Quantum Machine Learning in High Energy Physics

Examples from CERN

QUANTUM TECHNOLOGY INITIATIVE

Sofia.Vallecorsa@cern.ch

Quantum Technologies, Workshop INFN CSN4&5, June 7th-8th, 2023

QML: Quantum computing to “improve” ML

- Speed-up and complexity
- Sample efficiency
- Representational power
- Energy efficiency???

- Evaluate performance on realistic use cases

- QPU as accelerators within classical infrastructure?

Outline

- Introduction: the CERN Quantum Technology Initiative
- Quantum Machine Learning and Applications at CERN
- Anomaly detection
- Beam optimisation in linear accelerators
- Improving robustness
- Summary

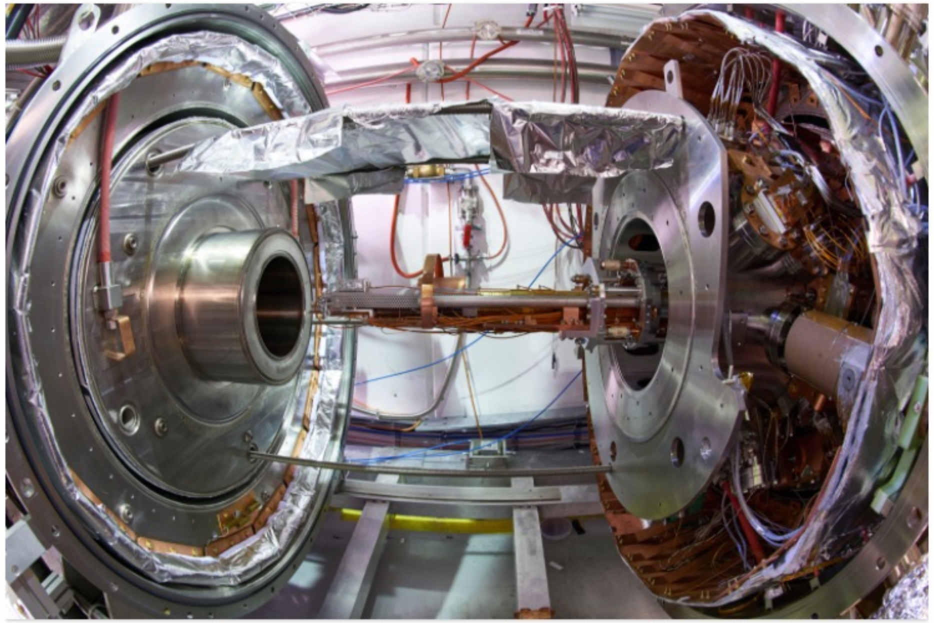
The CERN Quantum Technology Initiative

Voir en français

CERN meets quantum technology

The CERN Quantum Technology Initiative will explore the potential of devices harnessing perplexing quantum phenomena such as entanglement to enrich and expand its challenging research programme

30 SEPTEMBER, 2020 | By Matthew Chalmers



The AEgIS 1T antimatter trap stack. CERN's AEgIS experiment is able to explore the multi-particle entangled nature of photons from positronium annihilation, and is one of several examples of existing CERN research with relevance to quantum technologies. (Image: CERN)

Quantum simulation and HEP theory applications

Quantum Computing

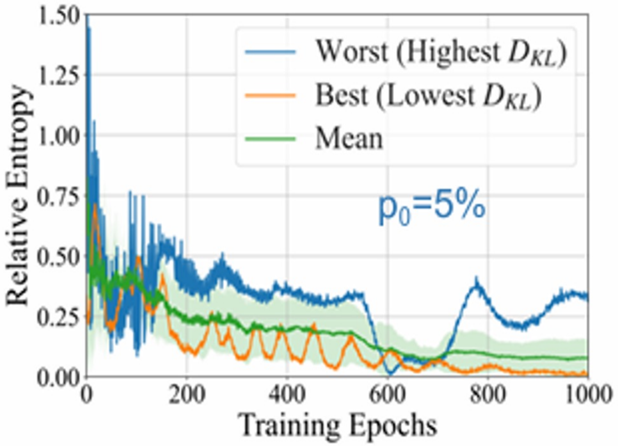
Quantum Sensing

Quantum Communication

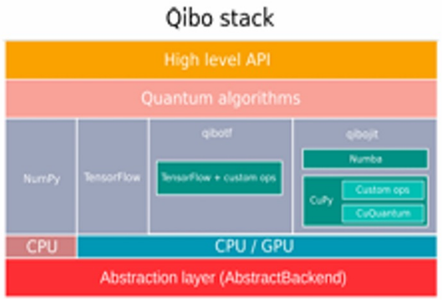
QTI Roadmap: https://doi.org/10.5281/zenodo.5553774

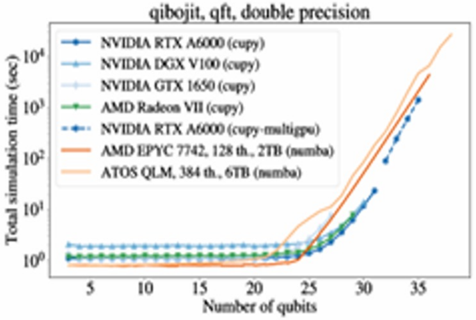
QC @ CERN

Borras, Kerstin, et al. "**Impact of quantum noise on the training of quantum Generative Adversarial Networks**." *arXiv preprint arXiv:2203.01007* (2022).

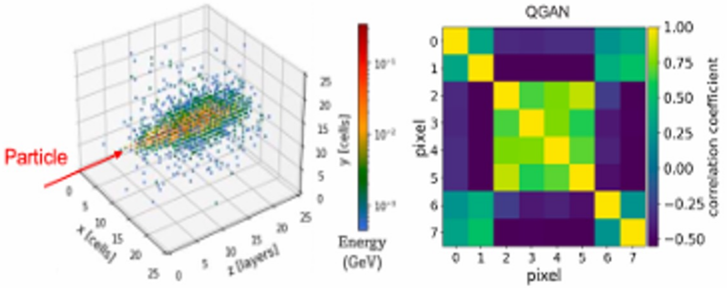


E.Stavros et all., **Quantum simulation with just-in-time compilation**, Quantum 2022

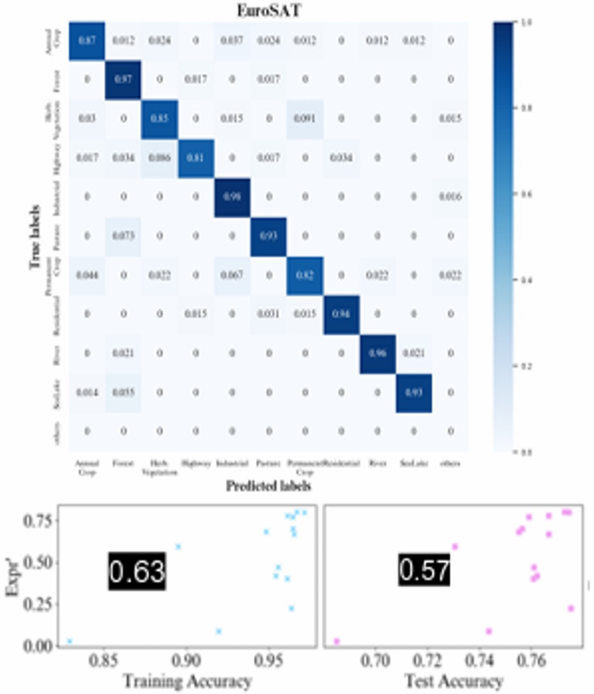




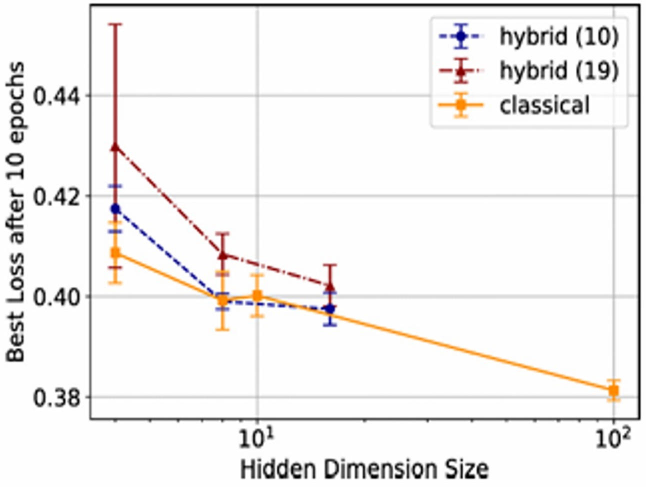
F.Rehm, **Full Quantum GAN Model for HEP Detector Simulations**, ACAT22



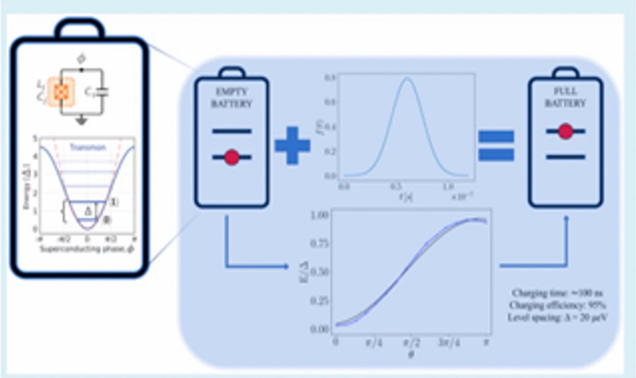
S.Chang, et all, **Hybrid Quantum-Classical Networks for Reconstruction and Classification of Earth Observation Images**, ACAT22



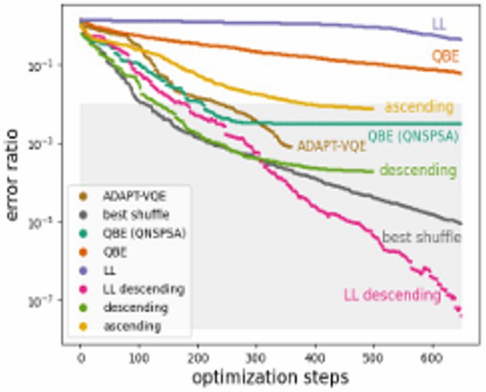
Tüysüz, Cenk, et al. "**Hybrid quantum classical graph neural networks for particle track reconstruction.**" *Quantum Machine Intelligence* 3.2 (2021): 1-20.



