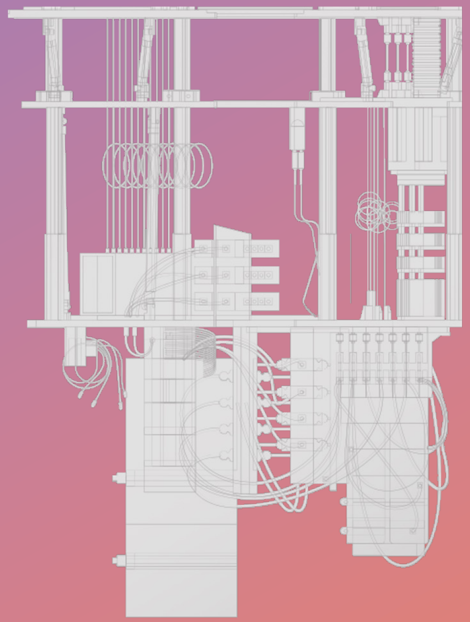
Quantum Machine Learning for Data Analysis in HEP

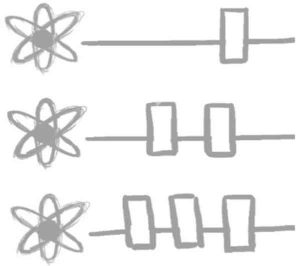


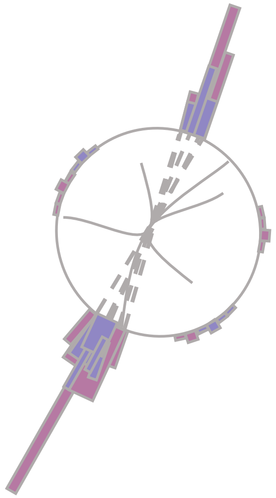
Andrea Delgado (she/her/ella)

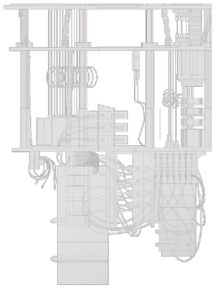
Oak Ridge National Laboratory

Roadmap of the talk

- ★ **Why, What, When Quantum Machine Learning?**
- ★ **Parameterized Quantum Circuits as Machine Learning Models**
- ★ ***Applications in Analysis of HEP Data***
- ★ ***Barren plateaus, and how to avoid them?***
- ★ ***Is Quantum Advantage the Right Goal for QML?***

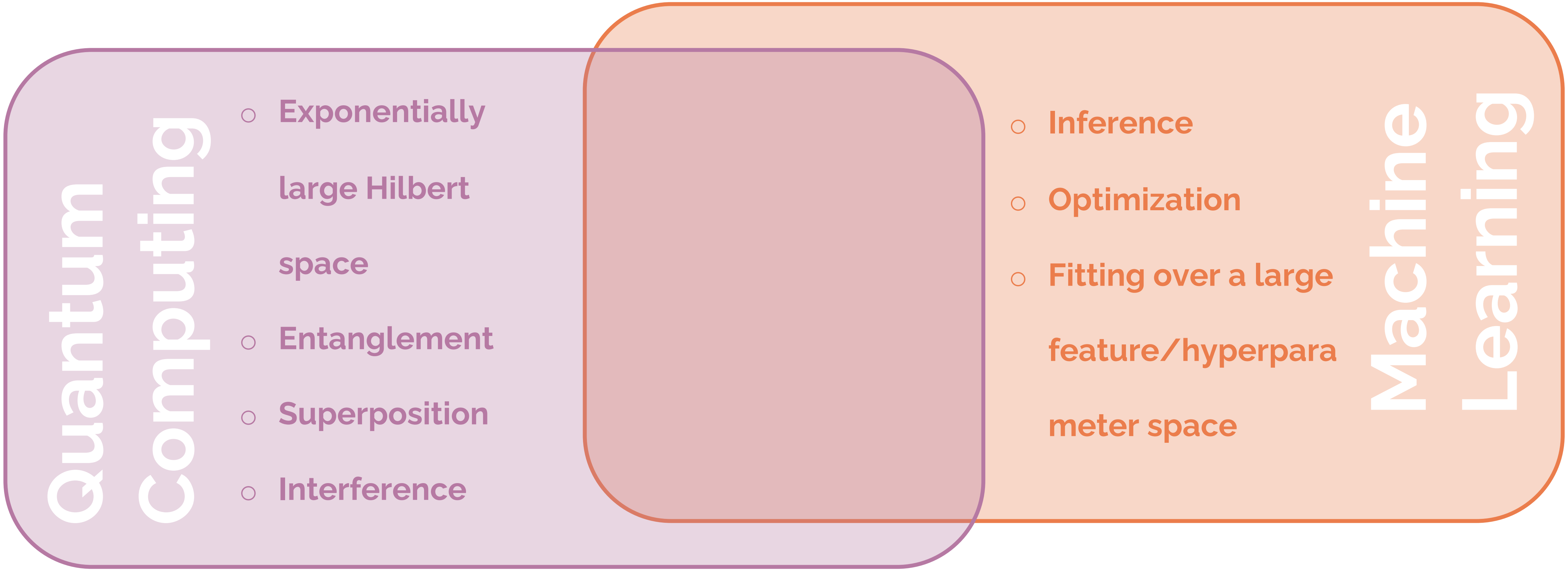






Quantum Machine Learning

The main goal of Quantum Machine Learning (QML) is to speed things up by applying what we know from quantum computing to machine learning



QML takes elements from classical machine learning theory, and views quantum computing from that lens

Quantum Machine Learning

The main goal of Quantum Machine Learning (QML) is to speed things up by applying what we know from quantum computing to machine learning

Quantum Computing

- Exponentially large Hilbert space
- Entanglement
- Superposition
- Interference

Quantum Computing ∩ Machine Learning

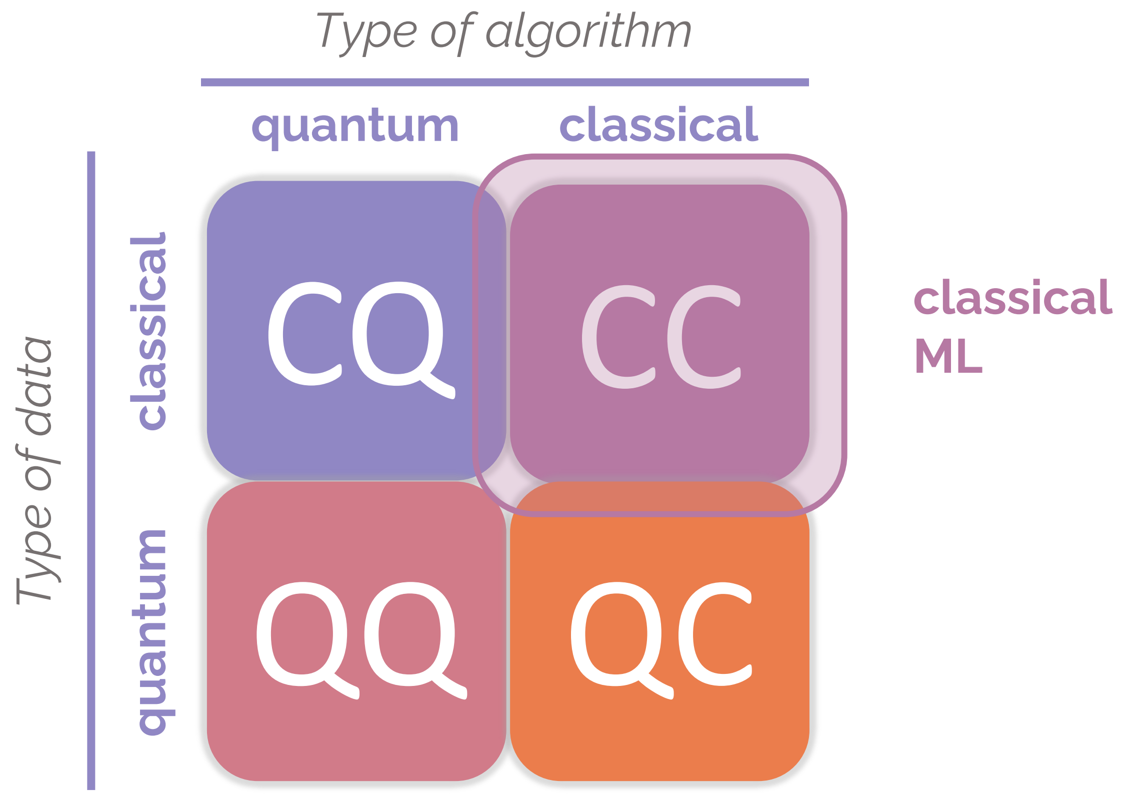
- Linear algebraic problems
- Kernel methods
- Optimization
- Deep quantum learning

Machine Learning

- Inference
- Optimization
- Fitting over a large feature/hyperparameter space

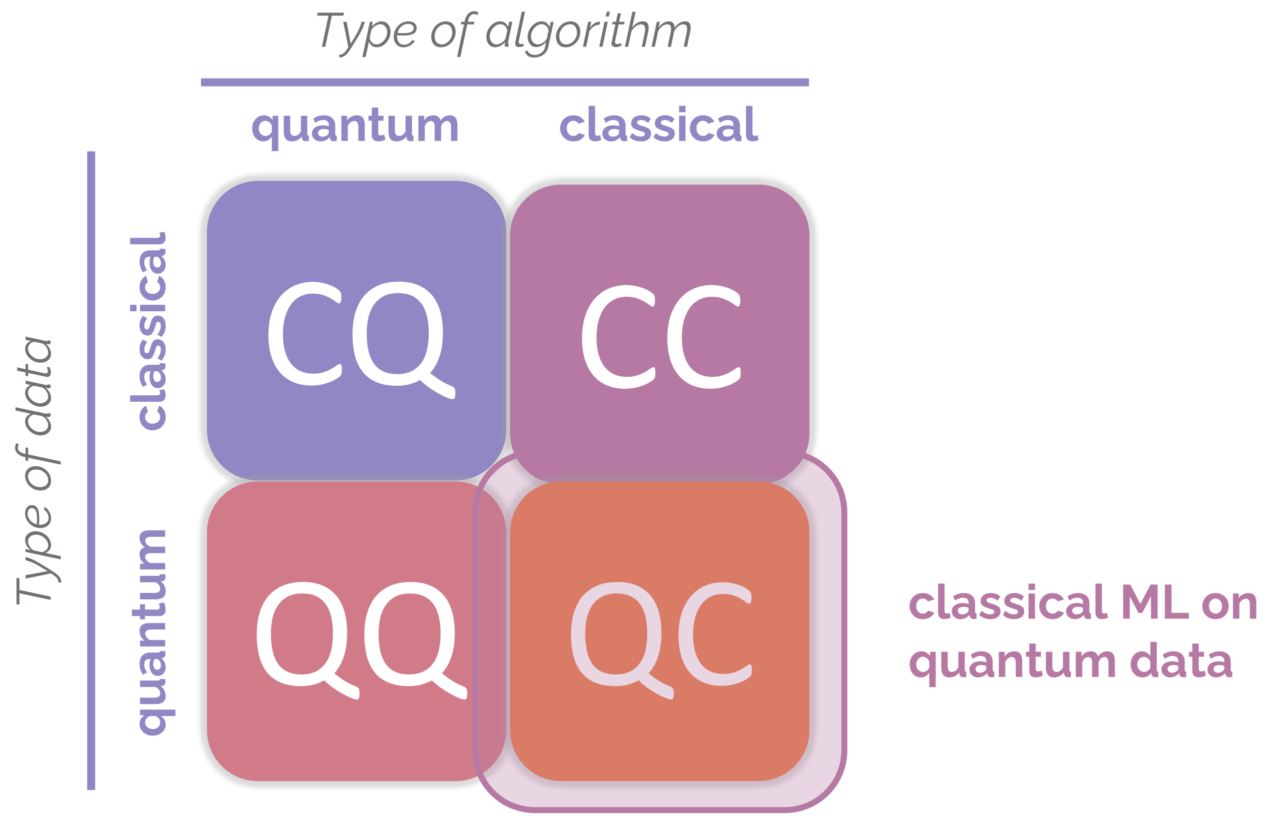
QCs can naturally solve certain problems with complex relations between inputs that can be incredibly hard for traditional, or "classical", computers. This suggest that learning models made on QC may be dramatically powerful for select applications, potentially boasting faster computation, better generalization on less data, or both.

Quantum Machine Learning



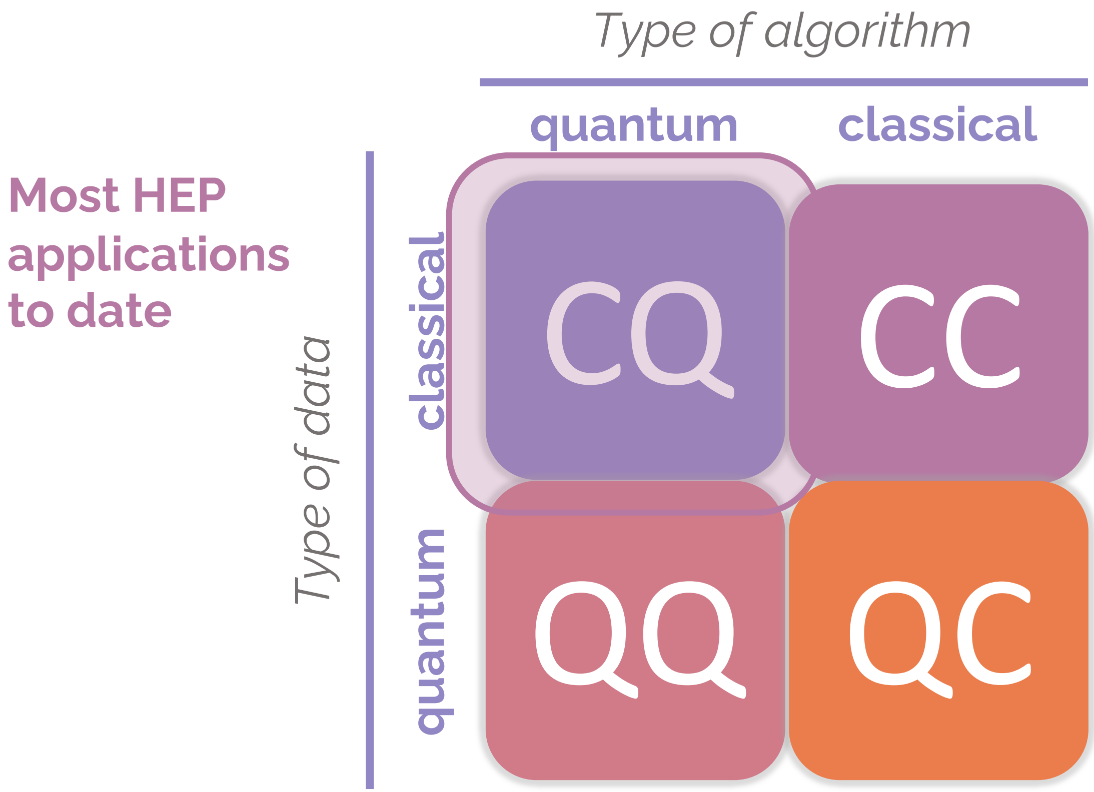
The intersection of quantum computing and ML is rich!

Quantum Machine Learning

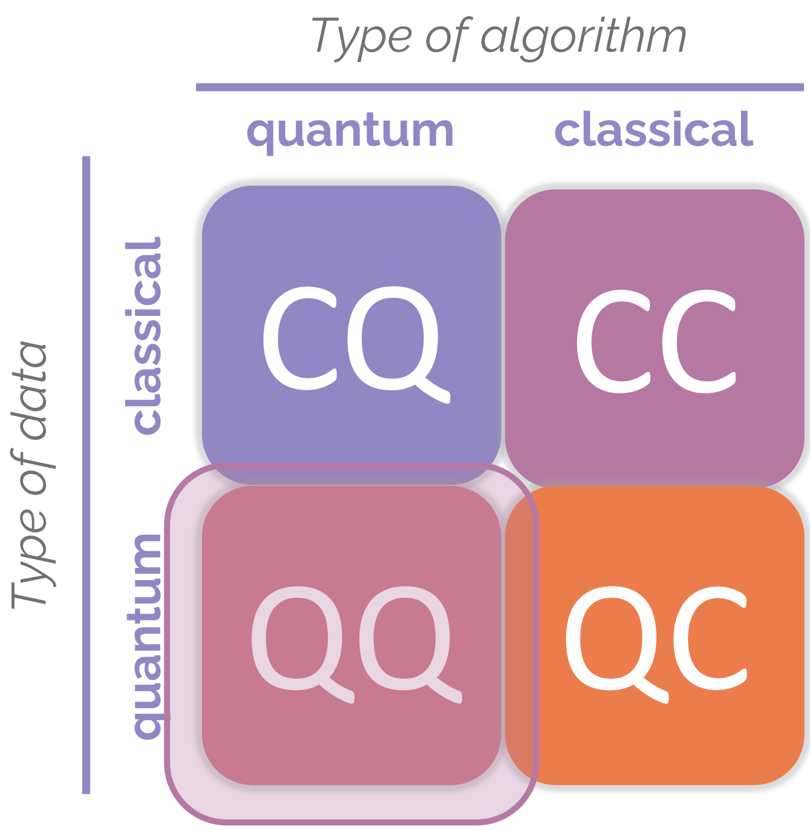


The intersection of quantum computing and ML is rich!

