorithms have been explored and benchmarked on various NISQ devices, several limitations are still associated with the adaption of QIS technologies in day-to-day analysis workflows.

One of the challenges explored in this study is the scalability of some of the traditional circuit designs available in the QML literature. The number of resources required for training QML models scales with the depth and size of the circuits. Thus, it is desirable to incorporate domain knowledge into the ansatz design to reduce the number of circuit evaluations and trainable parameters.

We extend the work presented in Ref. [9], where we successfully trained a 12-qubit QCBM by leveraging the symmetry associated with the circuit structure in the Efficient SU(2) ansatz. In unsupervised quantum circuit learning, we demonstrate that by considering the symmetry related to the Efficient SU(2) ansatz, we can pack many more free parameters per layer while ensuring a similar expressibility. These preliminary results show that although the number of parameters is considerably reduced, trainability is not affected, and learning performance is comparable to the over-parameterized circuit baseline. As a follow-up, we will test this methodology for circuits with n_{qubits} greater than seven qubits and expand the study to leverage symmetries introduced by two-qubit gates and other circuit designs.

ACKNOWLEDGMENTS

This work was partially supported by the Quantum Information Science Enabled Discovery (QuantISED) for High Energy Physics program at ORNL under FWP ERKAP61. This work was partially supported by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U. S. Department of Energy. This work was partially supported as part of the ASCR Testbed Pathfinder Program at Oak Ridge National Laboratory under FWP ERKJ332. This work was partially supported as part of the ASCR Fundamental Algorithmic Research for Quantum Computing Program at Oak Ridge National Laboratory under FWP ERKJ354. This research used quantum computing system resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. Oak Ridge National Laboratory manages access to the IBM Q System as part of the IBM Q Network.

REFERENCES

[1] Zeeshan et al. Ahmed. 2018. Quantum Sensing for High Energy Physics. https://doi.org/10.48550/ARXIV.1803.11306

[2] M. Sohaib Alam et al. 2022. Quantum computing hardware for HEP algorithms and sensing. In *2022 Snowmass Summer Study*. arXiv:2204.08605 [quant-ph]

[3] Christian W. Bauer et al. 2022. Quantum Simulation for High Energy Physics. (4 2022). arXiv:2204.03381 [quant-ph]

[4] Marcello Benedetti, Delfina Garcia-Pintos, Oscar Perdomo, Vicente Leyton-Ortega, Yunseong Nam, and Alejandro Perdomo-Ortiz. 2019. A generative modeling approach for benchmarking and training shallow quantum circuits. *npj Quantum Inf.* 5, 45 (2019).

[5] Ville et al. Bergholm. 2018. PennyLane: Automatic differentiation of hybrid quantum-classical computations. https://doi.org/10.48550/ARXIV.1811.04968

[6] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. 2021. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications* 12, 1 (mar 2021). https://doi.org/10.1038/s41467-021-21728-w

[7] Anthony Ciavarella. 2020. Algorithm for quantum computation of particle decays. *Phys. Rev. D* 102 (Nov 2020), 094505. Issue 9. https://doi.org/10.1103/PhysRevD.102.094505

[8] Brian Coyle, Daniel Mills, Vincent Danos, and Elham Kashefi. 2020. The Born supremacy: quantum advantage and training of an Ising Born machine. *npj Quantum Information* 6, 1 (2020), 1–11.

[9] Andrea Delgado and Kathleen E. Hamilton. 2022. Unsupervised Quantum Circuit Learning in High Energy Physics. (3 2022). arXiv:2203.03578 [quant-ph]

[10] Andrea Delgado and Jesse Thaler. 2022. Quantum Annealing for Jet Clustering with Thrust. (5 2022). arXiv:2205.02814 [quant-ph]

[11] Andrea et al. Delgado. 2022. Quantum Computing for Data Analysis in High-Energy Physics. https://doi.org/10.48550/ARXIV.2203.08805

[12] Alessio Gianelle, Patrick Koppenburg, Donatella Lucchesi, Davide Nicotra, Eduardo Rodrigues, Lorenzo Sestini, Jacco de Vries, and Davide Zuliani. 2022. Quantum Machine Learning for b-jet charge identification. *JHEP* 08 (2022), 014. https://doi.org/10.1007/JHEP08(2022)014 arXiv:2202.13943 [hep-ex]

[13] Erik J. Gustafson and Henry Lamm. 2021. Toward quantum simulations of $\mathbb{Z}_2$ gauge theory without state preparation. *Phys. Rev. D* 103, 5 (2021), 054507. https://doi.org/10.1103/PhysRevD.103.054507 arXiv:2011.11677 [hep-lat]

[14] Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles. 2022. Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus. *PRX Quantum* 3 (Jan 2022), 010313. Issue 1. https://doi.org/10.1103/PRXQuantum.3.010313

[15] Seth Lloyd. 2005. A theory of quantum gravity based on quantum computation. https://doi.org/10.48550/ARXIV.QUANT-PH/0501135

[16] Seth Lloyd and Christian Weedbrook. 2018. Quantum Generative Adversarial Learning. *Phys. Rev. Lett.* 121 (Jul 2018), 040502. Issue 4. https://doi.org/10.1103/PhysRevLett.121.040502

[17] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. 2018. Barren plateaus in quantum neural network training landscapes. *Nature Communications* 9, 1 (nov 2018). https://doi.org/10.1038/s41467-018-07090-4

[18] Shakir Mohamed and Balaji Lakshminarayanan. 2016. Learning in Implicit Generative Models. https://doi.org/10.48550/ARXIV.1610.03483

[19] Gregory Quiroz, Lauren Ice, Andrea Delgado, and Travis S. Humble. 2021. Particle track classification using quantum associative memory. *Nucl. Instrum. Meth. A* 1010 (2021), 165557. https://doi.org/10.1016/j.nima.2021.165557 arXiv:2011.11848 [quant-ph]

[20] Frederic Sauvage, Martin Larocca, Patrick J. Coles, and M. Cerezo. 2022. Building spatial symmetries into parameterized quantum circuits for faster training. https://doi.org/10.48550/ARXIV.2207.14413

[21] Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles. 2021. Noise-induced barren plateaus in variational quantum algorithms. *Nature Communications* 12, 1 (nov 2021). https://doi.org/10.1038/s41467-021-27045-6