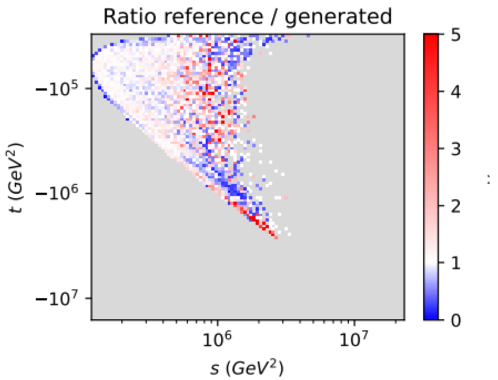
G. Gemme, M. Grossi et al, **IBM Quantum Platforms: A Quantum Battery Perspective**, *Batteries 8, 43 (2022)*



O. Kiss, **Quantum computing of the 6Li nucleus via ordered unitary coupled cluster**, 10.1103/PhysRevC.106.034325



Bravo-Prieto, Carlos, et al. "**Style-based quantum generative adversarial networks for Monte Carlo events.**" *Quantum 2022*

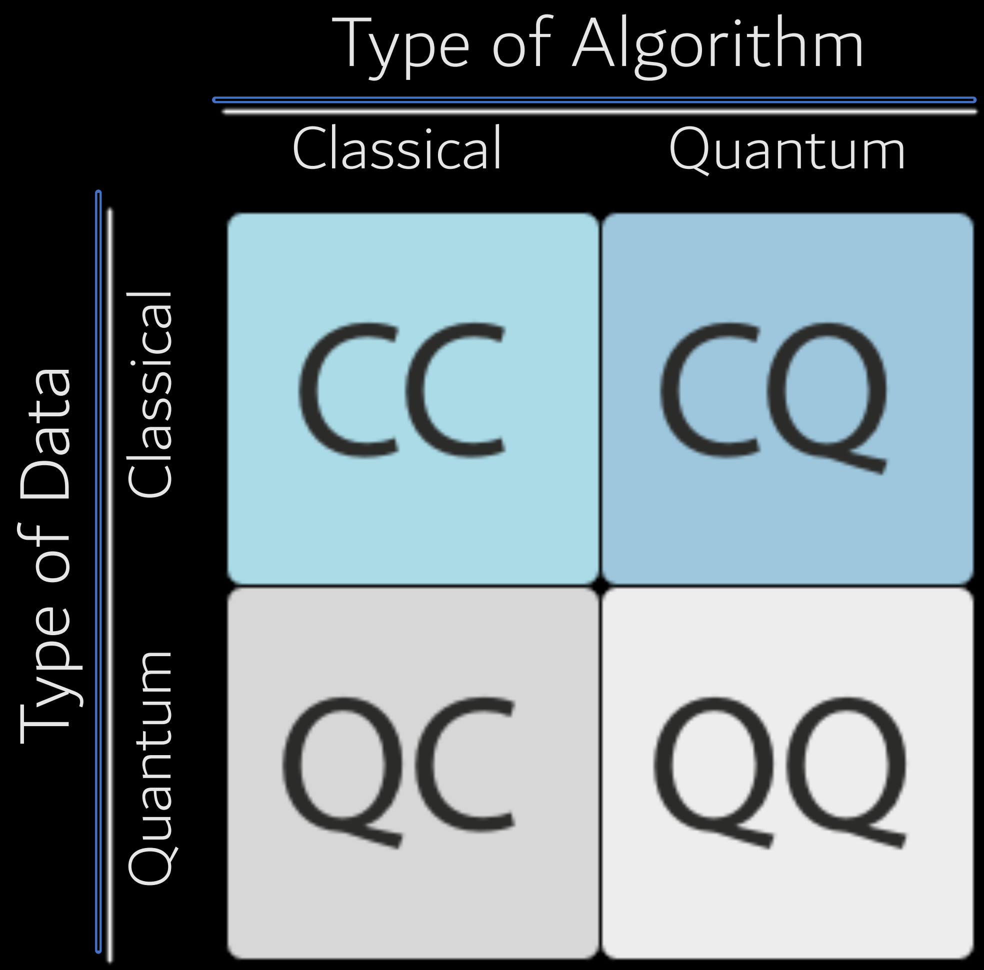


Quantum Machine Learning :

Introduction

QML in HEP

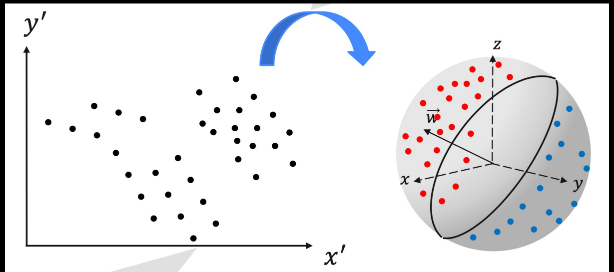
- Does it make sense to use QML in HEP?
- How do we understand when it is *useful* ?
- Which are the QML models we can leverage?



Classical Intractability:

- No established recipe for classical data
- Compromise between algorithm **expressivity** vs **trainability** and **generalization**

(Quantum) ML Lifecycle



Data Embedding

Data Preparation

Model Definition

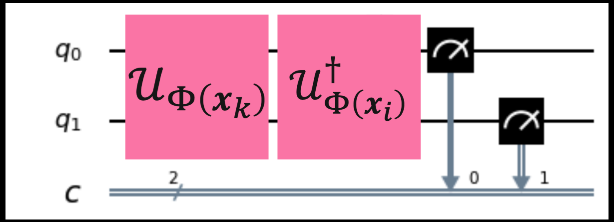
Model Training

Model Testing

Model Interpretation

Adapt classical learning models to quantum space

Readout and measurement "shots"



The advantage of many known QML algorithms is impeded today by I/O bottleneck

Models

Variational algorithms (ex. QNN)

Gradient-free or **gradient-based** optimization

Data Embedding can be **learned**

Ansatz design can leverage data symmetries[1]

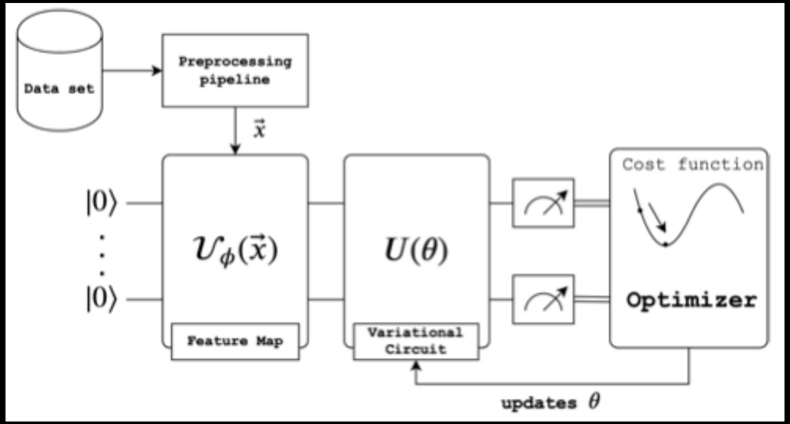


Image credit SwissQuantumHub

Representer theorem:

implicit models achieve **better accuracy**[3]

Explicit models exhibit **better generalization** performance

Kernel methods (ex. QSVM)

Feature maps as **quantum kernels**

Classical **kernel-based training** (**convex** losses)

Identify classes of kernels that relate to specific data **structures**[2]

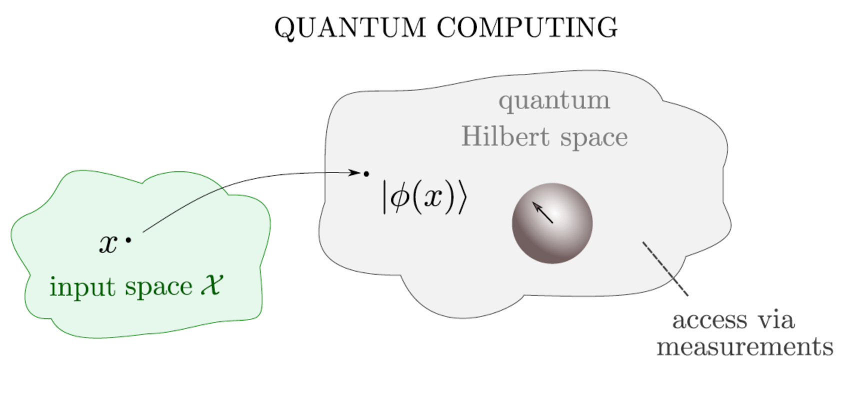


Image credit M. Schuld

Energy-based ML (ex. QBM)

Build network of stochastic binary units and optimise their energy.