Quantum Machine Learning



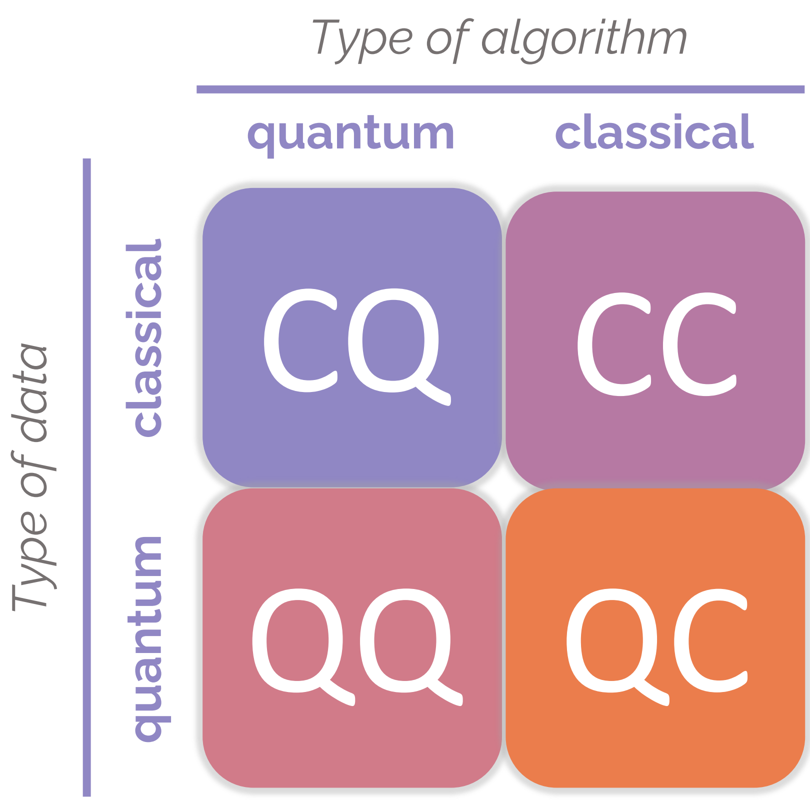
Quantum Machine Learning



- Chemical simulation
- Quantum matter simulation
- Quantum control
- Quantum networks
- Quantum metrology

Quantum Machine Learning – The Power of Data

- ★ Very **unlikely that QML will beat ML** performance on classical data.
- ★ Data generated by a quantum circuit that is hard to simulate classically is *not necessarily hard to learn for a classical model*.
- ★ Datasets that are hard for classical models and easy for quantum models to learn do exist.

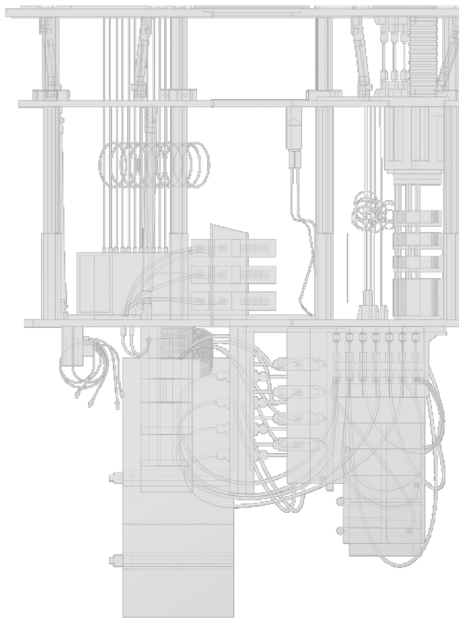


Understanding when a QC can help in a ML task depends not only on the task, but also on the data available, and a complete understanding of this must include both [*].

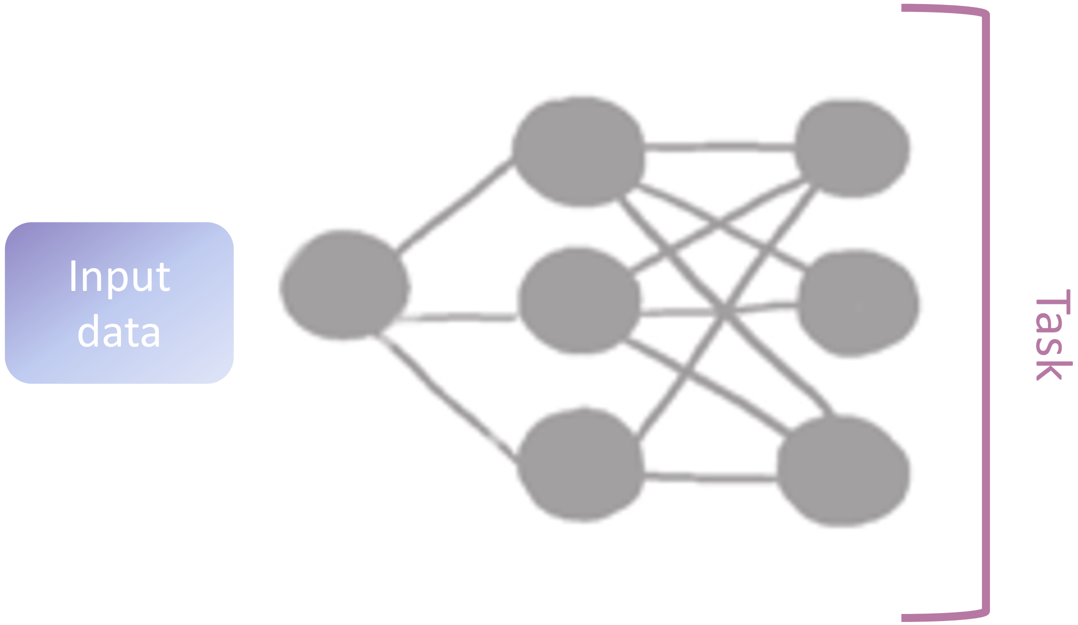
[*] Huang, HY., Broughton, M., Mohseni, M. et al. *"Power of data in quantum machine learning,"* Nat Commun 12, 2631 (2021). https://doi.org/10.1038/s41467-021-22539-9

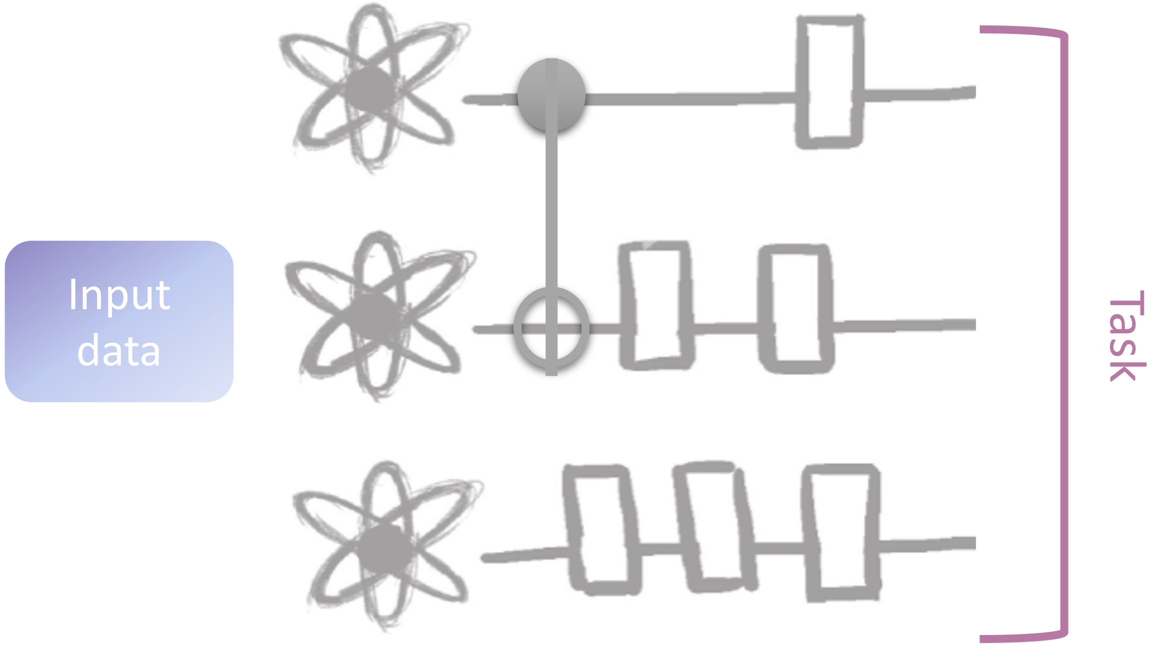
Quantum Machine Learning in the NISQ Era

- ★ Motivated by access to **cloud-based** processors and commercial applications.
- ★ Developed for deployment on **NISQ** devices.
    - o Few qubits,
    - o Noisy,
    - o Low gate fidelity.
- ★ Applications in **Quantum Machine Learning (QML)** spurred by the release of Xanadu's PennyLane / Google's Tensorflow.
- ★ **Co-design**:
    - o *Algorithmic development/research is adapting to match the pace of hardware development.*
- ★ Hybrid frameworks to leverage benefits of both classical and quantum computing - ***variational quantum circuits***.



Parameterized Quantum Circuits as ML Models



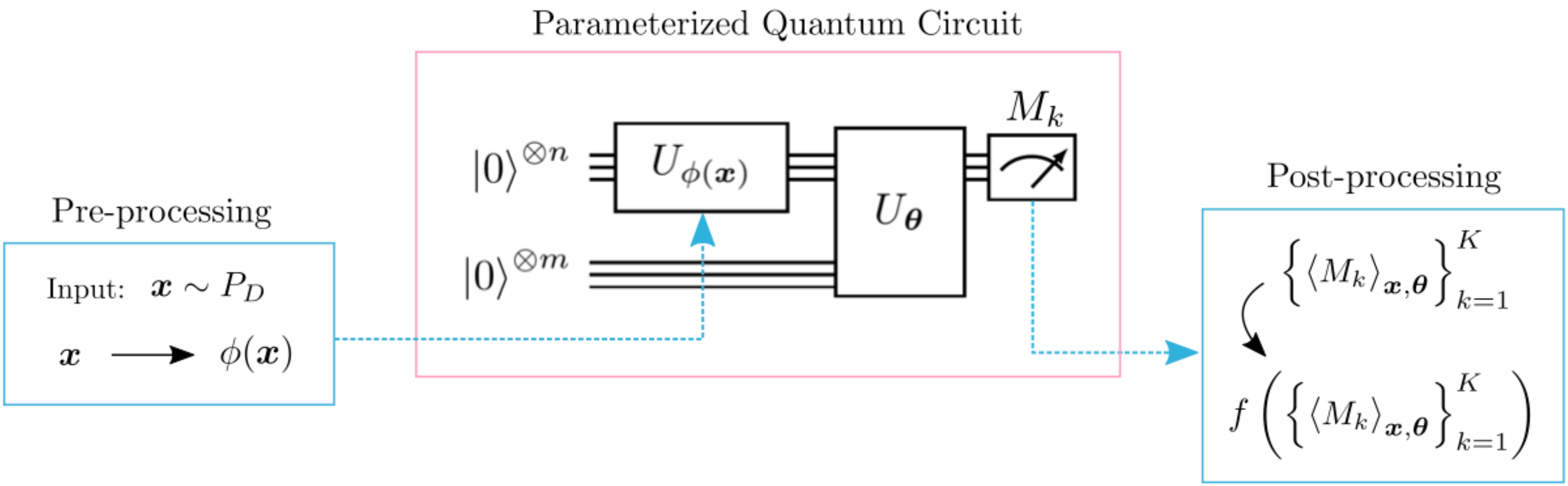


In both cases, learning describes the process of iteratively updating the model's parameters towards a goal

Benedetti, arXiv:1906.07682

Parameterized Quantum Circuits as ML Models

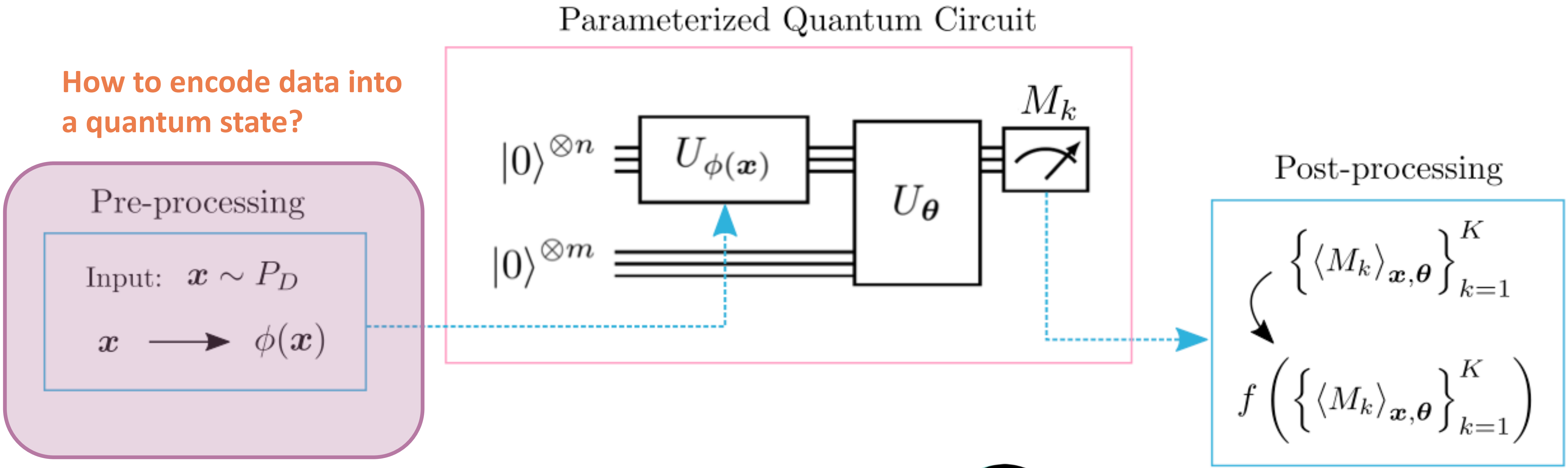
Benedetti, arXiv:1906.07682



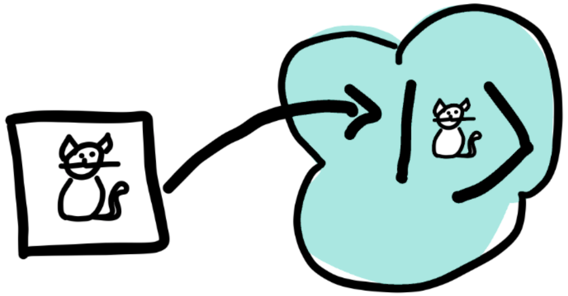
Parameterized Quantum Circuits as ML Models

Benedetti, arXiv:1906.07682

How to encode data into a quantum state?