QBM has quadratic energy function that follows the Boltzman distribution (Ising Hamiltonian)

1 Bogatskiy, Alexander, et al. "**Lorentz group equivariant neural network for particle physics**." PMLR, 2020.

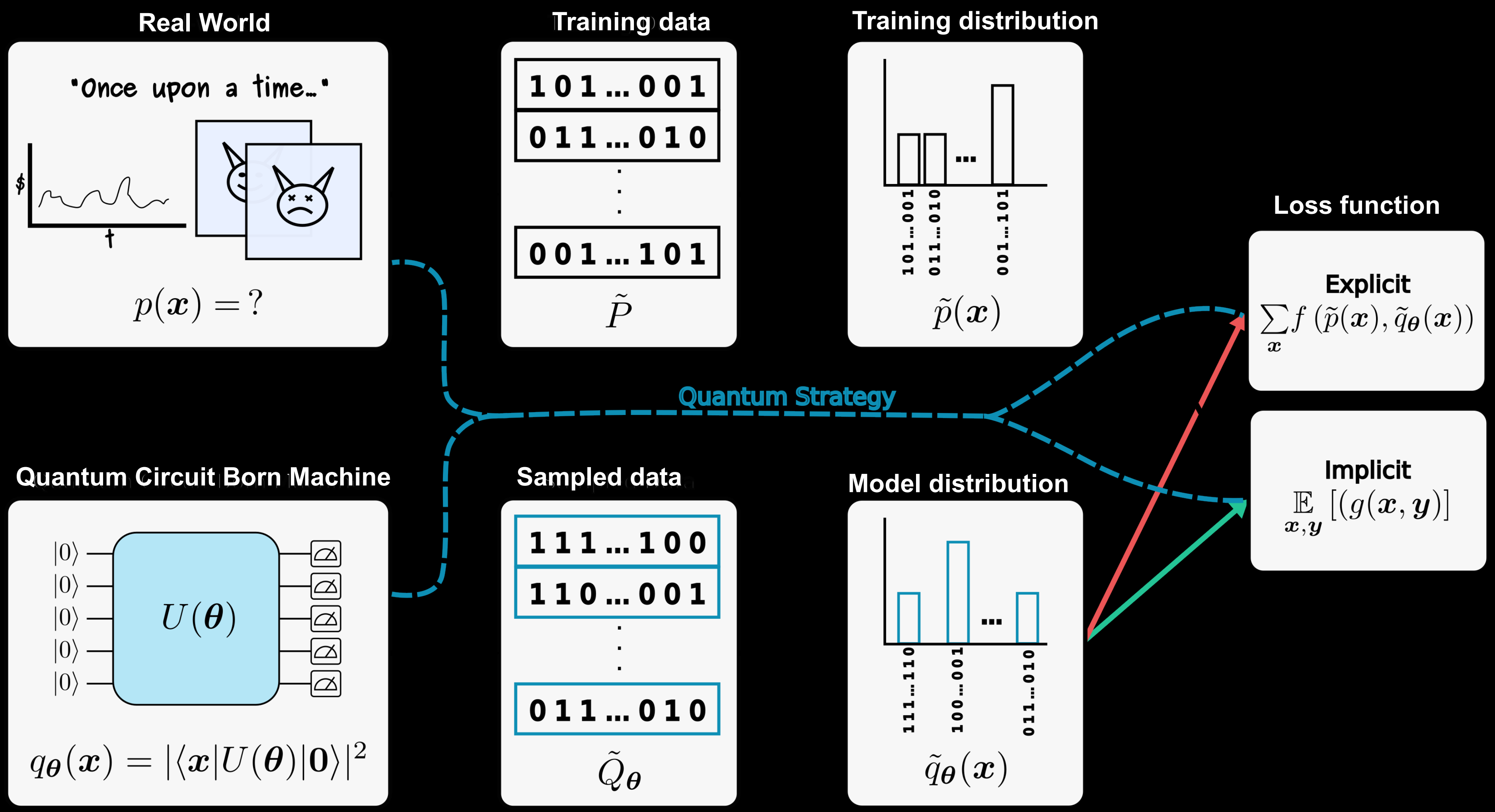
2 Glick, Jennifer R., et al. "**Covariant quantum kernels for data with group structure**." *arXiv:2105.03406* (2021).

[3]Jerbi, Sofiene, et al. "**Quantum machine learning beyond kernel methods**." *arXiv preprint arXiv:2110.13162* (2021).

Rudolph, M. S., Lerch, S., Thanasilp, S., Kiss, O., Vallecorsa, S., Grossi, M., & Holmes, Z. (2023). **Trainability barriers and opportunities in quantum generative modeling.** *arXiv:2305.02881*.

Generative QML and trainability barriers

Representation learning: encoding probability distributions

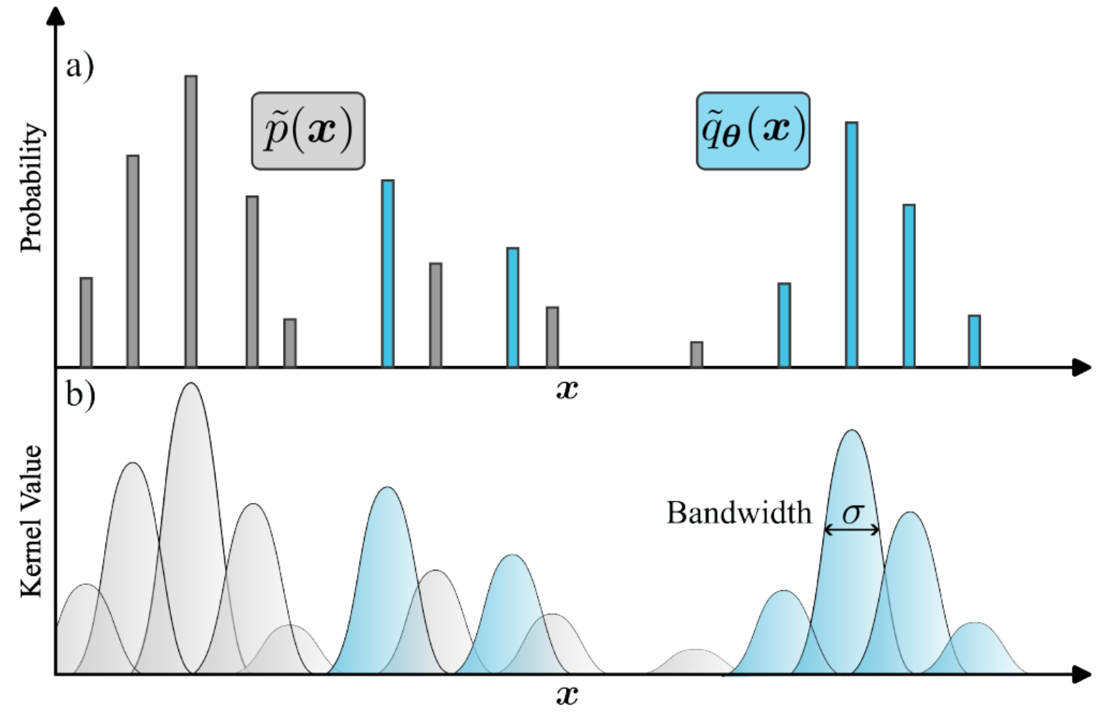


Rudolph, M. S., Lerch, S., Thanasilp, S., Kiss, O., Vallecorsa, S., Grossi, M., & Holmes, Z. (2023). **Trainability barriers and opportunities in quantum generative modeling.** *arXiv:2305.02881*.

Explicit and Implicit Losses

The use of explicit losses hinders trainability of implicit generative models as the system size increases since it requires an exponentially larger number of shots.

Need implicit losses!



Explicit

$$\mathcal{L}^{\mathrm{KLD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} p(\boldsymbol{x}) \log\left(\frac{p(\boldsymbol{x})}{q_{\boldsymbol{\theta}}(\boldsymbol{x})}\right)$$

Implicit

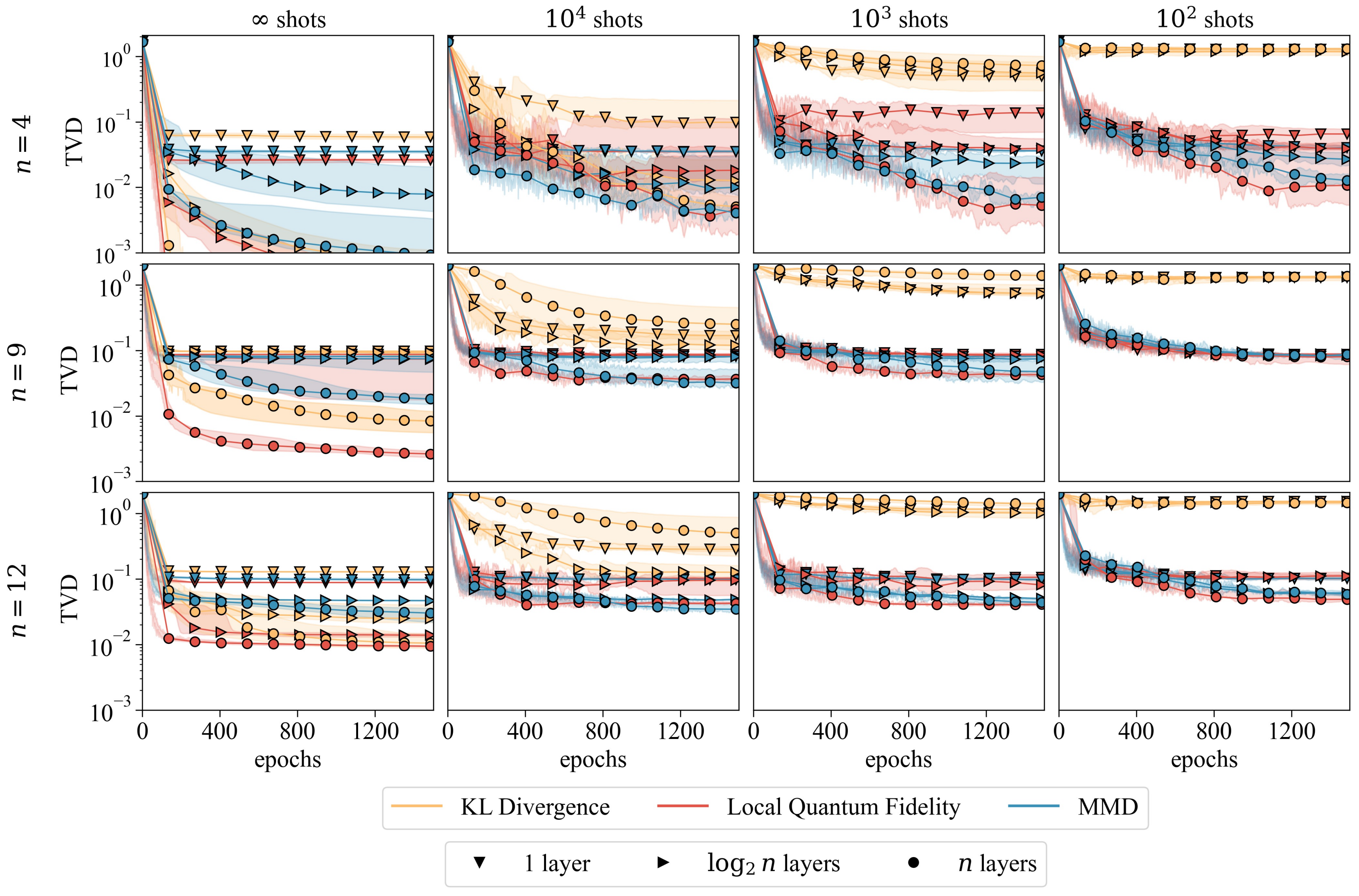
$$\mathcal{L}_{\mathrm{MMD}}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{x},\boldsymbol{y}\sim q_{\boldsymbol{\theta}}}[K(\boldsymbol{x},\boldsymbol{y})] - 2\mathbb{E}_{\boldsymbol{x}\sim q_{\boldsymbol{\theta}},\boldsymbol{y}\sim p}[K(\boldsymbol{x},\boldsymbol{y})] + \mathbb{E}_{\boldsymbol{x},\boldsymbol{y}\sim p}[K(\boldsymbol{x},\boldsymbol{y})]\,,$$

Rudolph, M. S., Lerch, S., Thanasilp, S., Kiss, O., Vallecorsa, S., Grossi, M., & Holmes, Z. (2023). **Trainability barriers and opportunities in quantum generative modeling.** *arXiv:2305.02881*.

Quantum Circuit Born Machine for HEP

QCBM

Sample variational pure state $|\psi(\theta)\rangle$ by projective measurement through **Born rule:** $\mathbf{p}_{\boldsymbol{\theta}}(\mathbf{x}) = |\langle \mathbf{x}|\boldsymbol{\psi}(\boldsymbol{\theta})\rangle|^2$.



Quantum embedding for classical data

Compromise between **exponential compression and circuit depth**

Ex: **Amplitude Encoding**

$$|\phi(x)\rangle = \frac{1}{\|x\|}\sum_{i=0}^{N} x_i|i\rangle$$

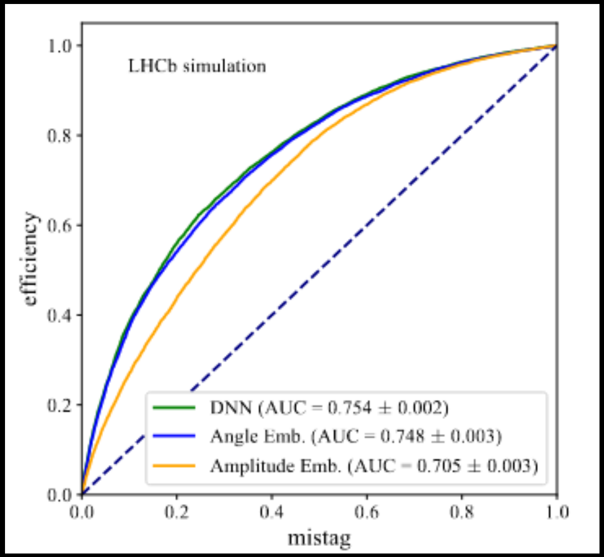
Exponential compression

$\mathbf{n_{qubit}} \propto \mathbf{O(\log(N))}$

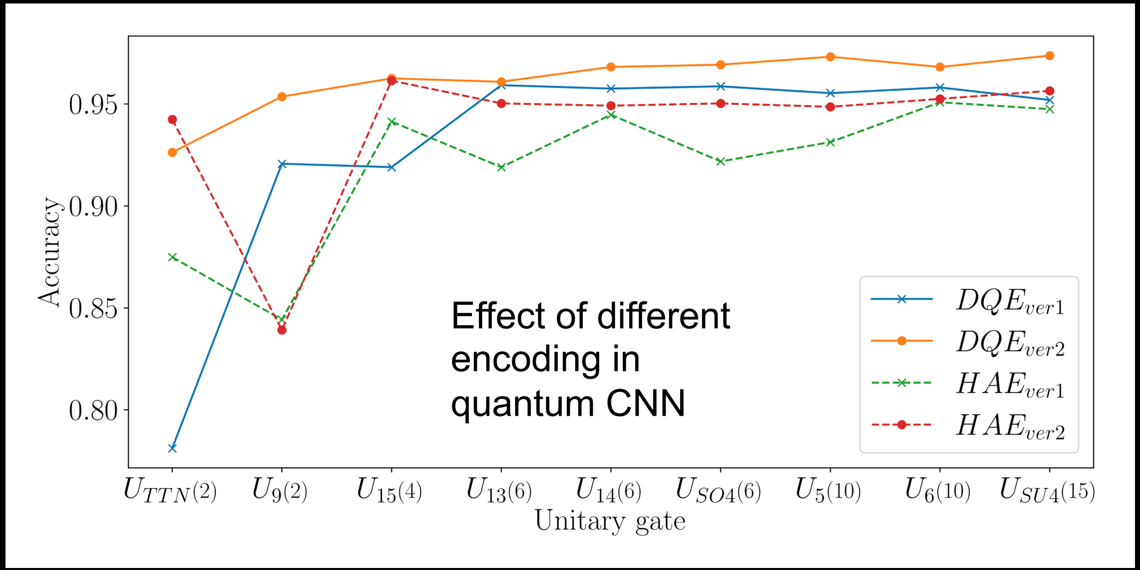
Polynomial number of gates

$\mathbf{n_{gate}} \propto \mathbf{O(poly(N))}$

Gianelle, A., Koppenburg, P., Lucchesi, D. *et al.* **Quantum Machine Learning for *b*-jet charge identification**. *J. High Energ. Phys.* **2022,** 14 (2022). https://doi.org/10.1007/JHEP08(2022)014

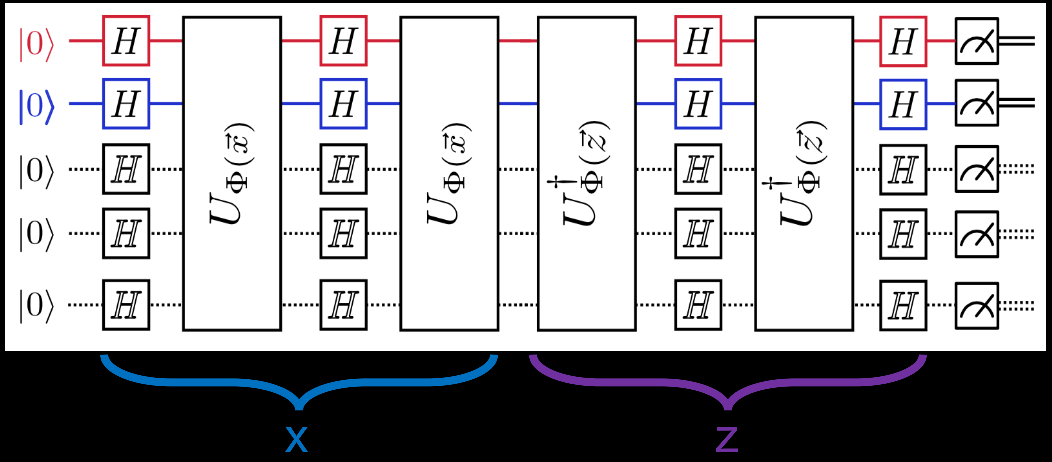


S.Y. Chang, poster at "**Quantum Tensor Network in Machine Learning**, NeurIPS 2021



Quantum embedding and kernel methods

- Create **classically intractable features in the Hilbert space**
- Estimate **Fidelity kernel**
- Use classical training (convex losses)



$$\hat{y} = \text{label}(z) = \text{sign}\left(\sum \alpha_i y_i K(x_i, z) + b\right)$$

$$|\langle \Phi(\vec{x}) | \Phi(\vec{z}) \rangle|^2 = |\langle 0^m | U^\dagger_{\Phi(\vec{x})} U_{\Phi(\vec{z})} | 0^m \rangle|^2$$