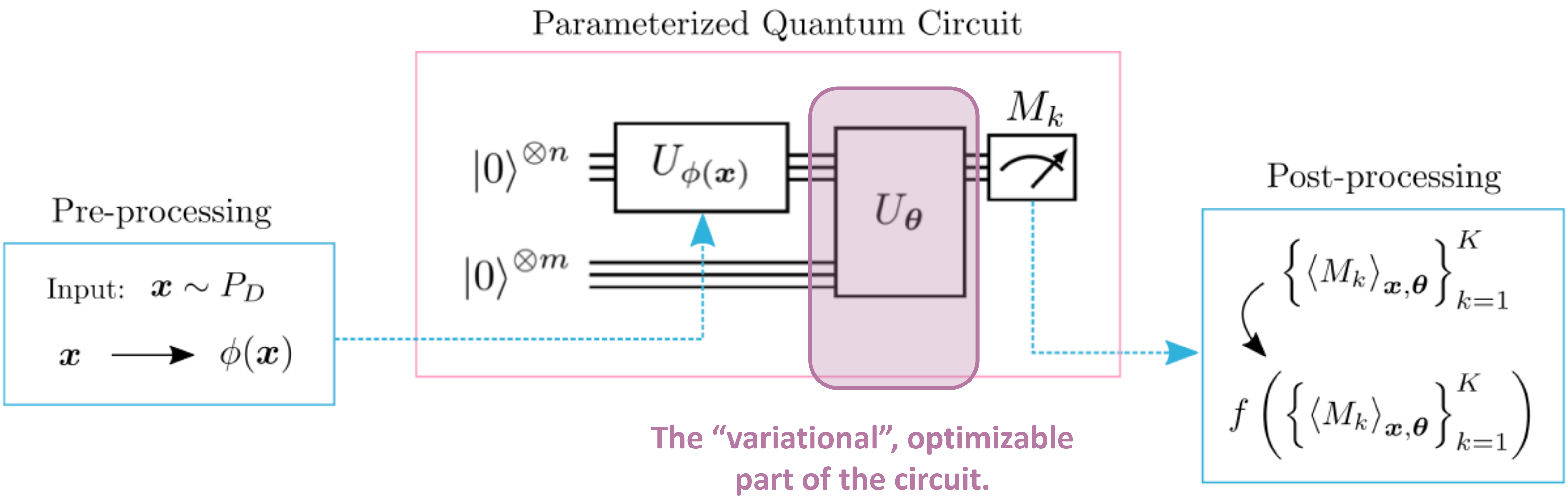
1. Start from a feature vector ***x***.
2. <u>Optional:</u> dimensionality reduction, PCA, etc.
3. Quantum embedding through a quantum feature map: *Basis embedding, amplitude embedding.*



- ★ Havlicek, et al, arXiv:1804.11326
- ★ Schuld, Killoran, arXiv:1803.07128
- ★ Lloyd, Schuld, et al, arXiv:2001.03622

Parameterized Quantum Circuits as ML Models

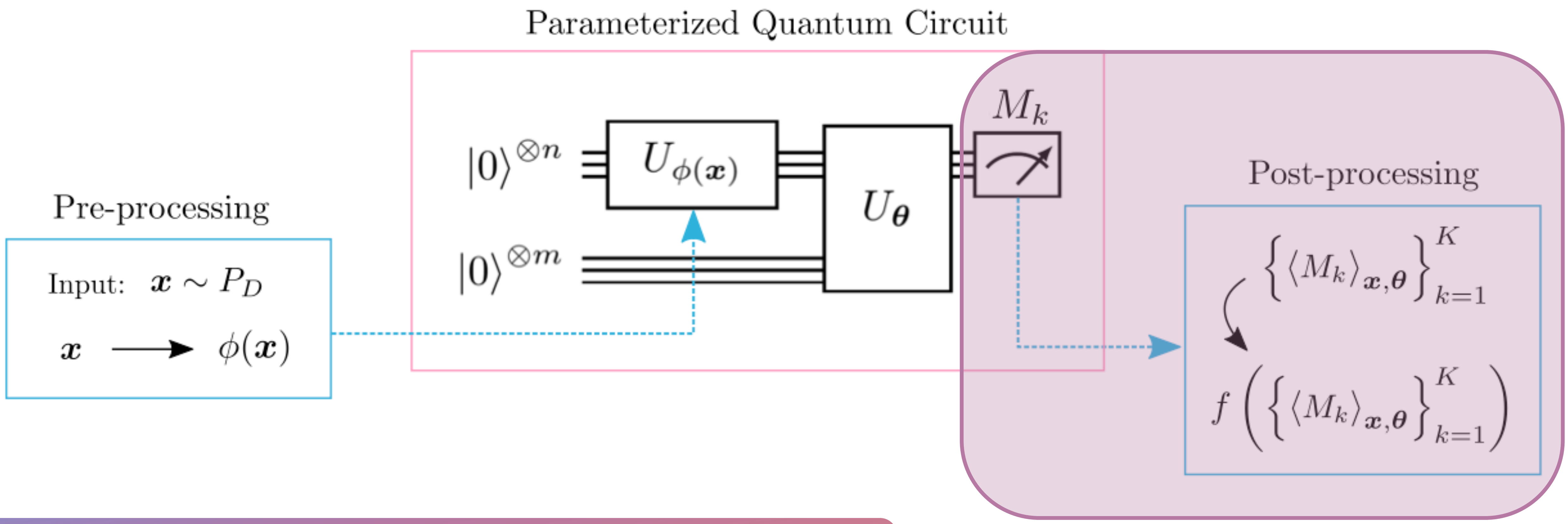
Benedetti, arXiv:1906.07682



The "guess" or trial function is the unitary U parameterized by a set of free parameters θ that will be updated during training.

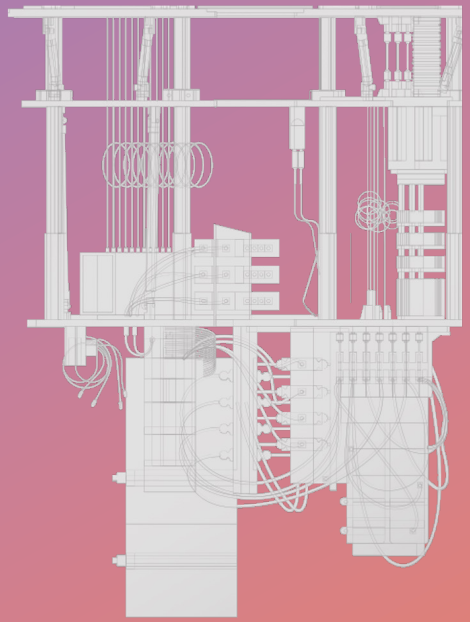
Parameterized Quantum Circuits as ML Models

Benedetti, arXiv:1906.07682



Quantum information is turned back into classical information by evaluating the expectation value of an observable, or measurement.

The measurement output is then used to construct a decision function, a probability distribution, a boundary, etc.



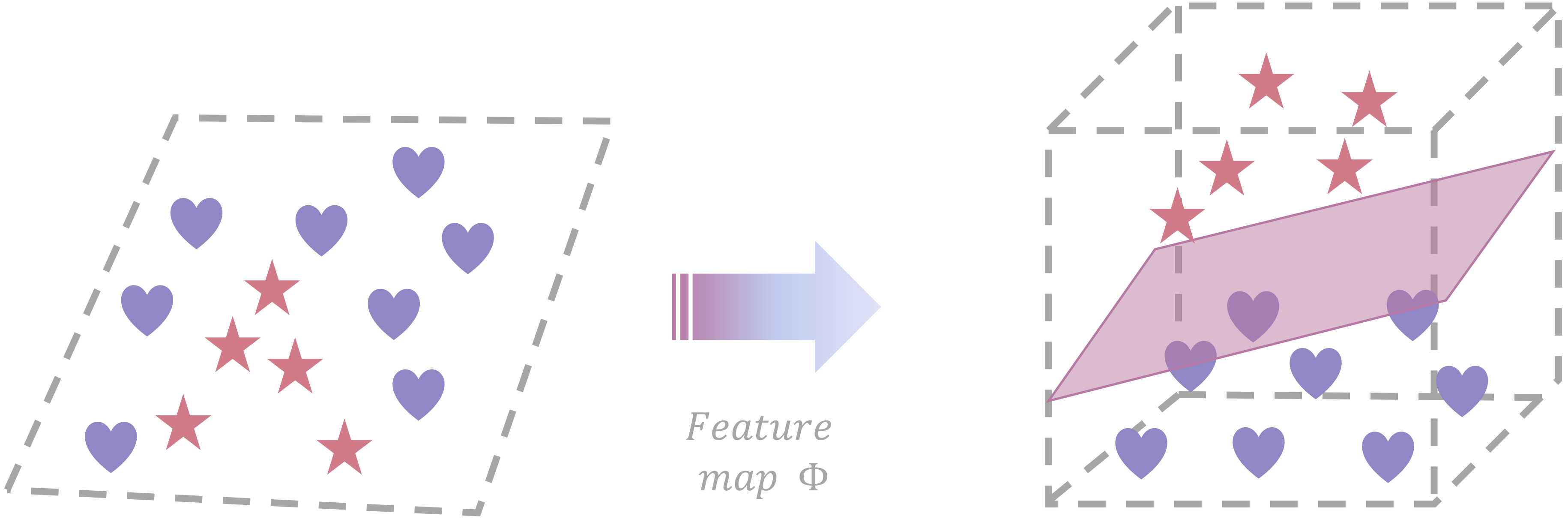
Applications

Supervised Learning with Kernel-based Quantum Models

Quantum machine learning models for supervised learning and kernel methods are based on a similar principle.

A high-level overview, for more details check references: *arXiv:2101.11020, Phys. Rev. Lett. 122, 040504 (2019), Nature. vol. 567, pp. 209-212 (2019)*

Kernel Trick



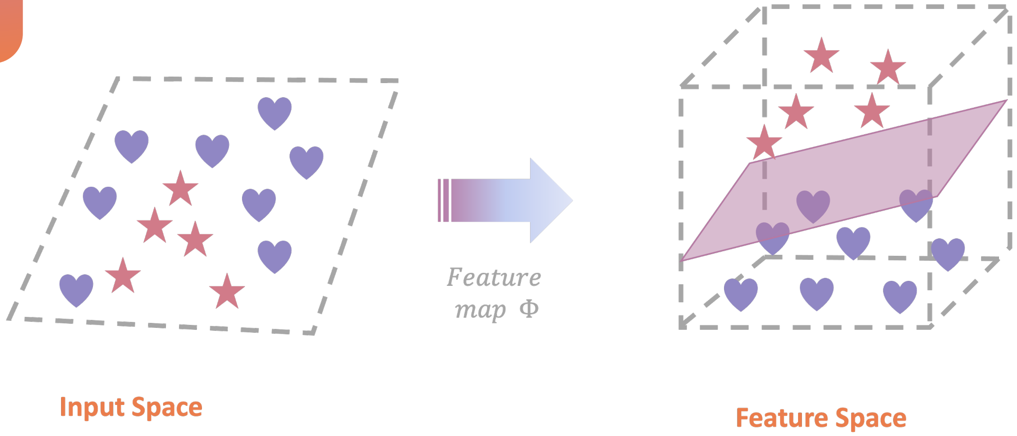
Input Space

Feature Space

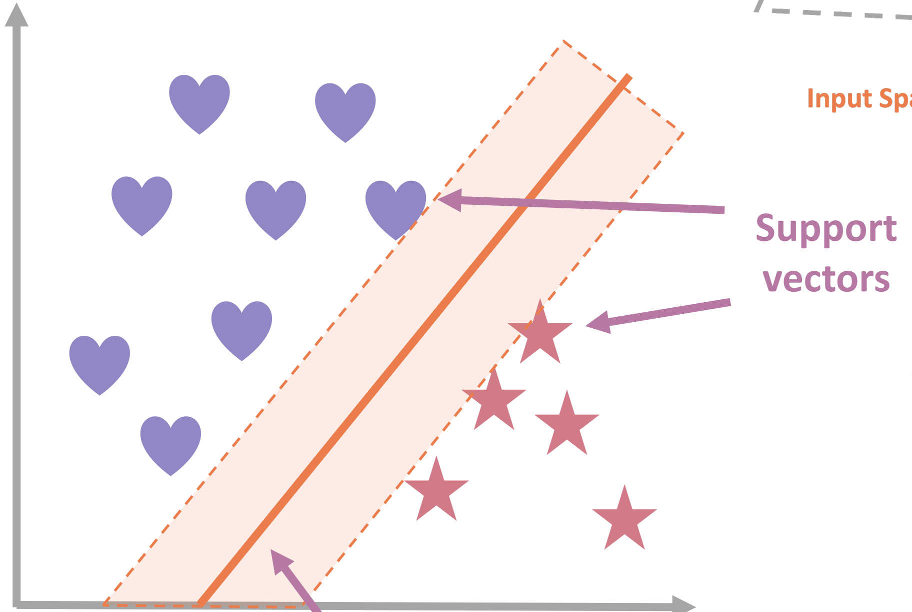
Supervised Learning with Kernel-based Quantum Models

Quantum machine learning models for supervised learning and kernel methods are based on a similar principle.

A high-level overview, for more details check references: *arXiv:2101.11020, Phys. Rev. Lett. 122, 040504 (2019), Nature. vol. 567, pp. 209-212 (2019)*



Support Vector Machine



$$\tilde{m}(\vec{s}) = sign\left(\sum_{i=1}^{t} y_i \alpha_i^{*} K(\vec{x}_i, \vec{s}) + b\right)$$

To optimize a loss function of the form

$$L_D(\alpha) = \sum_{i=1}^{t} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{t} y_i y_j \alpha_i \alpha_j \boxed{K(\vec{x}_i, \vec{x}_j)}$$

Kernel

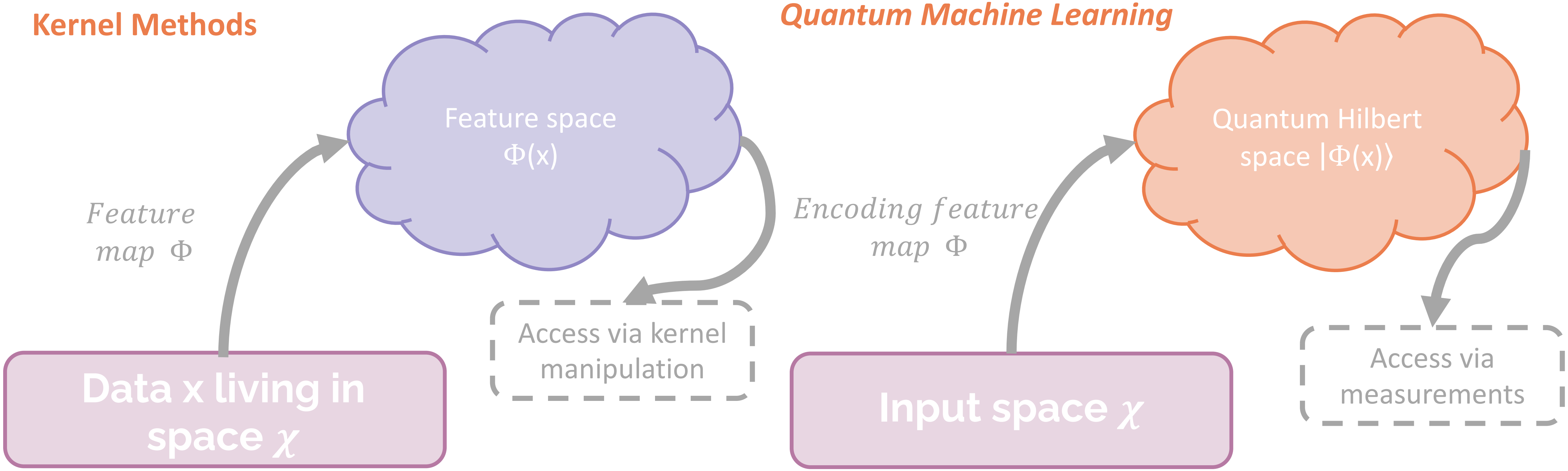
Supervised Learning with Kernel-based Quantum Models

Quantum machine learning models for supervised learning and kernel methods are based on a similar principle.

A high-level overview, for more details check references: *arXiv:2101.11020, Phys. Rev. Lett. 122, 040504 (2019), Nature. vol. 567, pp. 209-212 (2019)*

Kernel Methods

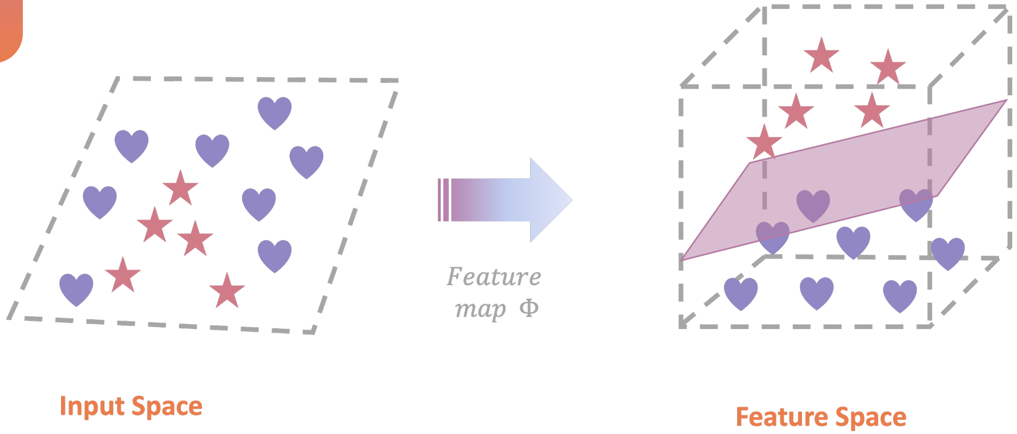
Quantum Machine Learning



Supervised Learning with Kernel-based Quantum Models

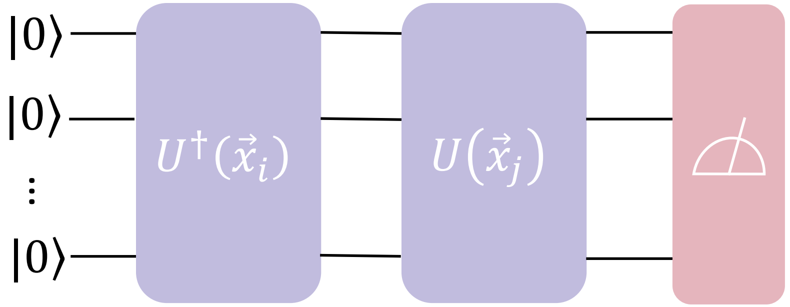
Quantum machine learning models for supervised learning and kernel methods are based on a similar principle.

A high-level overview, for more details check references: *arXiv:2101.11020, Phys. Rev. Lett. 122, 040504 (2019), Nature. vol. 567, pp. 209-212 (2019)*



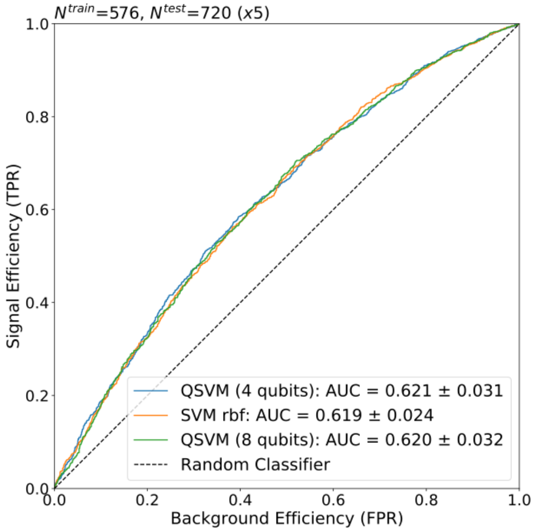
Construct a kernel matrix of the form

Support Vector Machine

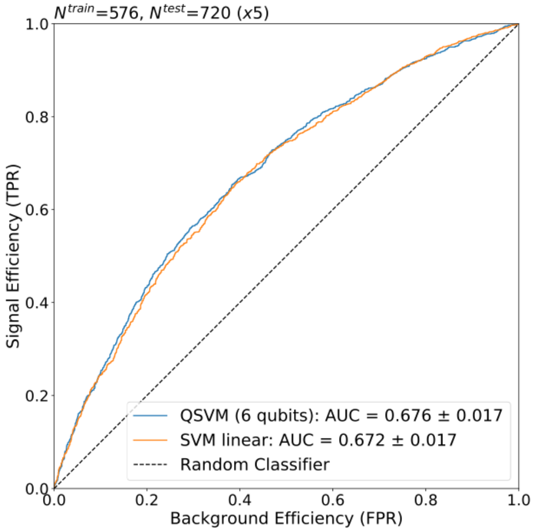


$$K_{ij} = \left|\langle 0|U^\dagger(\vec{x}_i)U(\vec{x}_j)|0\rangle\right|^2$$

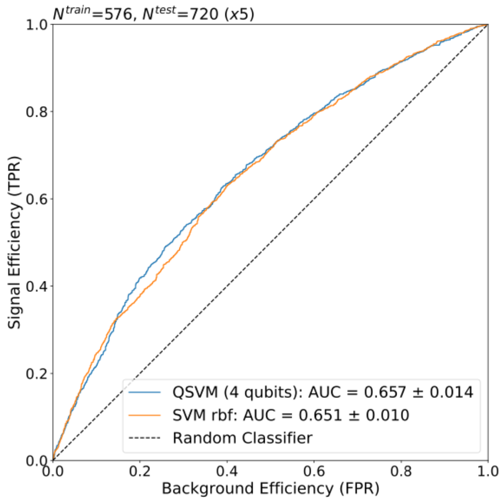
Supervised Learning with Kernel-based Quantum Models



(a) Models trained on the AE latent space features (16).



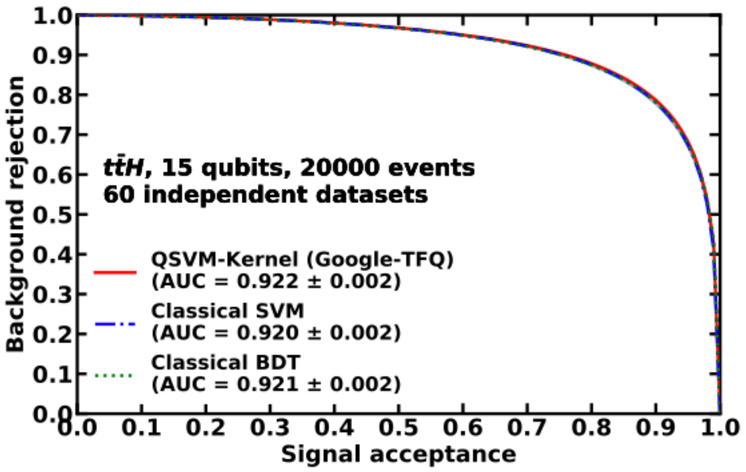
(b) Models trained on the original input features (67), discarding the 3 least informative ones (64).



(c) Models trained on 16 selected features of the input space according to their individual AUC values.

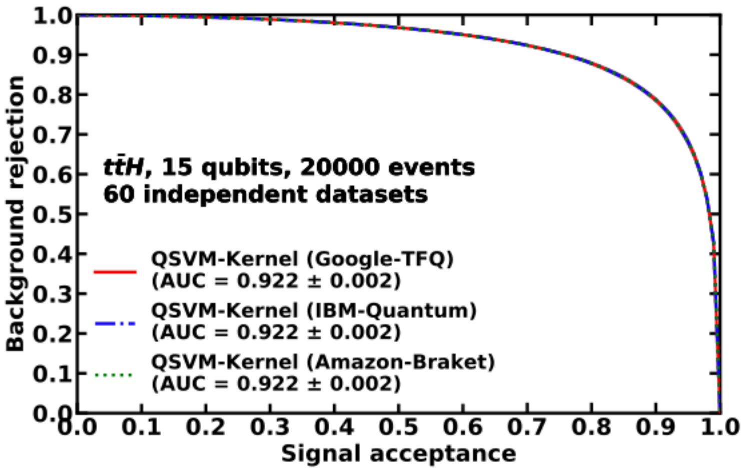
$t\bar{t}H(b\bar{b})$classification problem on 4,6, and 8 qubit QSVM

"Higgs analysis with quantum classifiers", Belis, Gonzalez-Castillo, et al., arXiv:2104.07692 (2021)



(a)

$t\bar{t}H$classification problem on 20 qubit QSVM in simulation and 15 qubits in HW



(b)

"Application of Quantum Machine Learning Using the Quantum Kernel Algorithm on High-Energy Physics Analysis at the LHC", Wu, Sun, Guan, et al., arXiv:2104.05059 (2021)

Supervised Learning with Kernel-based Quantum Models

Signal/background classification problem on 6 qubit QSVM

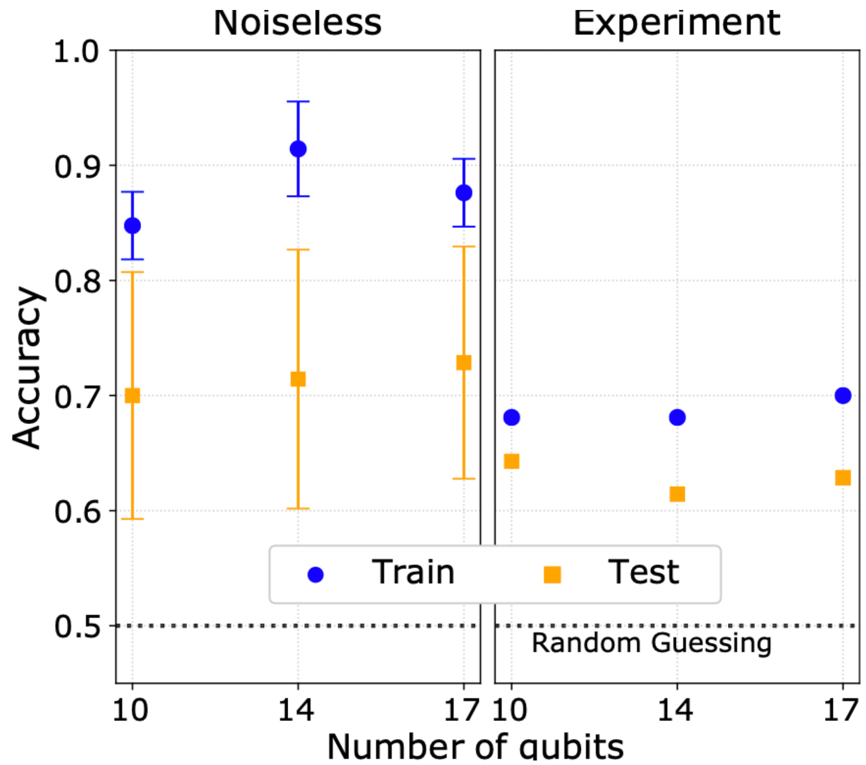
Table 1 Average results from 10 random dataset samples obtained by classically simulating various encoding circuits using Qiskit *statevector_simulator* with 60,000 training events and 10,000 testing events in each sample

Encoding circuit	Accuracy	AUC
Combinatorial encoding	0.762	0.822
Separate particle encoding	0.776	0.835
Bloch sphere encoding	0.764	0.836
Separate particle with bloch	0.771	0.848
Classical RBF kernel SVM	0.728	0.793
XGBoost	0.590	0.621

The uncertainty on each of the mean values stated is ± 0.001

"Quantum Support Vector Machines for Continuum Suppression in B meson Decays", Heredge, Hill, Hollenberg, Sevior, Computing and Software for Big science (2021) 5:27

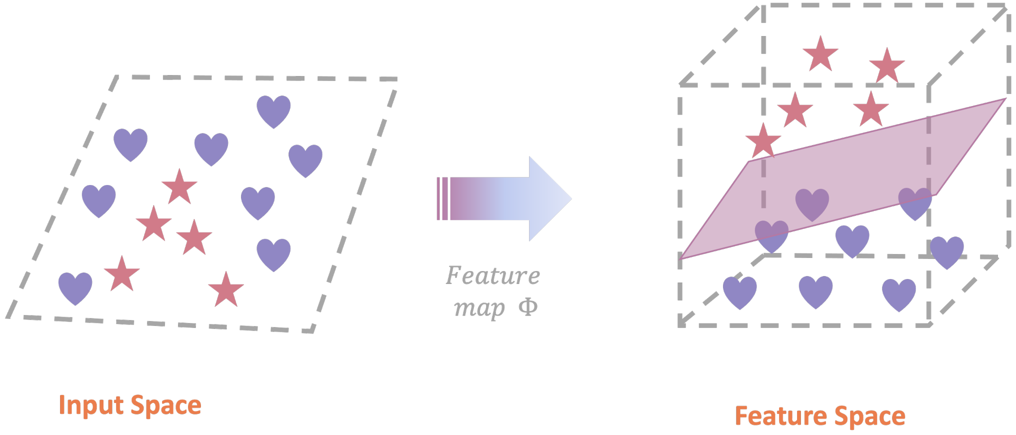
Supernovae classification with QSVM at Google's Sycamore processor

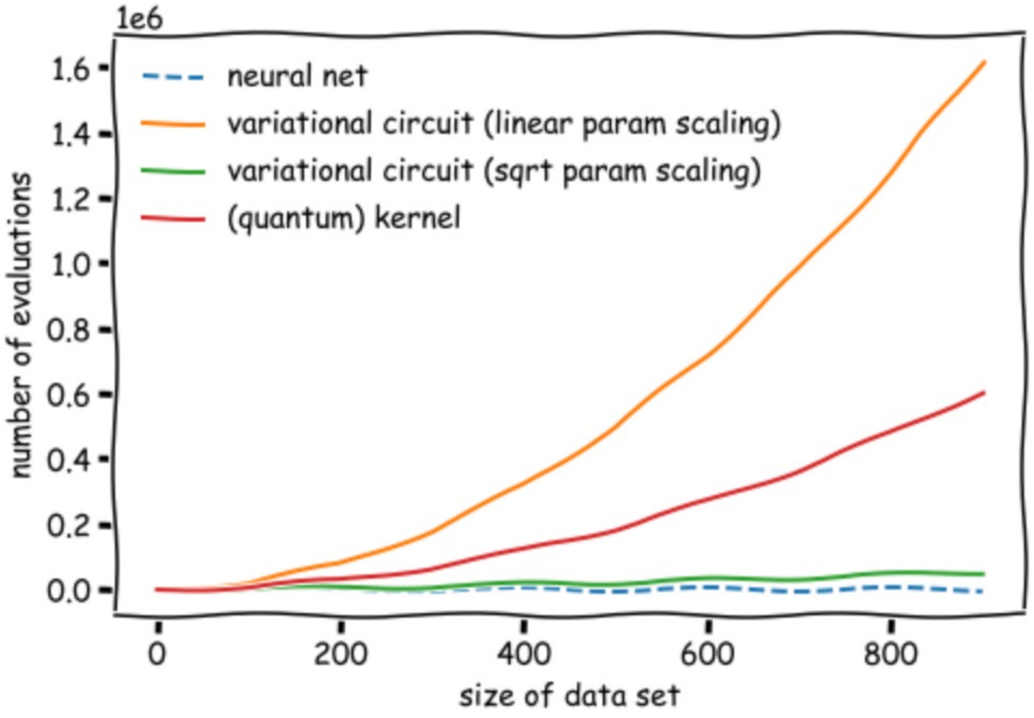


"Machine Learning of high-dimensional data on a noisy quantum processor", Peters, Caldeira, Ho, et al., npj Quantum Information (2021) 7:161

On Kernel-based Quantum Models...

- ★ **Kernel methods** are essentially based on feature maps that allow for *classification on a higher-dimensional space*.
- ★ Quantum machine learning models based on kernel methods might provide an **advantage** when *the kernel is hard to estimate classically*.
  - ★ But... the efficiency of kernel-based methods compared to variational circuits depends on the number of parameters used in a variational model.
  - ★ Meaning that for specific applications, if the number of parameters scales linearly, most likely your application is better suited for a VQC.



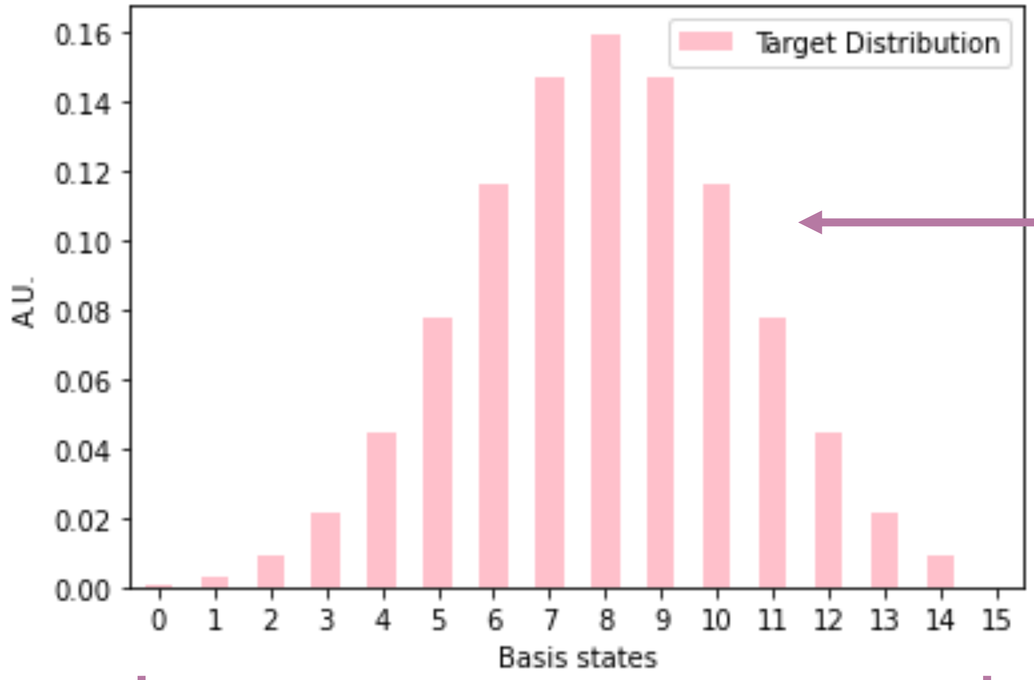


Checkout PennyLane tutorial on "Kernel-based training of a quantum models with scikit-learn" https://pennylane.ai/qml/demos/tutorial_kernel_based_training.html

Unsupervised Generative Modeling

Quantum Circuit Born Machines are generative models which represent the probability distribution of a classical dataset as quantum pure states

A high-level overview, for more details check references: *Phys. Rev. A 98, 062324 (2018), arXiv:2203.03578*