Hilbert space is exponentially larger

↓

Sparser data

↓

Loss of predictive power

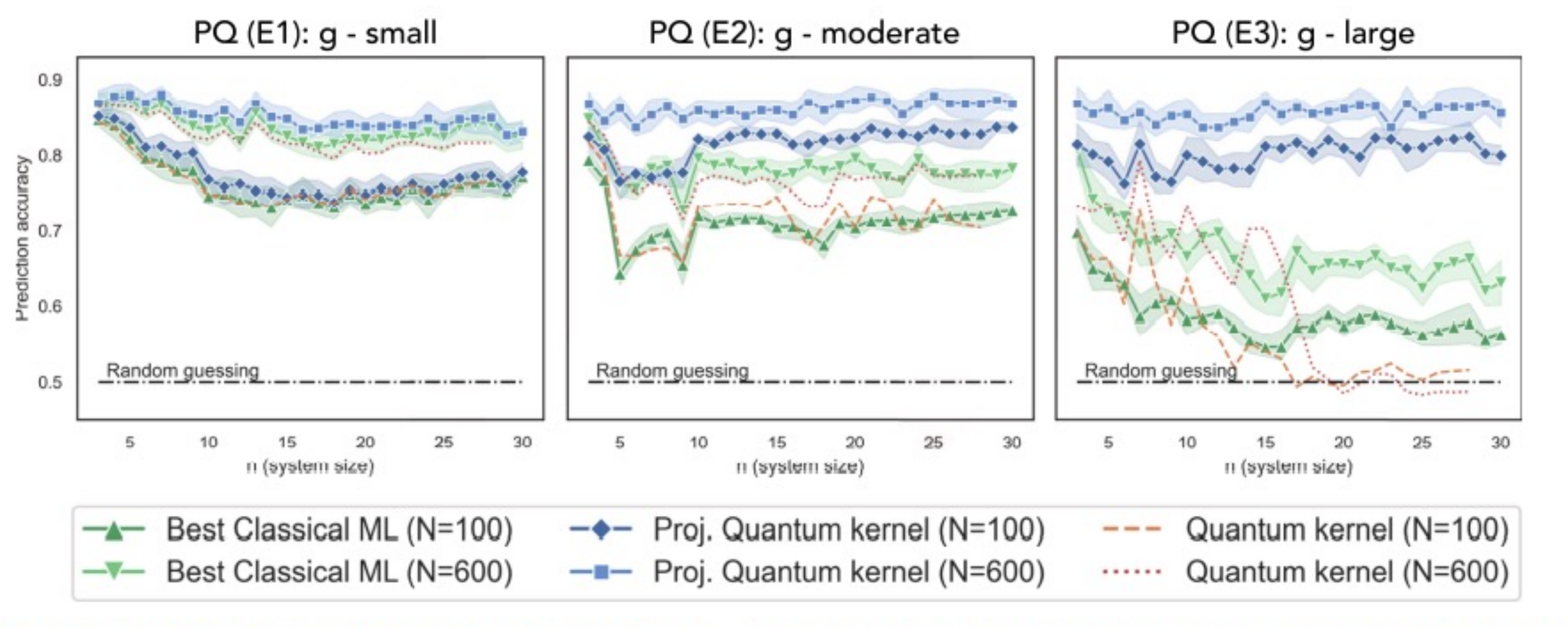
Projected Quantum Kernel

$$k^{lp}(x_i, x_j) = \sum_{k=1}^{m} \frac{Tr[\rho_k(x_i)\rho_k(x_j)]}{m}$$

Project quantum kernels to lower dimensionality (i.e. local density matrix):

- Improved generalizion while keeping features into states classically hard

- g_{CQ}: geometric difference between classical and quantum embeddings



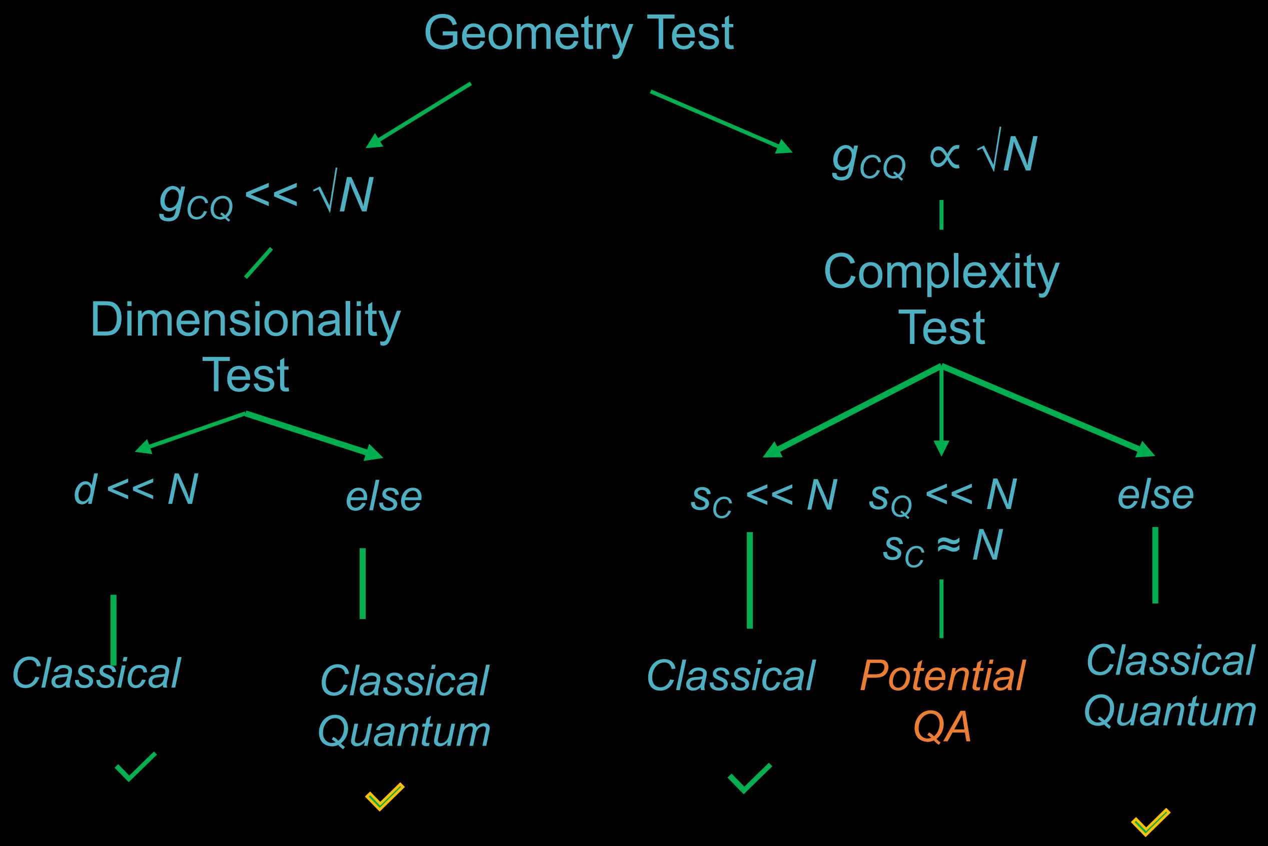
Huang, Hsin-Yuan, et al. "Power of data in quantum machine learning." *Nature communications* 12.1 (2021): 2631.

Quantum Advantage

Define an upper bound on classical and quantum kernels prediction error

$$\mathbb{E}_{\mathbf{x}}|h(\mathbf{x}) - y(\mathbf{x})| \leq \mathcal{O}\left(\sqrt{\frac{s_{K,\lambda}(N)}{N}}\right)$$

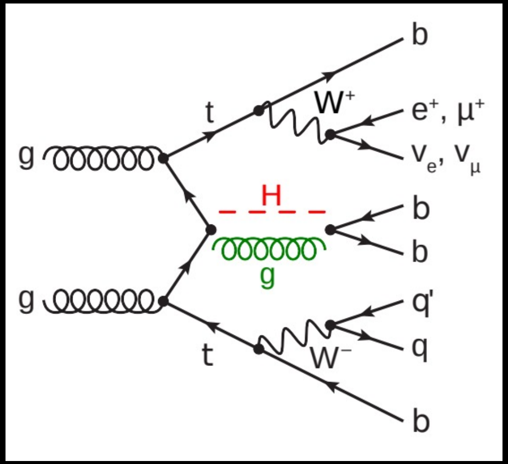
- N training events
- ***g_{CQ}:*** geometric difference between classical and quantum embeddings
- ***S* (N)**: model complexity
- **d** *:* feature space dimension

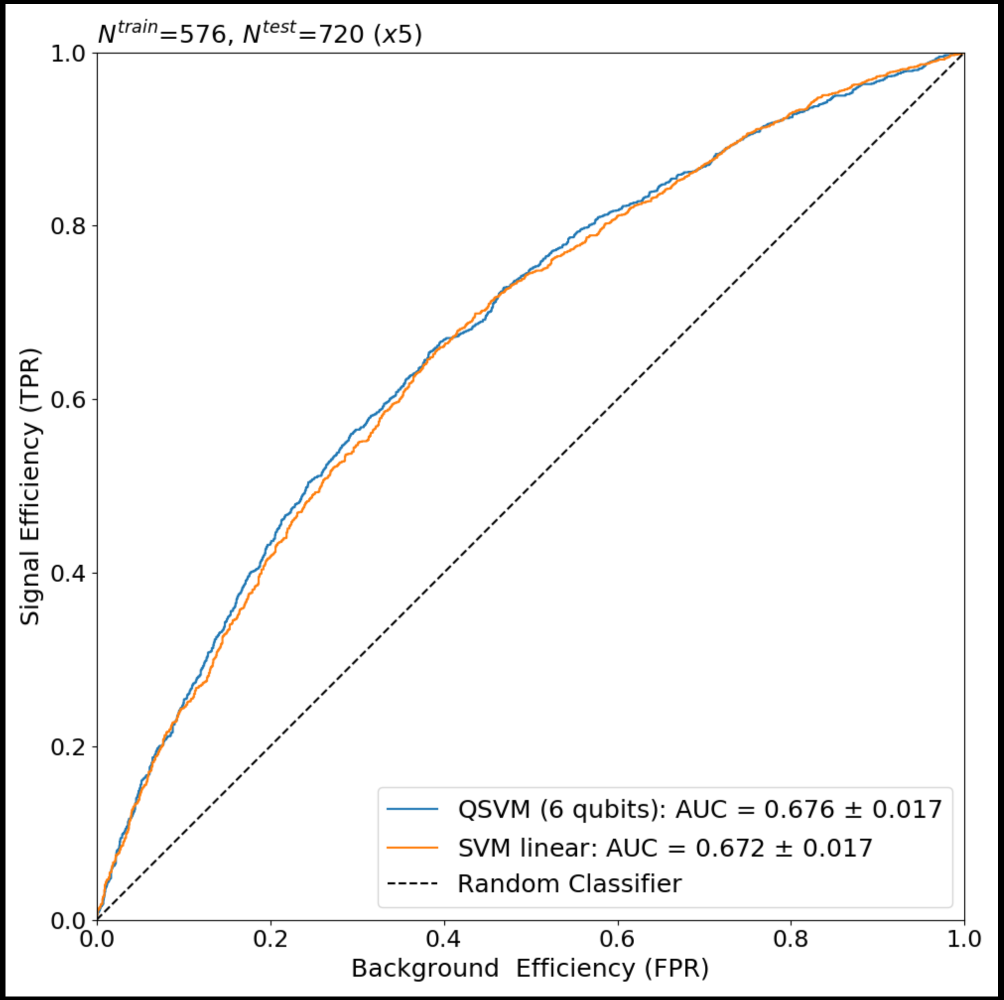


Huang, Hsin-Yuan, et al. "Power of data in quantum machine learning." *Nature communications* 12.1 (2021): 2631.