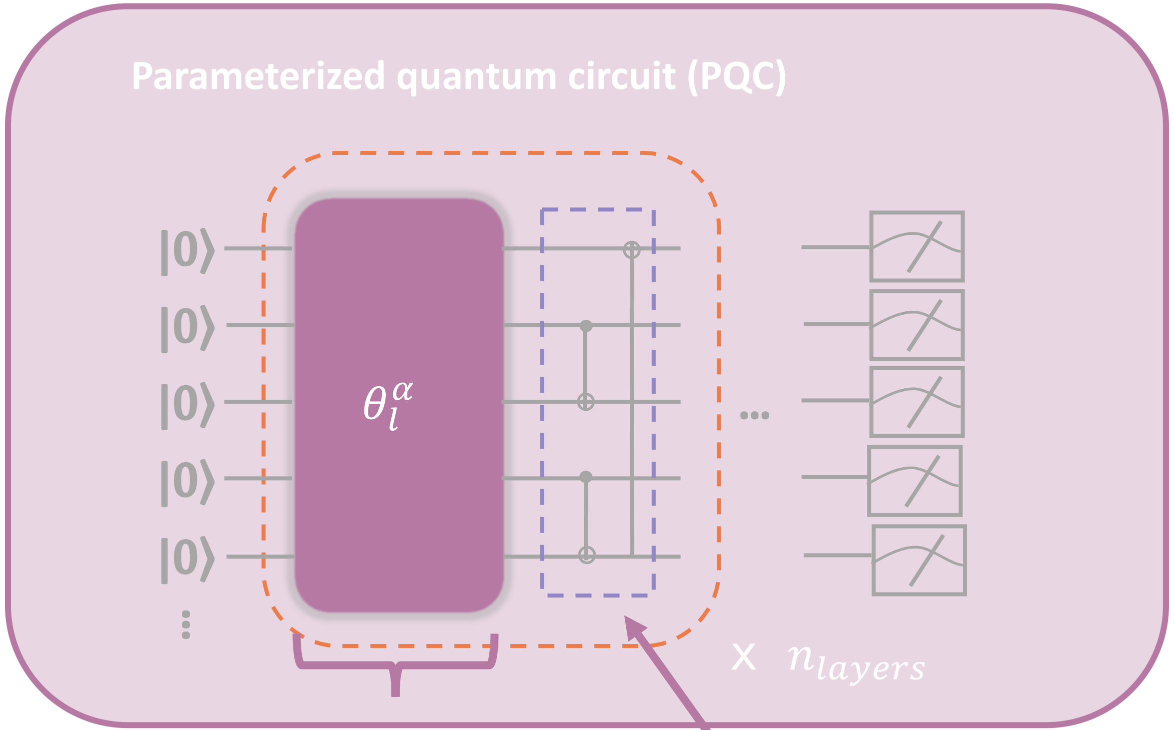
Discretized Gaussian probability distribution over $2^{n_{qubits}}$ basis states or bins.

$2^{n_{qubits}}$ basis states or bins, i.e., 0000, 0001, 0010, etc.

Unsupervised Generative Modeling

Quantum Circuit Born Machines are generative models which represent the probability distribution of a classical dataset as quantum pure states

A high-level overview, for more details check references: *Phys. Rev. A 98, 062324 (2018), arXiv:2203.03578*



Block of rotation gates, with tunable parameters

Block of entangling gates

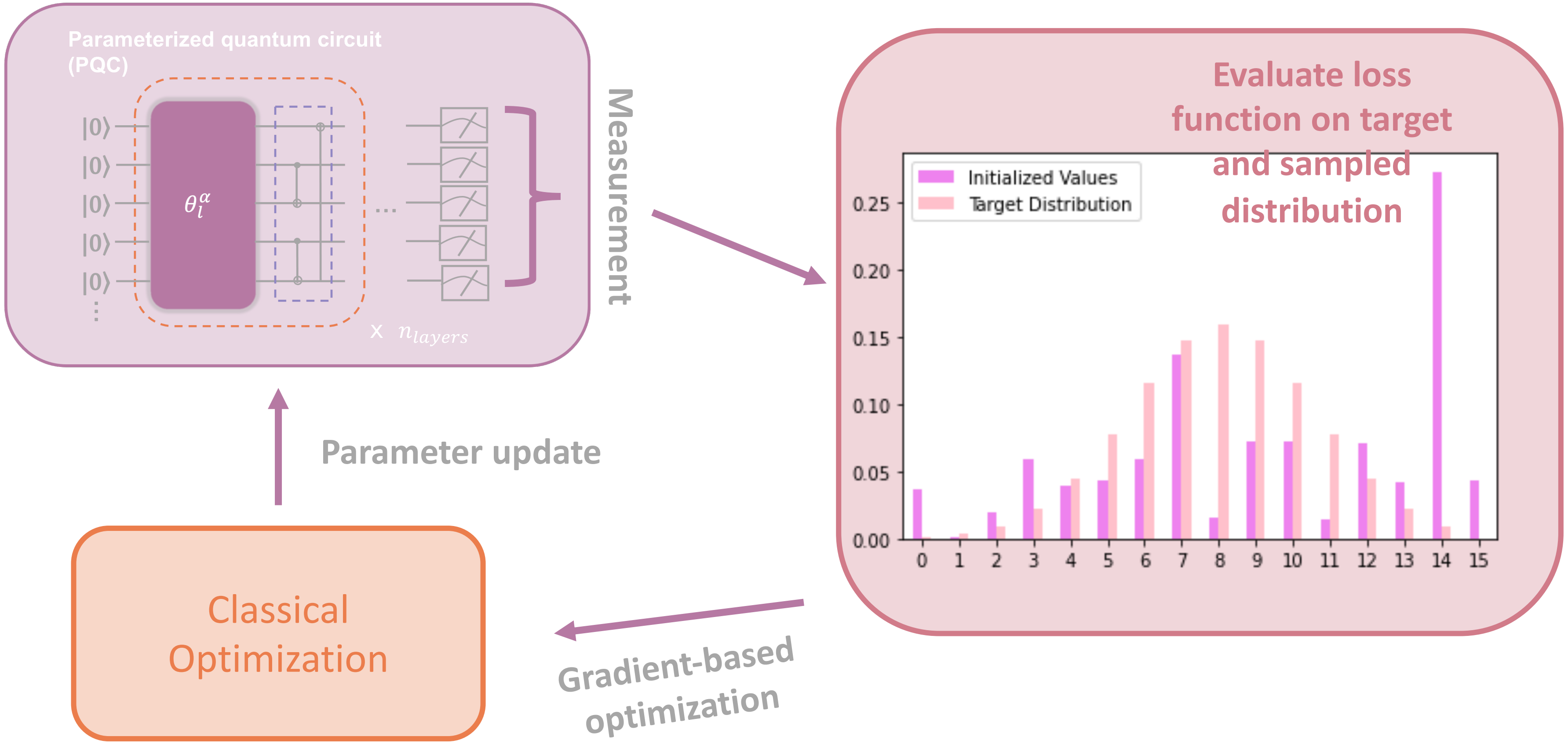
A PQC consists of layers or blocks of rotational and entangling gates that can be repeated to maximize the circuit's expressibility.

Discussion on smart Ansatz choices in a few slides ☺

Unsupervised Generative Modeling

Quantum Circuit Born Machines are generative models which represent the probability distribution of a classical dataset as quantum pure states

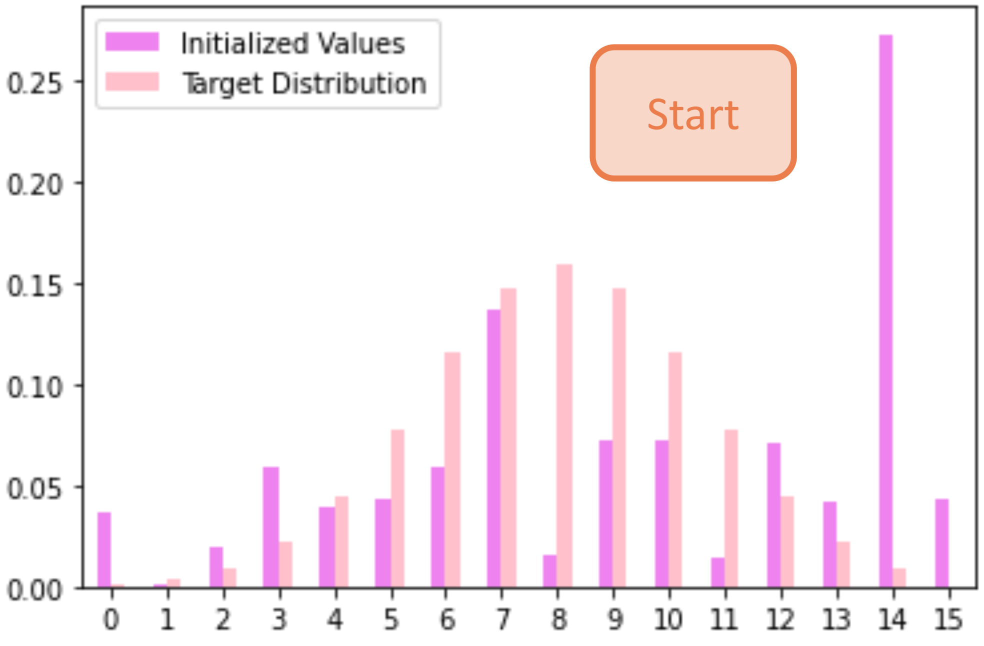
A high-level overview, for more details check references: *Phys. Rev. A 98, 062324 (2018), arXiv:2203.03578*

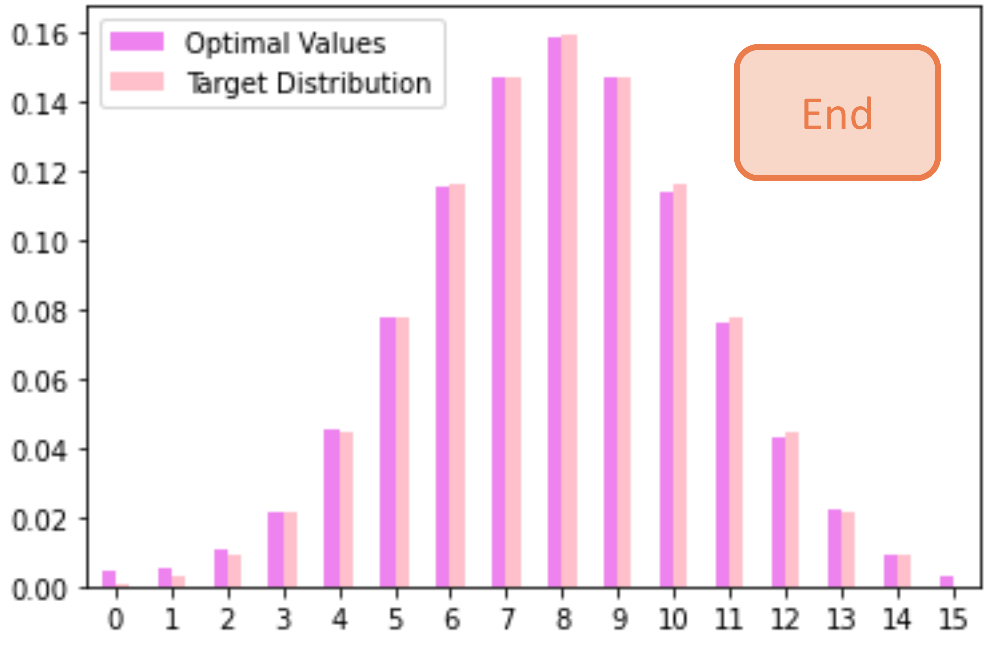


Unsupervised Generative Modeling

Quantum Circuit Born Machines are generative models which represent the probability distribution of a classical dataset as quantum pure states

A high-level overview, for more details check references: *Phys. Rev. A 98, 062324 (2018), arXiv:2203.03578*



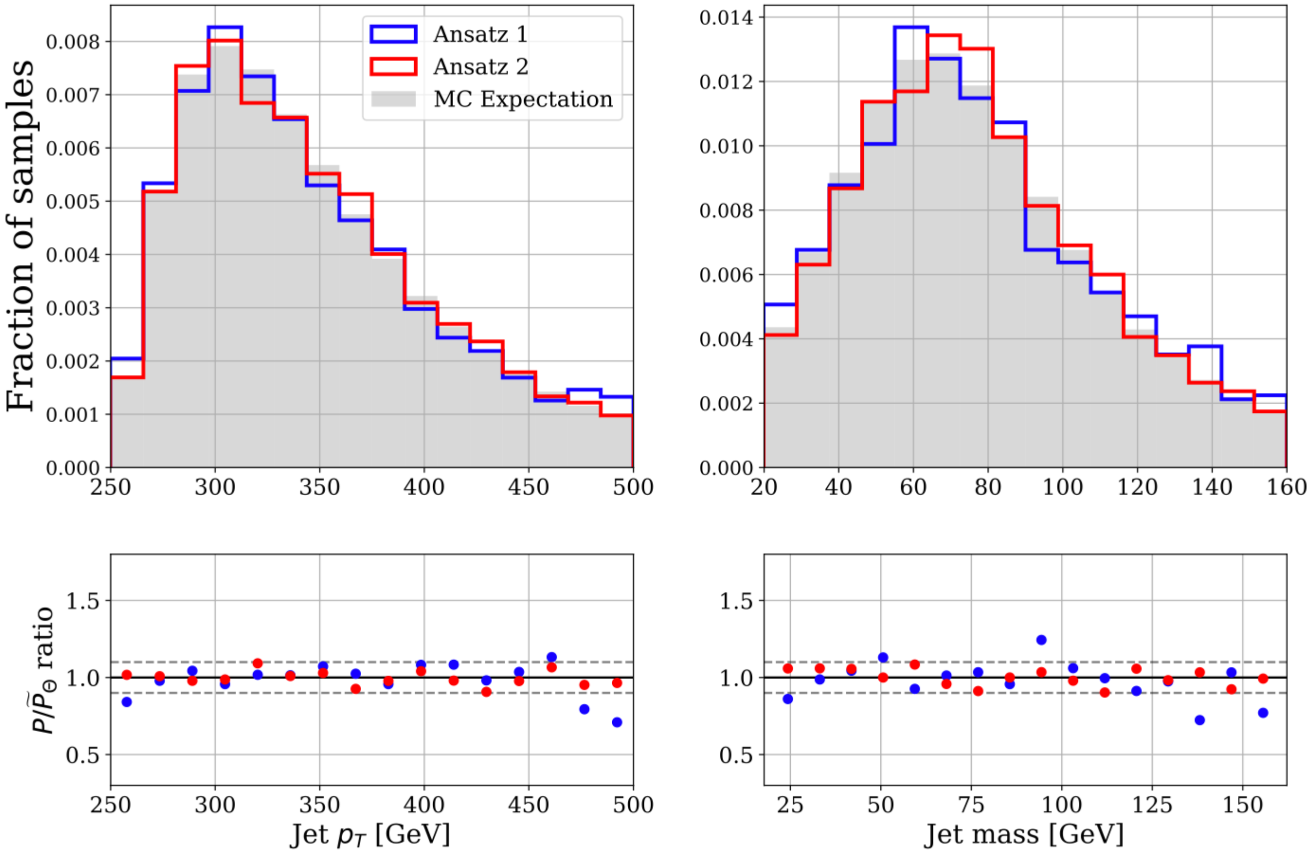


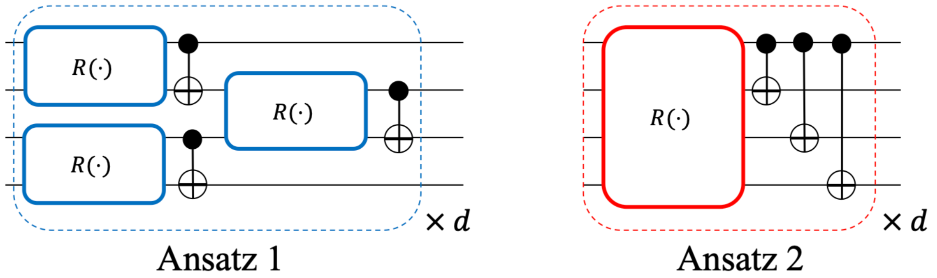
QCBM trained on 4 qubits using cosine distance metric optimized using gradient-based optimizer (Adam). Hyperparameters: learning rate = 0.1, number of steps = 100 , 8192 shots.

Unsupervised Generative Modeling

Can QCBM's learn joint distributions?

Yes!





(a) Monte Carlo (Ground Truth)

	p_T	mass
p_T	-	0.2
mass	0.2	-

(b) Ansatz 1

	p_T		mass	
	$\vert 0\rangle^{\otimes 8}$	$\vert \Phi^+\rangle^{\otimes 4}$	$\vert 0\rangle^{\otimes 8}$	$\vert \Phi^+\rangle^{\otimes 4}$
p_T	-		0.19	0.12
mass	0.19	0.12	-	

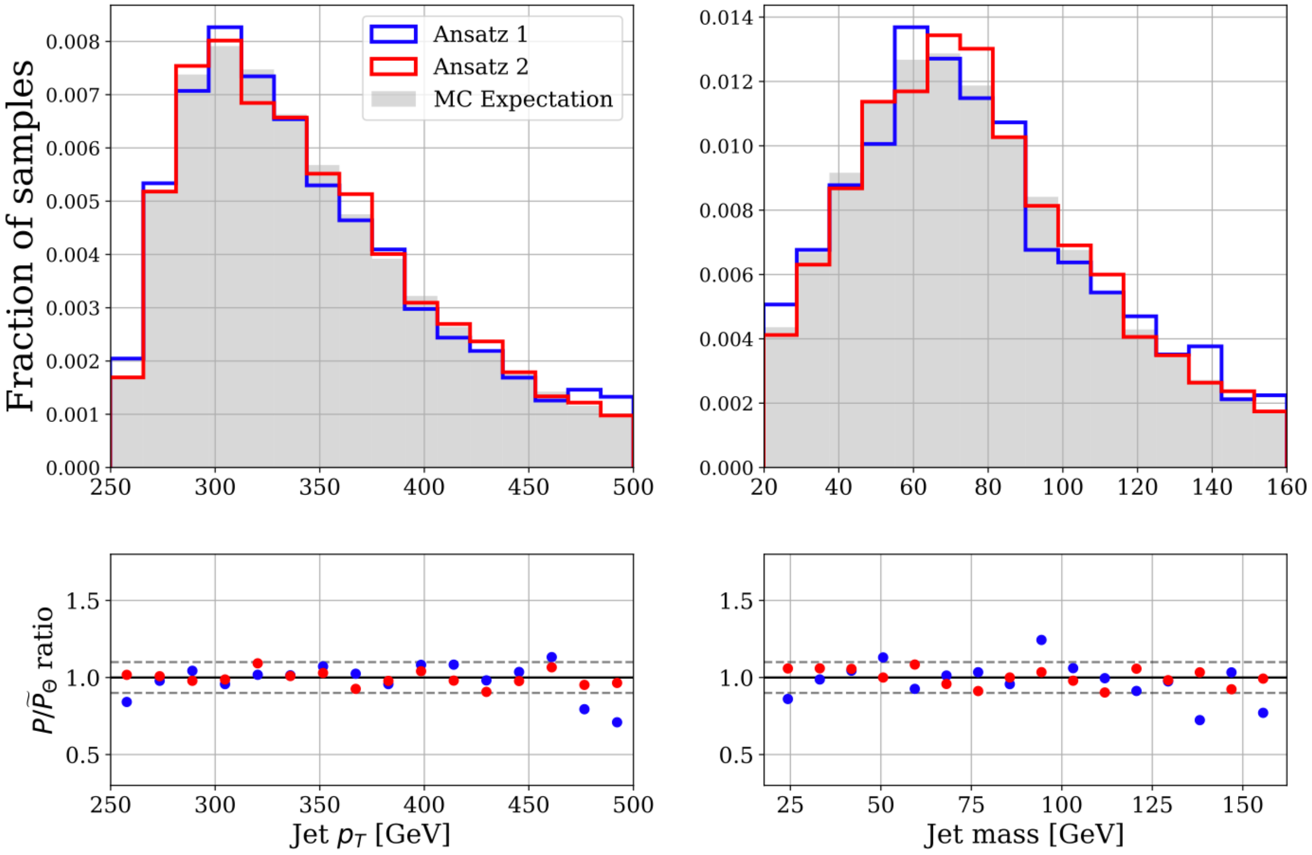
(c) Ansatz 2

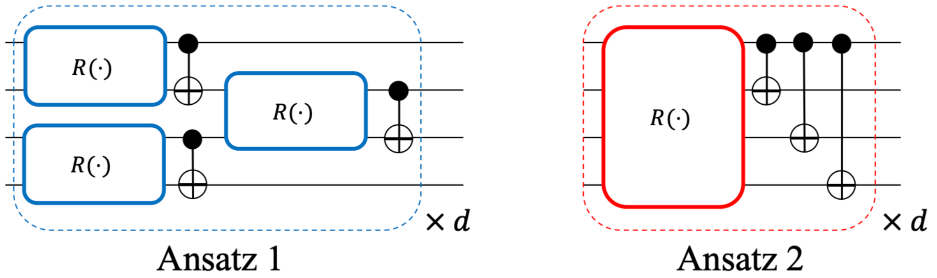
	p_T		mass	
	$\vert 0\rangle^{\otimes 8}$	$\vert \Phi^+\rangle^{\otimes 4}$	$\vert 0\rangle^{\otimes 8}$	$\vert \Phi^+\rangle^{\otimes 4}$
p_T	-		-1.0e-3	-9.1e-3
mass	-1.0e-3	-9.1e-3	-	

Delgado, A., Hamilton, K. E., "Unsupervised Quantum Circuit Learning in High-Energy Physics" arXiv:2203.03578

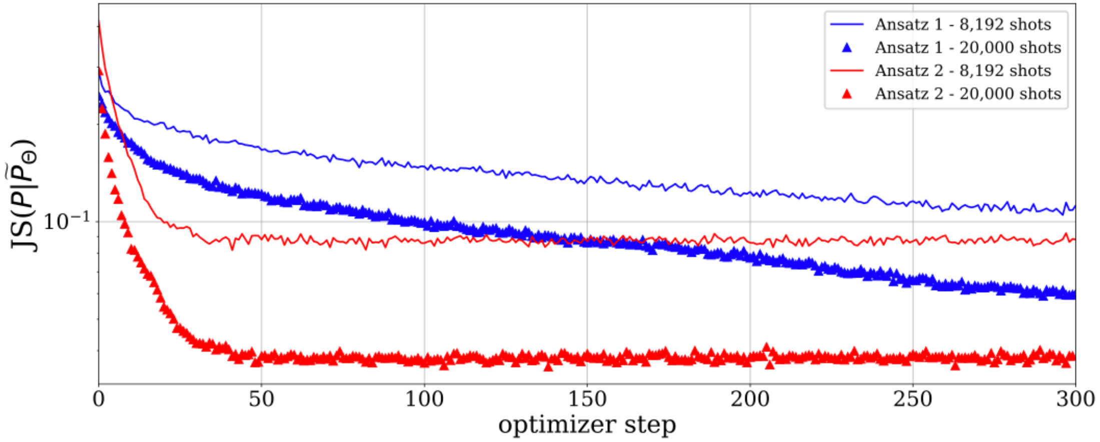
Unsupervised Generative Modeling

Can QCBM’s learn joint distributions?





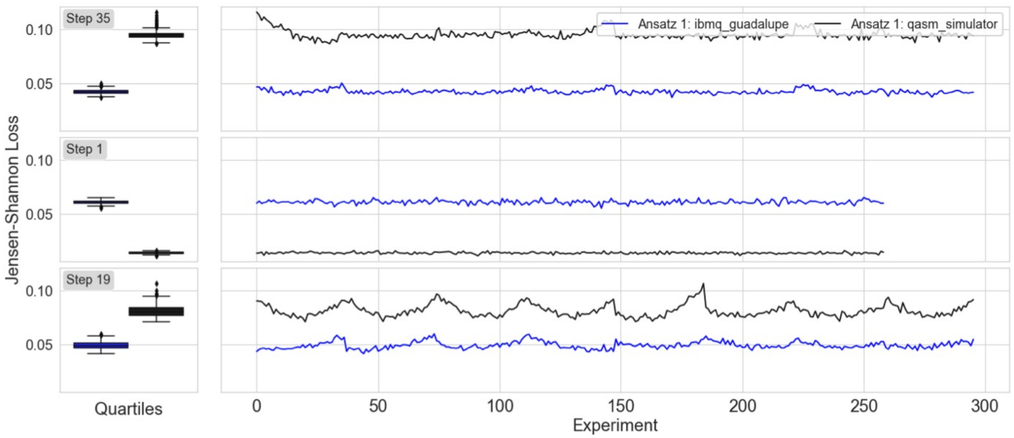
The effect of number of shots in training

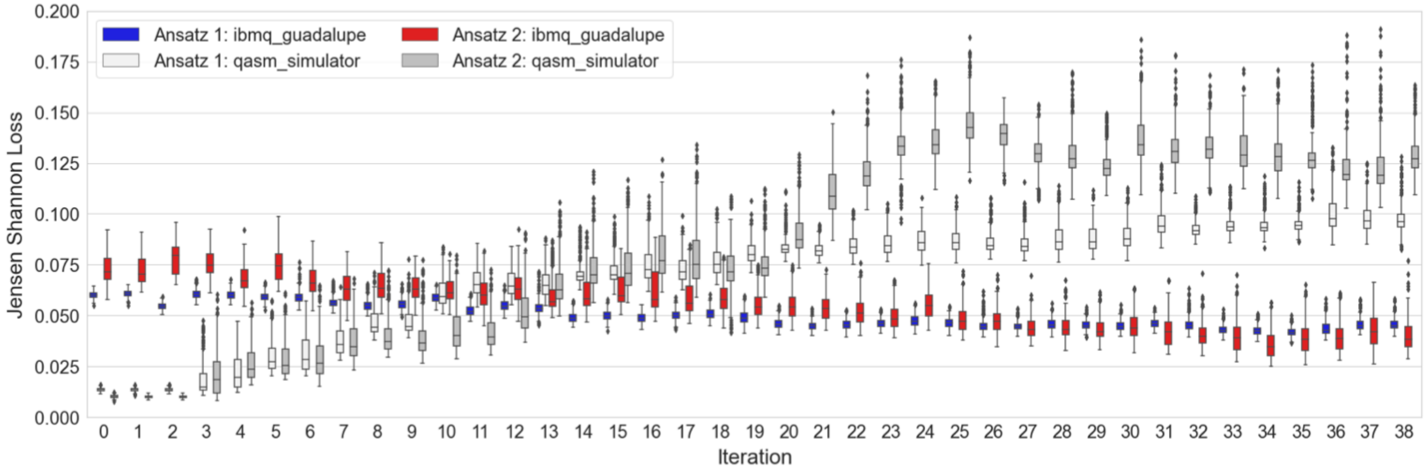


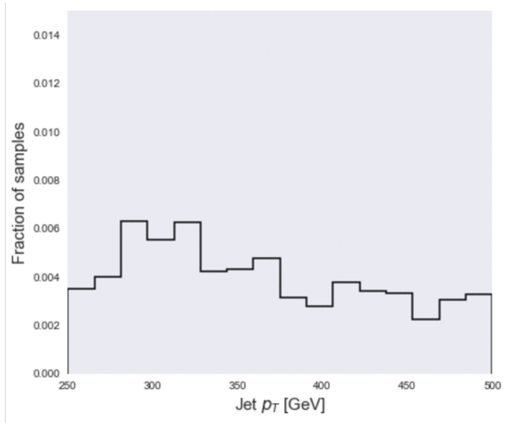
Delgado, A., Hamilton, K. E., “Unsupervised Quantum Circuit Learning in High-Energy Physics” arXiv:2203.03578

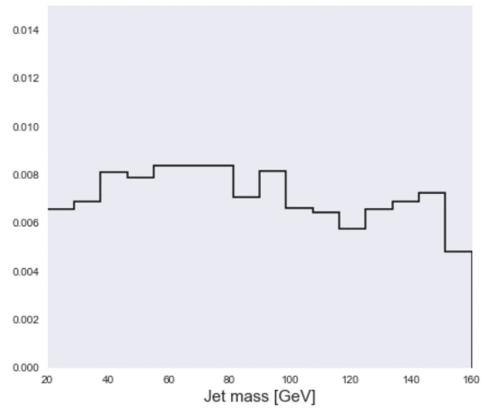
Unsupervised Generative Modeling

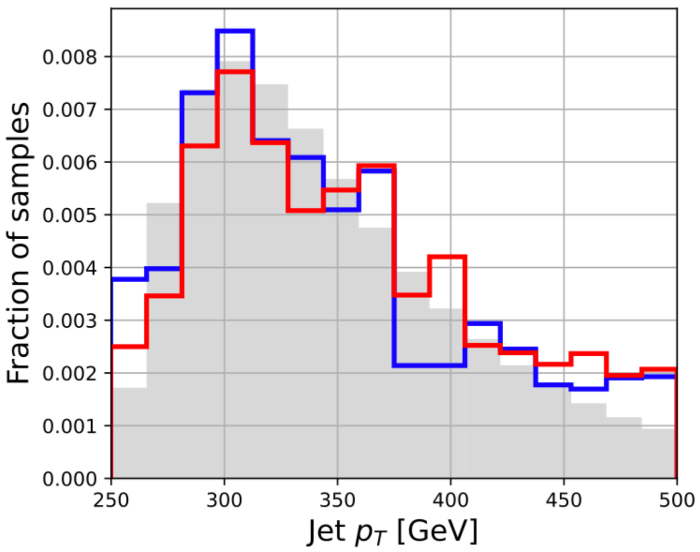
What about hardware noise?

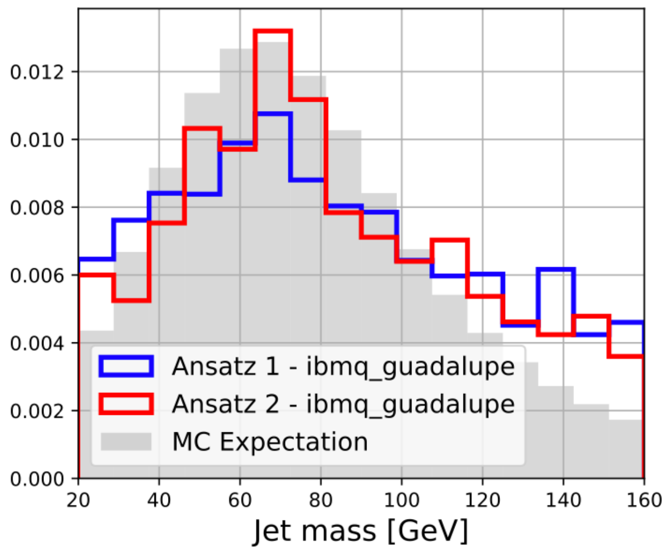








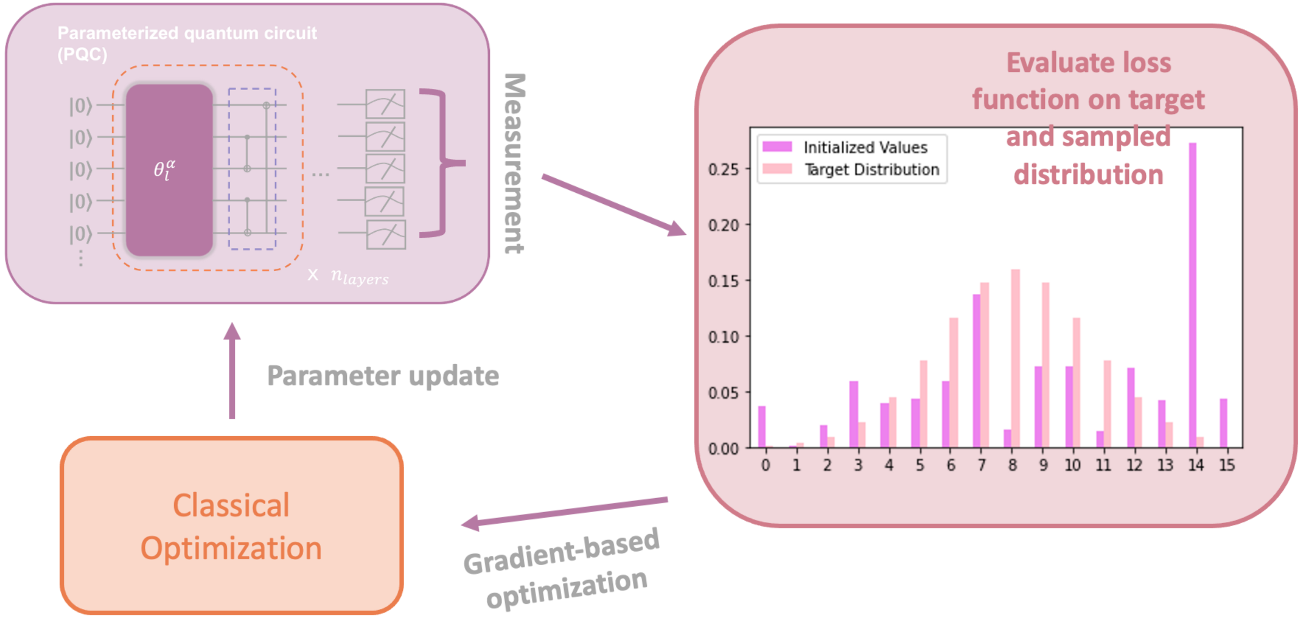




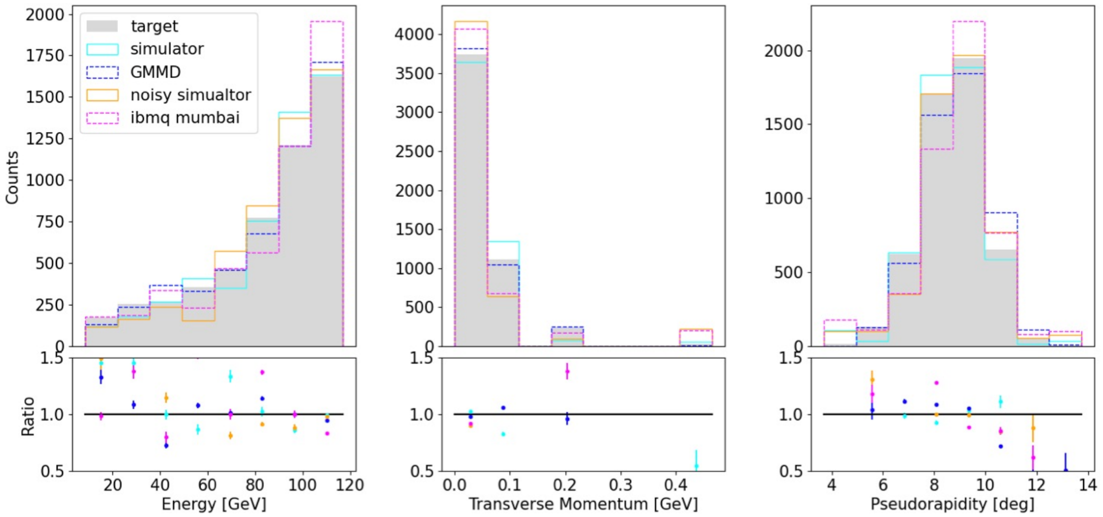
Delgado, A., Hamilton, K. E., "Unsupervised Quantum Circuit Learning in High-Energy Physics" arXiv:2203.03578