Higgs classification

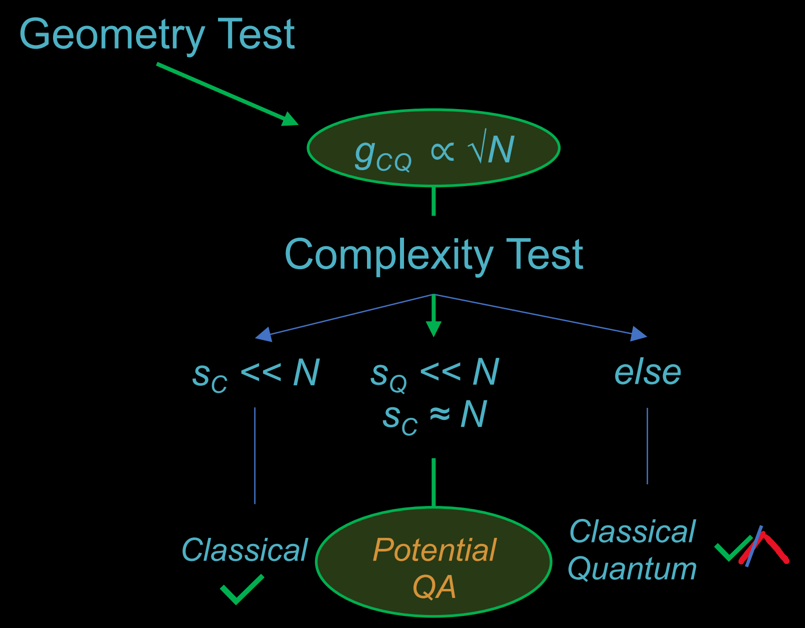
Quantum Support Vector Machine for the *ttH(bb)* event classification[5]





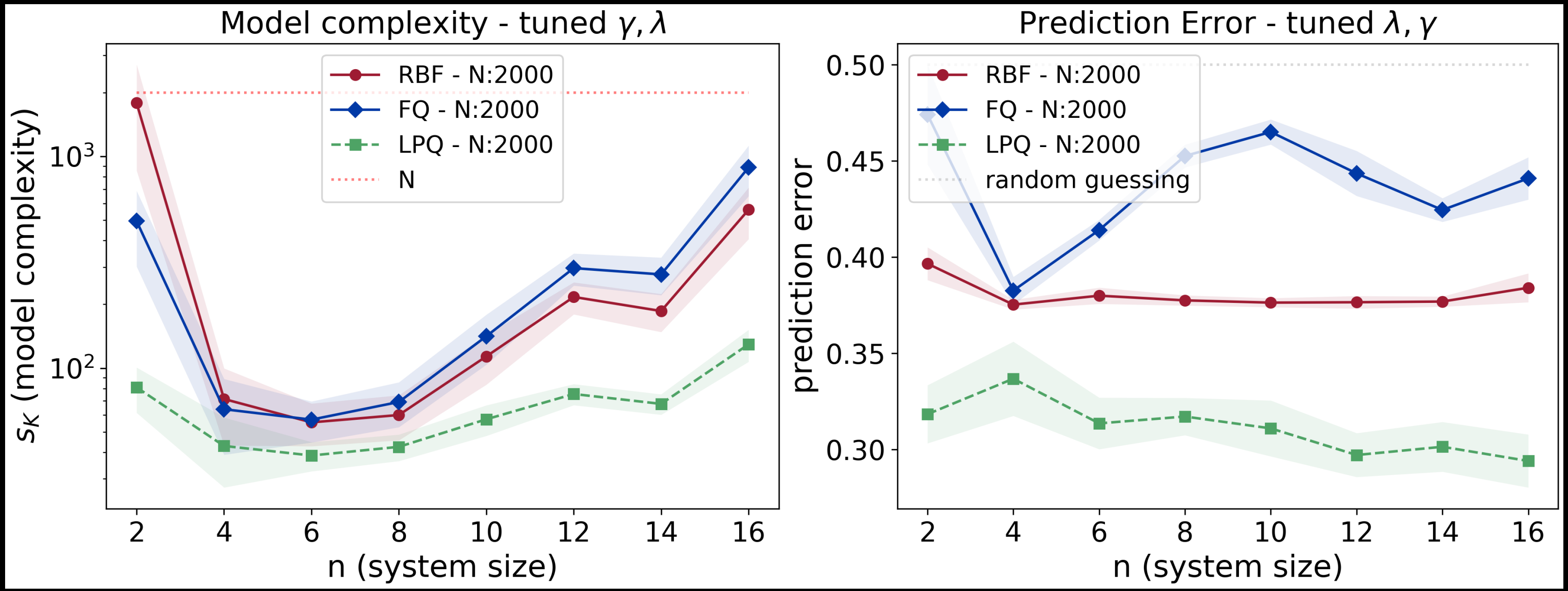
[1] V Belis et al, (2021), *Higgs Analysis with Quantum Classifiers,* EPJ Web Conf

Projected kernels work best



Optimized quantum and classical kernels

- g_{CQ} moderate to √N
- s_C and s_Q moderate/comparable to N

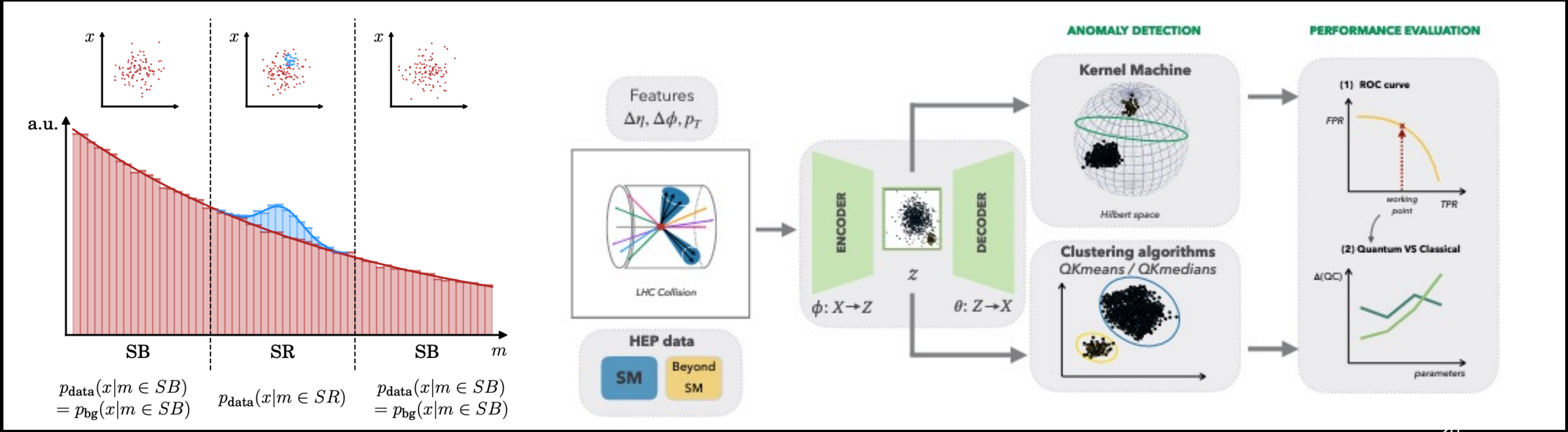


Quantum Machine Learning examples:

Analysis and Anomaly Detection

Unsupervised learning for Anomaly Detection

Anomaly detection can point to new physics at the LHC



Standard Model jets

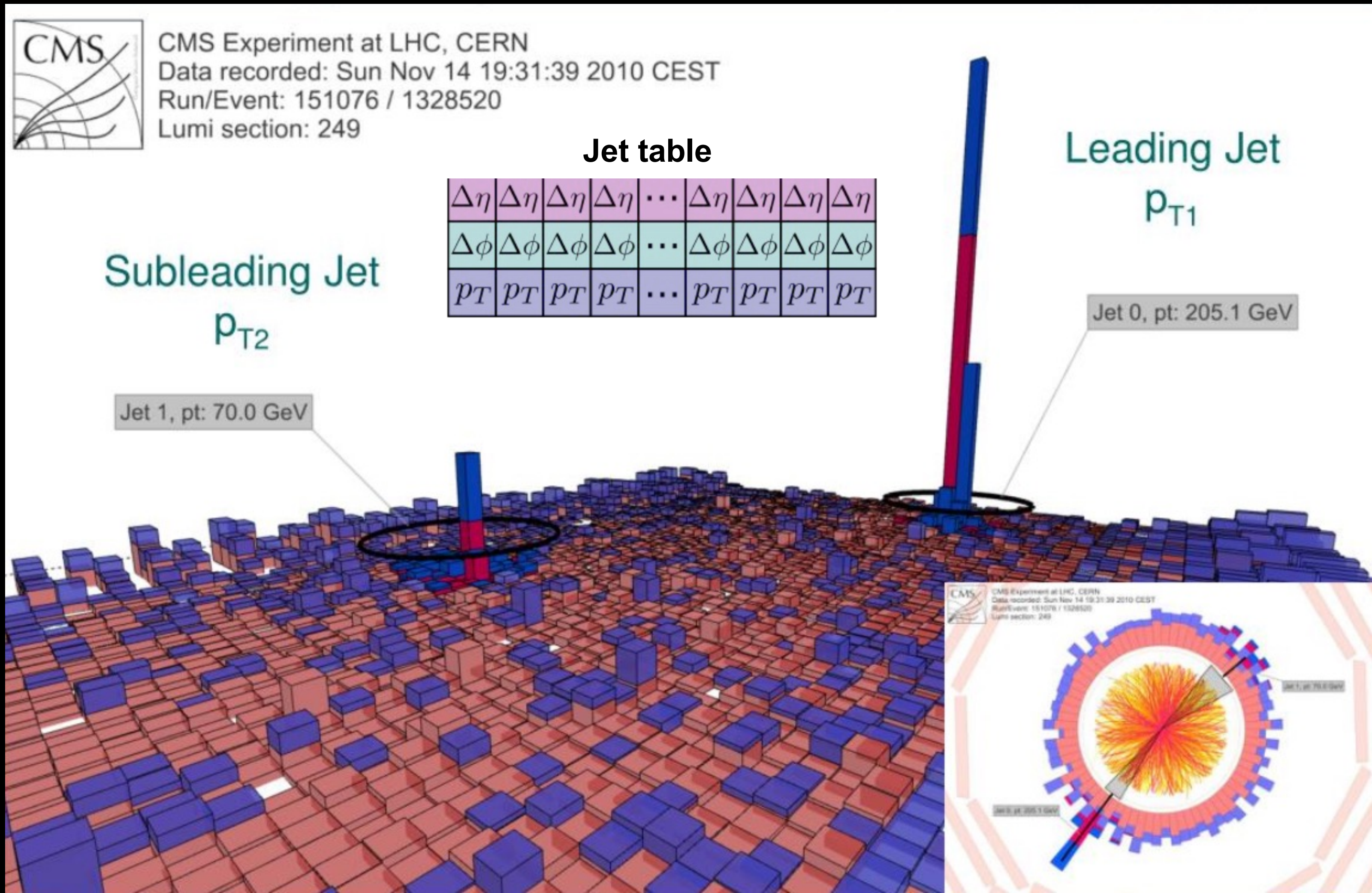
Simulate QCD multi-jets at the LHC

Build jet from 100 highest pt particles

Apply realistic event selection

Convolutional AutoEncoder
learns the jet **internal structure**

$$\mathbb{R}^{300} \to \mathbb{R}^{\ell}\,,\ \ell = 4, 8, 16$$



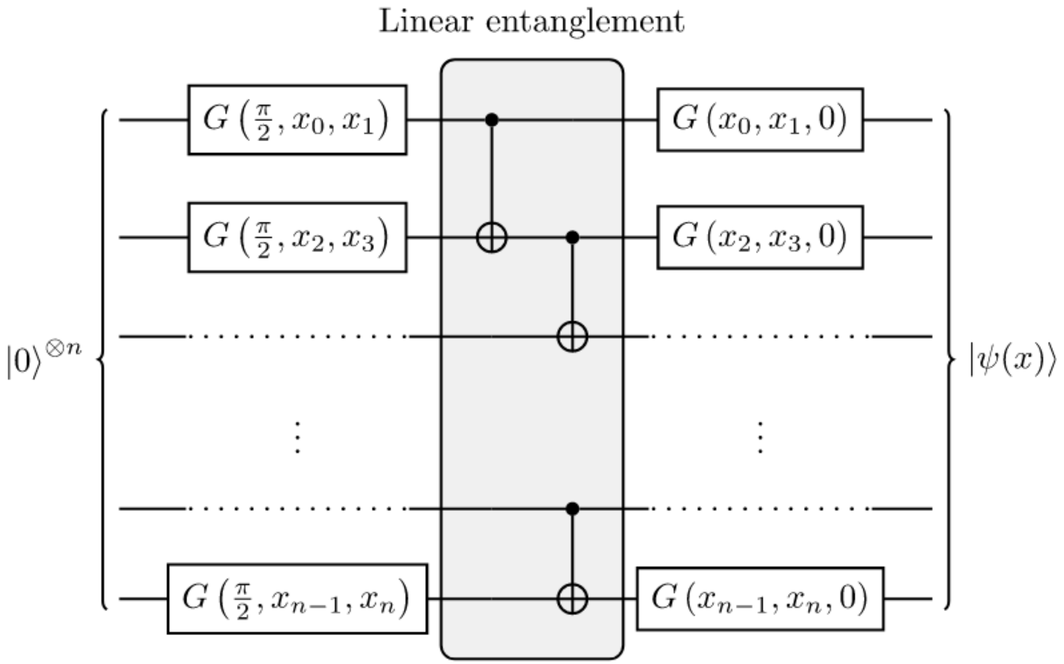
Unsupervised kernel machine

Find the hyperplane that maximizes the distance of the data from the origin of the feature vector space

Upper bound on fraction of anomalies in training data at 0.01 (at most 1% QCD training data are falsely flagged)

$$k(x_i, x_j) := \mathrm{tr}[\rho(x_i)\rho(x_j)] = \left|\langle 0|U^\dagger(x_i)U(x_j)|0\rangle\right|^2$$

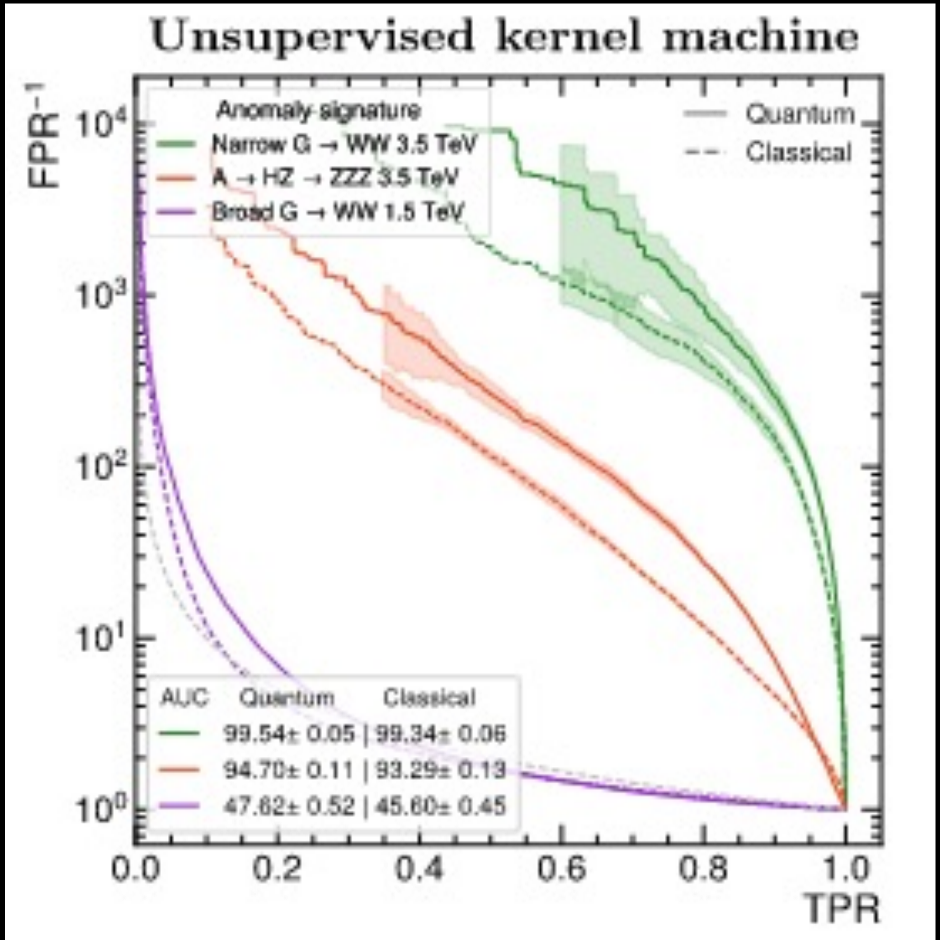
$$\rho(x_i) := U(x_i)\,|0\rangle\,\langle 0|\,U^\dagger(x_i)$$



$$\min_{w\in\mathcal{F},\, \xi\in\mathbb{R}^\ell,\, \rho\in\mathbb{R}} \quad \frac{1}{2}||w||^2 + \frac{1}{\nu\ell}\sum_i \xi_i - \rho$$

$$\text{subject to} \quad w\cdot\Phi(x_i) \geq \rho - \xi_i,\ \xi_i \geq 0,\ \forall i \qquad \nu \in (0,1)$$