On Unsupervised Quantum Generative Models…

- ★ **Quantum generative models** are currently a candidate for quantum advantage in QML, with current performance comparable to classical methods… when trained on simulation.
- ★ There are sill a lot of open questions.
  - ★ Scalability?
  - ★ Ansatz choice
  - ★ Training
  - ★ Scalable error correction

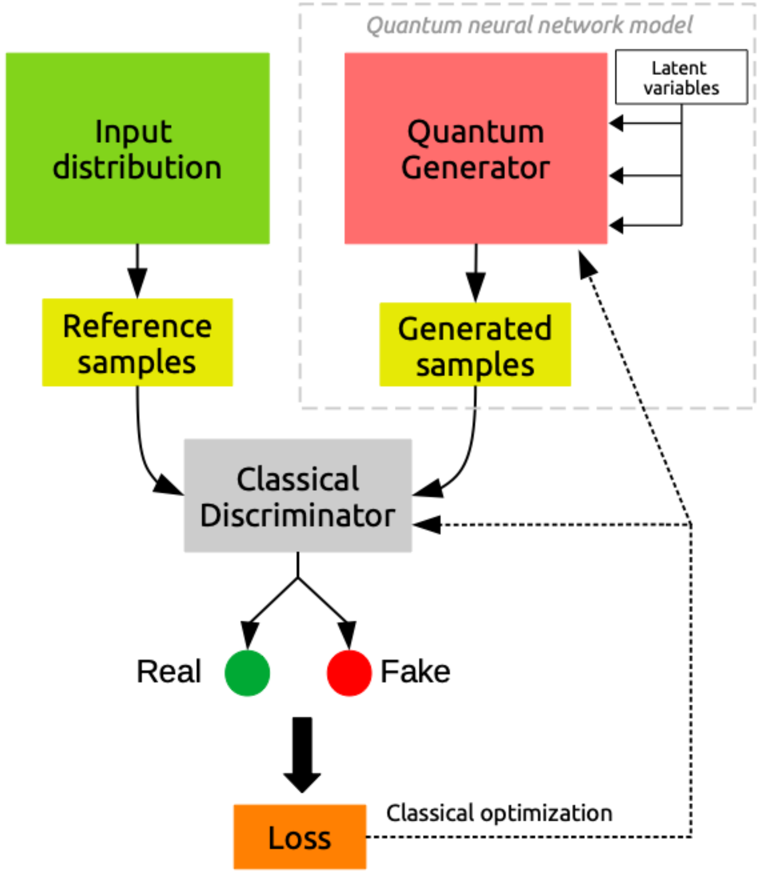


Also check out:

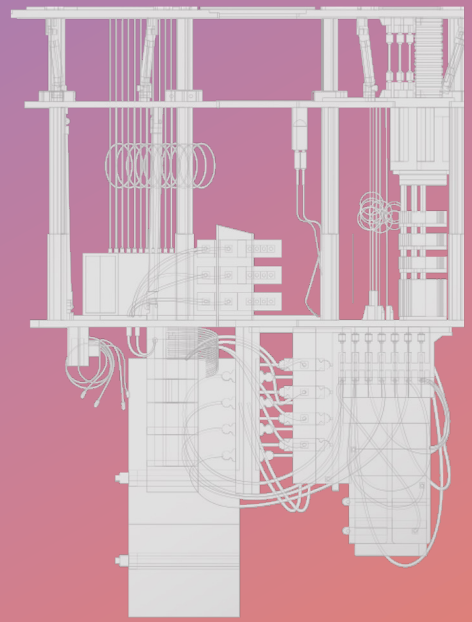


(a) Energy (b) Transverse momentum (c) pseudorapidity

”Conditional Born machine for Monte Carlo events generation”, Kiss, O., Grossi, M., Kajomovitz, E., Vallecorsa, S., arXiv:2205.07674



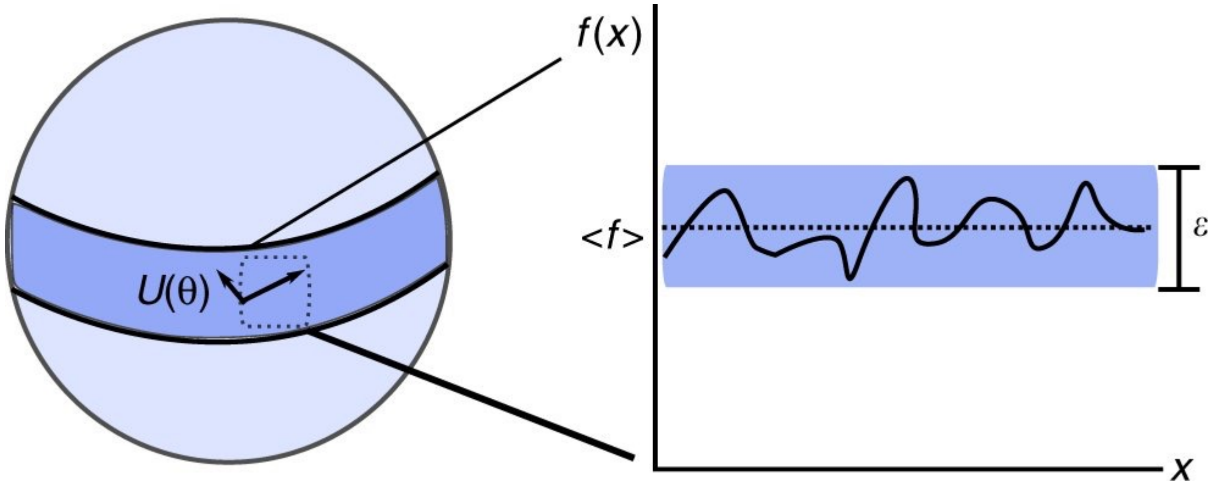
”Style-based quantum generative adversarial networks for Monte Carlo events”, Bravo-Prieto, C., Baglio, J., Ce, M., Francis, A., Grabowska, D., Carrazza, S., arXiv: 2110.06933

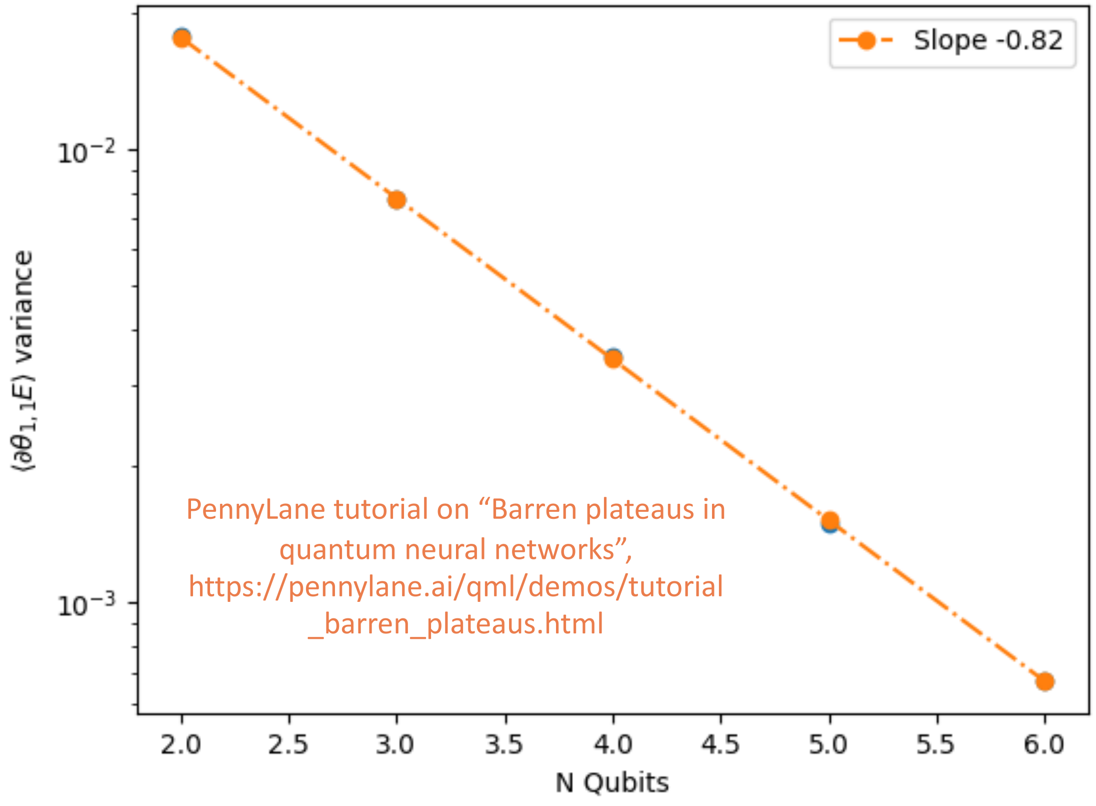


Barren Plateaus, and how to avoid them

(Overcoming) barren plateaus

- A **barren plateau (BP)** occurs in PQCs if the variational ansatz circuit is too random.
  - Prepares a random state with nearly maximal entropy.
- The wave function is essentially spread over the exponentially large Hilbert space.
  - Larger number of measurements to estimate observable.
  - Extrapolates to gradient calculation, to navigate the optimization landscape
- How to overcome them?
  - Several techniques have been explored





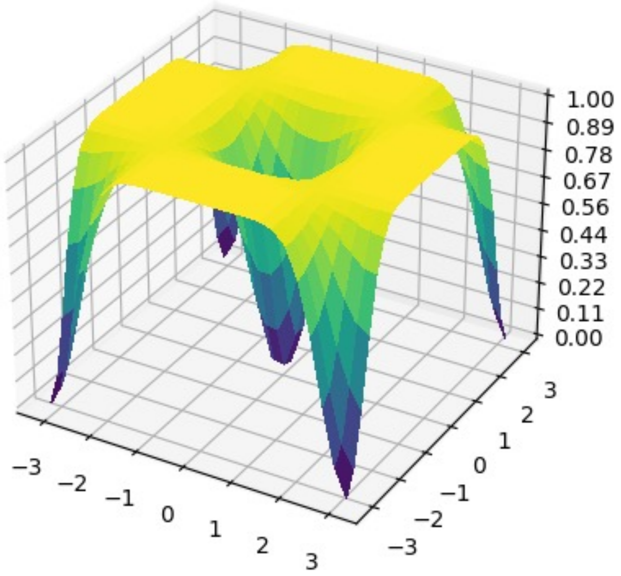
(Overcoming) barren plateaus

(1) Local vs Global Functions

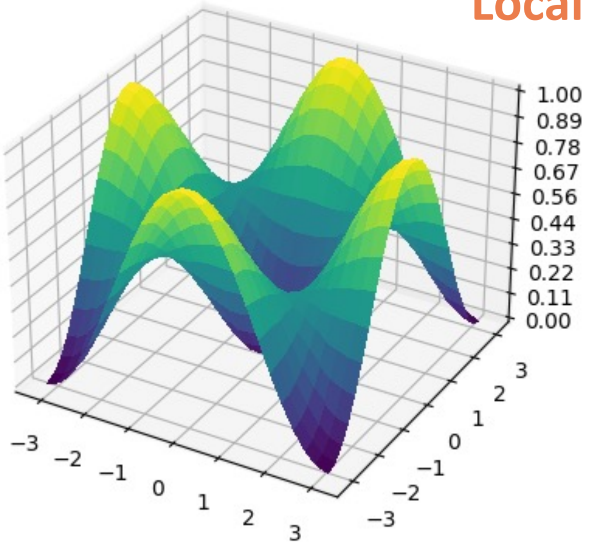
- ★ Most applications to date use global cost functions
  - ★ We just saw the effect of vanishing gradient with increasing number of qubits.
- ★ Evaluate cost function on shallower circuits.
  - ★ i.e., local cost function.

"Cost Function Dependent Barren Plateaus in Shallow Parameterized Quantum Circuits," M. Cerezo, A. Sone, T. Volkoff, L. Cincio, P. J. Coles, arXiv:2001.00550

Global cost function



Local cost function

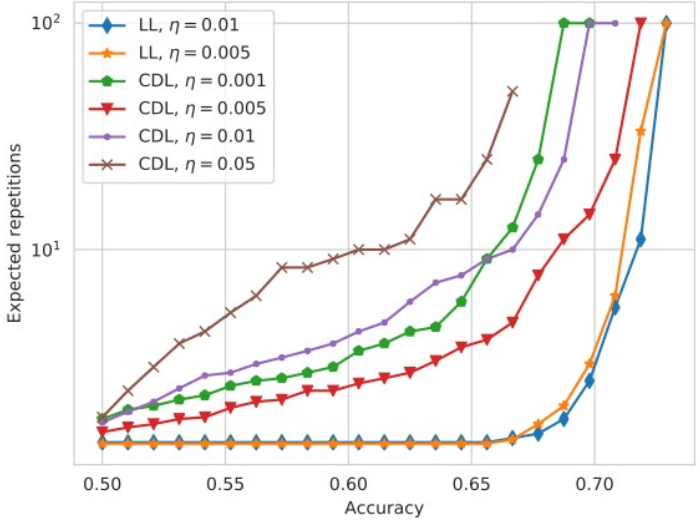


PennyLane tutorial on "Alleviating barren plateaus with local cost functions", https://pennylane.ai/qml/demos/tutorial_local_cost_functions.html

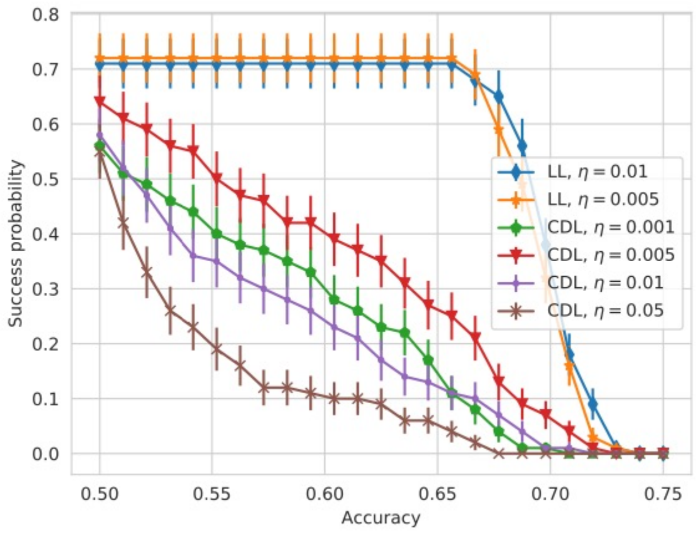
(Overcoming) barren plateaus

(2) Layerwise Learning (LL)

- ★ Method consists on incrementally growing the circuit during optimization,
  - ★ As opposed to training all layers.
- ★ Only subsets of parameters are updated in each training step.
- ★ Essentially restoring back to training shallow circuits.
- ★ LL is expected to:
  - ★ Decrease runtime,
  - ★ Increase probability of success on random restarts.



(a) expected number of repetitions



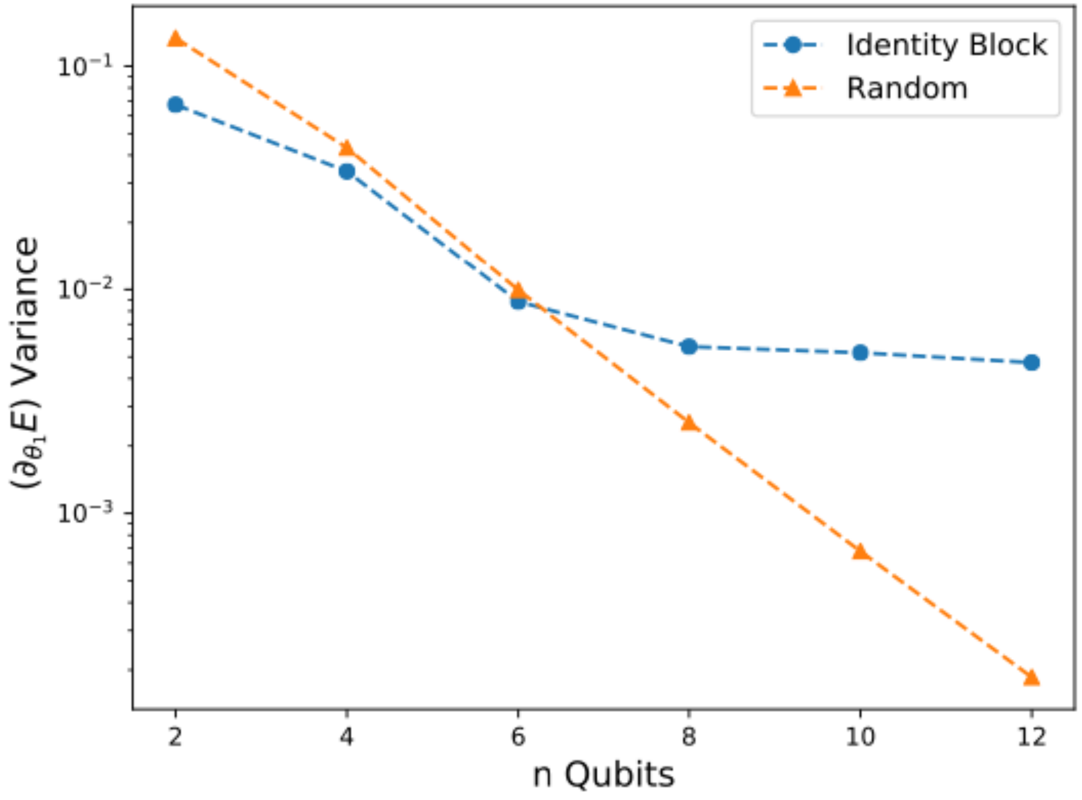
(b) success probability

"Layerwise learning for quantum neural networks," A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, M. Leib, arXiv:2006.14904

(Overcoming) barren plateaus

(3) Block Identity Initialization

- ★ Create the circuit in blocks such that the blocks satisfy: $U^{\dagger}U = 1$.
- ★ The resulting state is a product state (no entanglement) and this, no BP.
- ★ Too random -> Nearly maximal entanglement entropy.



"An initialization strategy for addressing barren plateaus in parameterized quantum circuits," E. Grant, L. Wossnig, M. Ostaszewski, M. Benedetti, arXiv:1903.05076

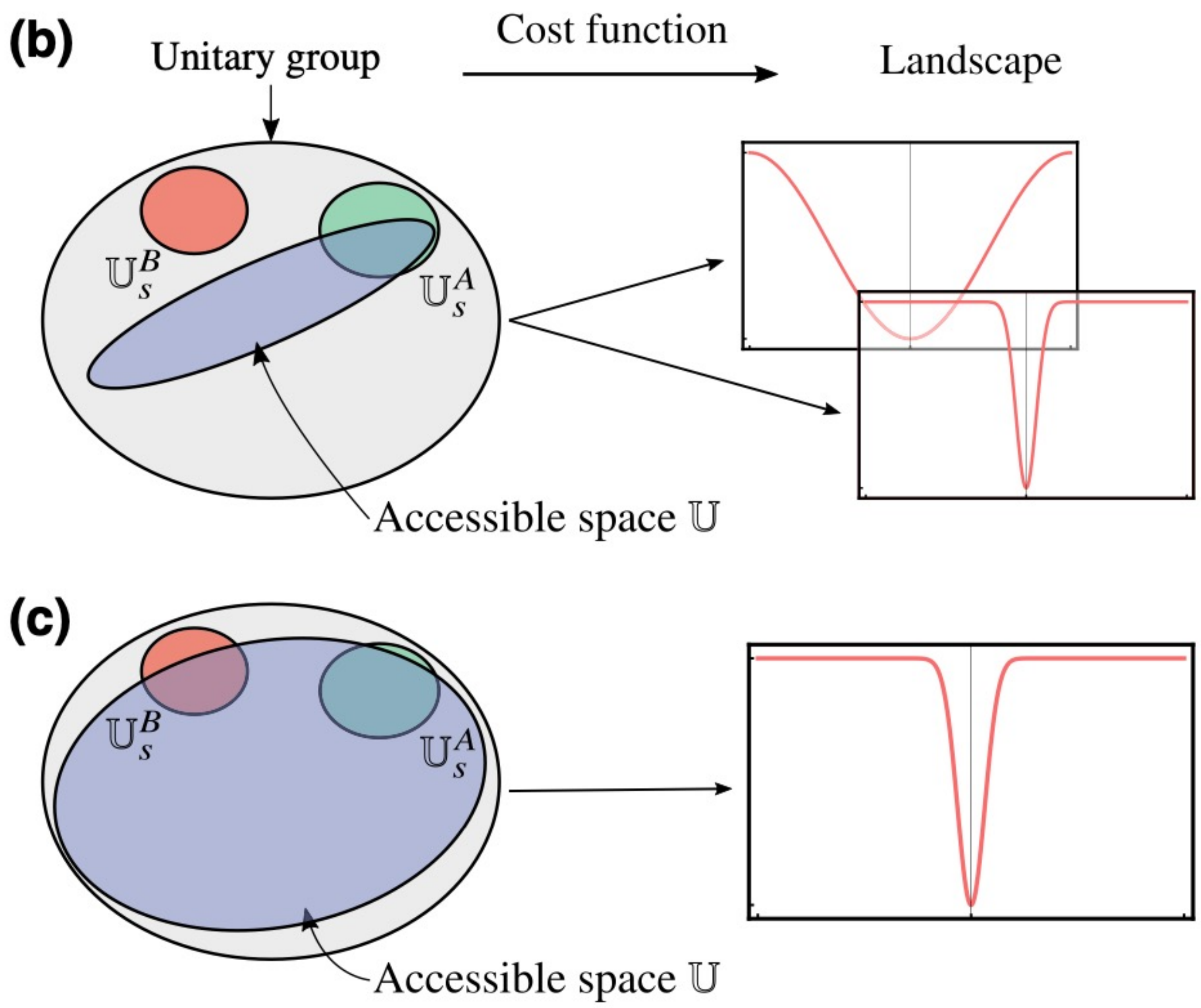
(Overcoming) barren plateaus

(4) Small qubit rotations

★ Studies showed that using small single qubit rotation angles can slow down the growth of entanglement.

★ By controlling the magnitude of the angles, the BP can be delayed to arbitrary circuit depth.

"Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus," Z. Holmes, K. Sharma, M. Cerezo, P. J. Coles, PRX Quantum 3, 010313 (2022)

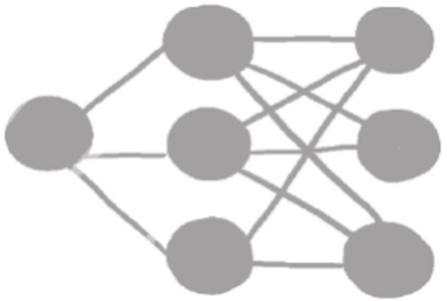


Is Quantum Advantage the Right Goal for QML?

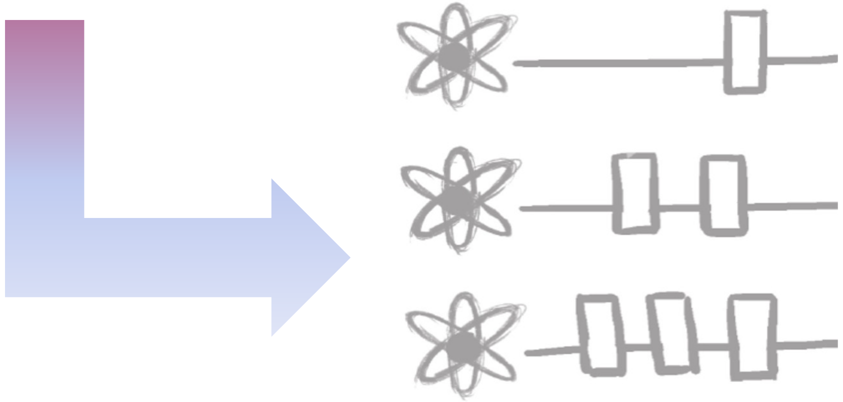
Based on the Perspective Manuscript by M. Schuld and N. Killoran, PRX Quantum 3, 030101 (2022)

★ ML is a **hard problem**!
- ❑ There is no rigorous basis for generalization.
- ❑ NNs are sequences of linear and non-linear transformations, making them unwieldy for mathematical modeling.

★ Once we add **"quantumness"** to the mix
- ❑ We only have minimal access to empirical results from *"just running the algorithm"*.
- ❑ We cannot say much about the behavior that quantum models will have at a scale beyond *what can be "simulated"*.



★ What architecture is best suited for a problem?
★ What affects trainability?
★ Model expressibility
★ Generalization power?

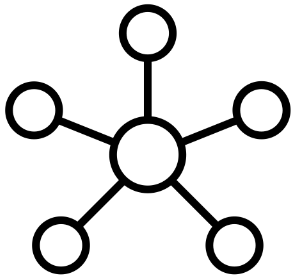


"The question on whether quantum computers can really play a role in identifying practical ML application is still wide open, and It is unlikely to be decided by theoretical proofs or small-scale experiments"

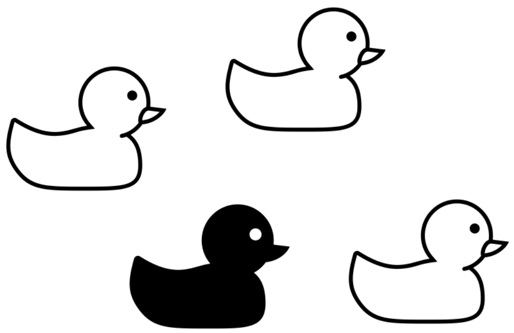
But also… what about?

Quantum Machine Learning on Quantum Data?

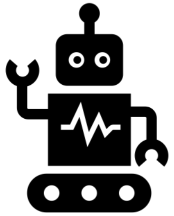
Ensemble learning methods for network of quantum sensors?



Anomaly detection?

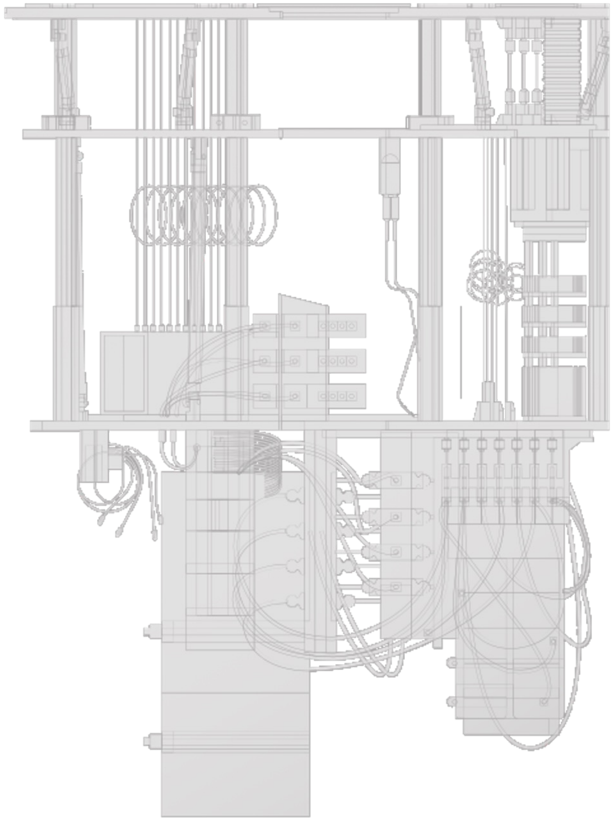


System Control



Summary

- ★ **Machine Learning algorithms based on parameterized quantum circuits are a prime candidate for near-term applications on noisy quantum computers. But…**
  - ★ We still don’t understand how these QML models compare, both mutually and to classical ML models.



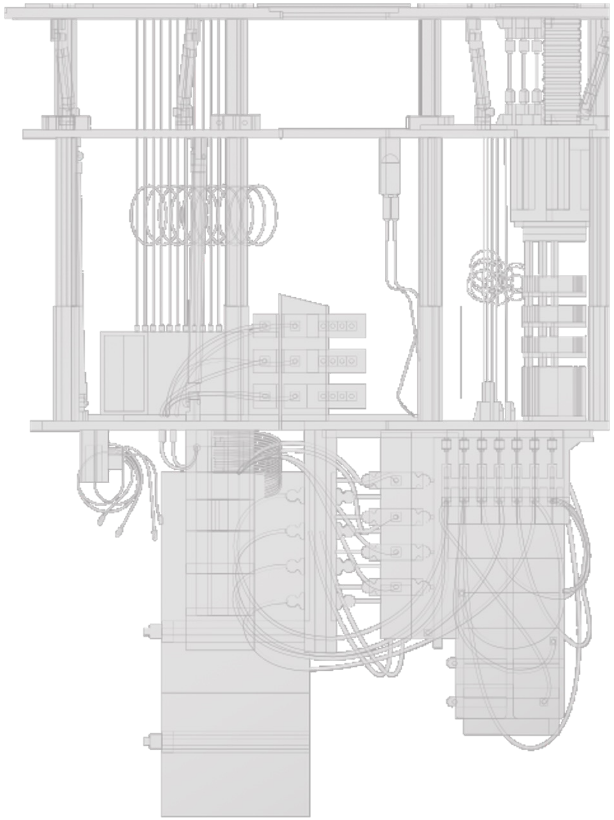
Summary

★ Machine Learning algorithms based on parameterized quantum circuits are a prime candidate for near-term applications on noisy quantum computers. But…

- ★ We still don't understand how these QML models compare, both mutually and to classical ML models.

★ There are several things I didn't cover today, but I encourage you to read about them:

- ★ Continuous variable quantum machine learning
- ★ Tensor networks as ML models.



Summary

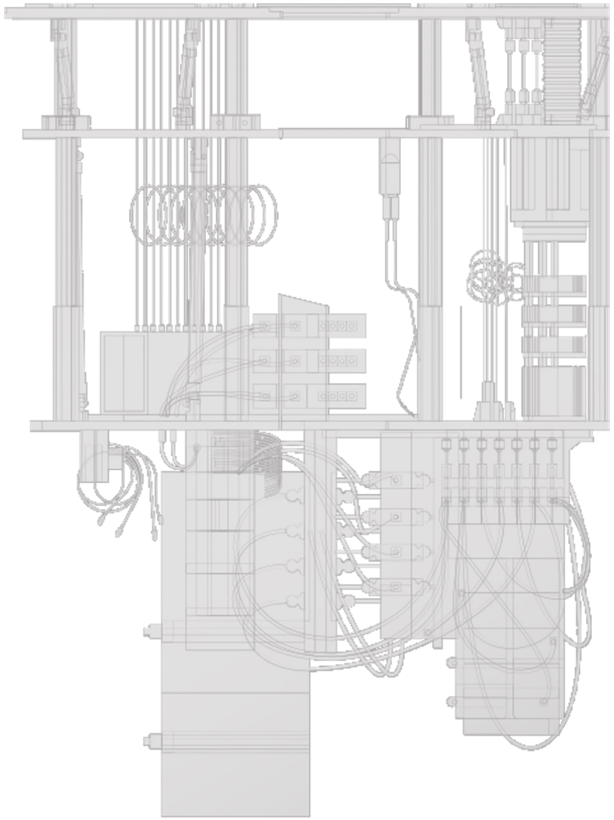
★ Machine Learning algorithms based on parameterized quantum circuits are a prime candidate for near-term applications on noisy quantum computers. But…

- ★ We still don't understand how these QML models compare, both mutually and to classical ML models.

★ There are several things I didn't cover today, but I encourage you to read about them:

- ★ Continuous variable quantum machine learning
- ★ Tensor networks as ML models.

★ Its an exciting time for QML!



Thank you!

delgadoa@ornl.gov