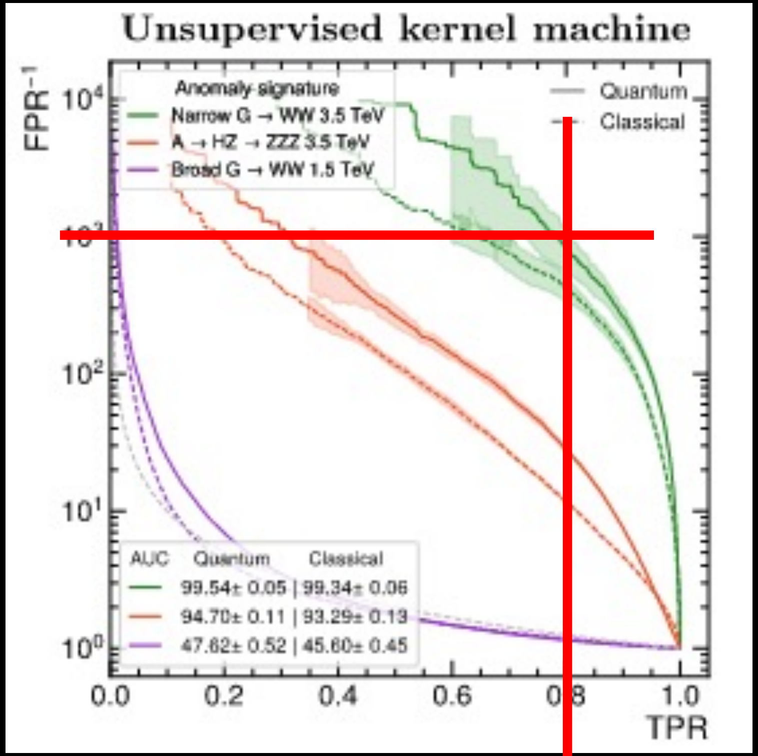
Results



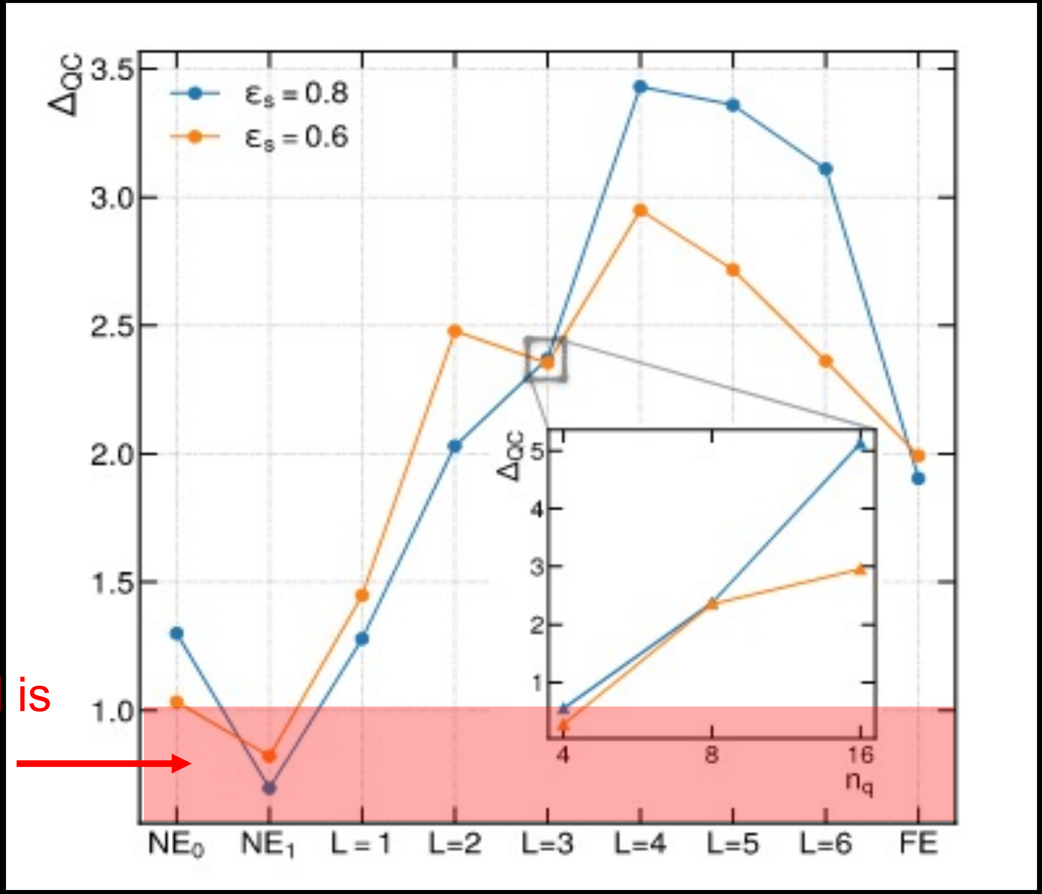
Is this an «advantage» we can use?

Quantum anomaly detection in the latent space of proton collision events at the LHC
Vasileios Belis *et al.*, *arXiv:2301.10780*.

In reality....



Classical is better



Higher is better

Increasing entanglement & expressivity

Quantum anomaly detection in the latent space of proton collision events at the LHC
Vasileios Belis *et al.*, *arXiv:2301.10780*.

More AD results...

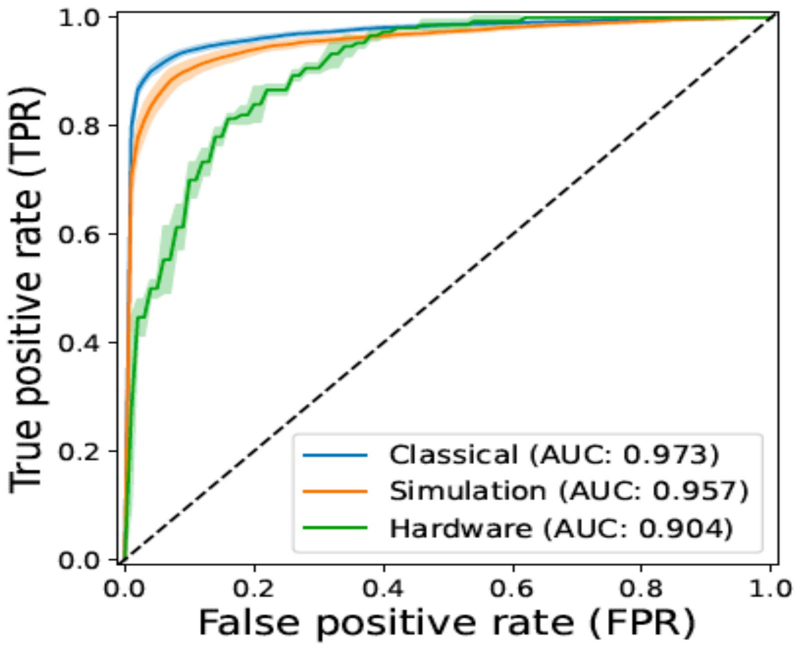


FIG. 4. ROC-AUC curve of the classification between SM events and artificial anomalies. The kernel matrices for the classification were provided by a classical kernel function (blue), a simulated quantum kernel (orange), and a quantum kernel estimated using the quantum device *ibm_cairo* (green).

Unravelling physics beyond the Standard Model with unbiased classical and quantum anomaly detection
Julian Schumacher *et al.*, *arXiv:2301.10787*.

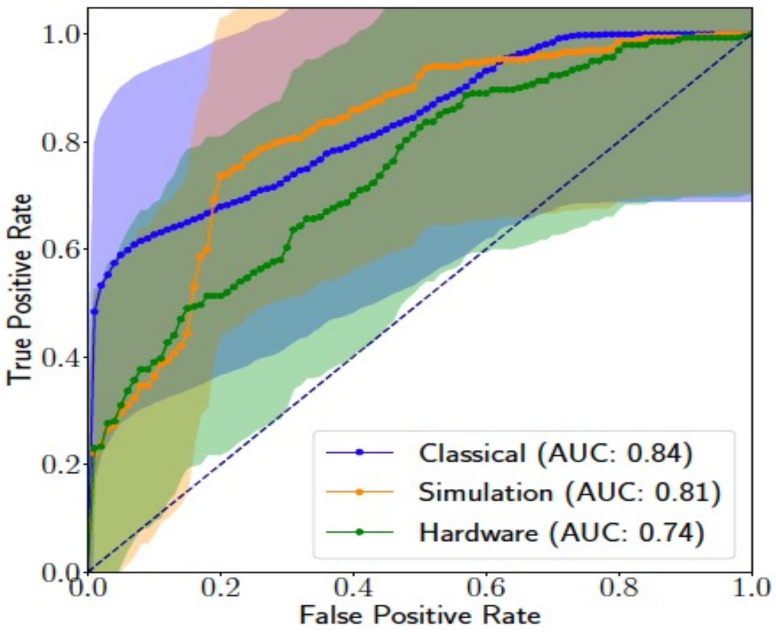


Fig. 6: ROC-AUC curve of the classification between SM and Graviton events for the classical GAN (blue), the noiseless simulation of a qGAN (orange), and the qGAN executed on the IBM Quantum processor *ibmq_belem*. All models are trained on 3 features of the SM data set, and evaluated on SM and Graviton events.

Quantum Generative Adversarial Networks For Anomaly Detection In High Energy Physics
Elie Bermot *et al.*, *arXiv:2304.14439*.

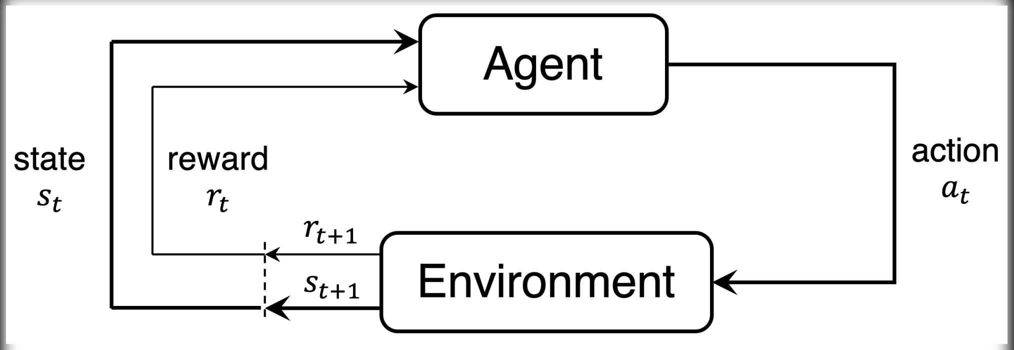
Quantum Machine Learning examples:

Reinforcement Learning

Schenk, M *et al.* Hybrid actor-critic algorithm for quantum reinforcement learning at CERN beam lines. *arXiv preprint arXiv:2209.11044.*, CHEP2023

Reinforcement learning

… in a nutshell



RL book: Sutton & Barto

Trial-and-error learning

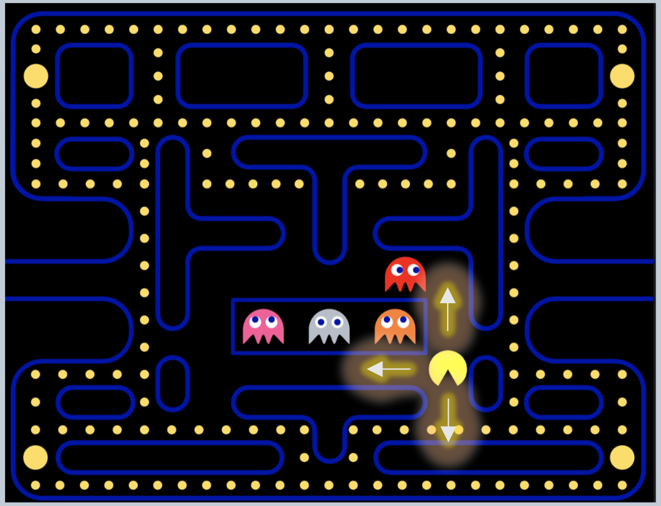
- Agent **takes actions** in environment and **collects rewards**

Q-learning

- Estimate **return using Q-function** $\mathrm{Q}(\boldsymbol{s}, \mathrm{a})$
- **Learn iteratively** using **collected interactions**
- Once trained, **select action greedily**

$$a = \arg\max_{\mathrm{a}} Q(\boldsymbol{s}, \boldsymbol{a})$$

Example: Pacman



State
where am I? Where are ghosts, snacks, cookies?

Actions
up, down, left, right

Reward
food (+), ghosts (-)

Return
how much food am I going to eat over time

Free-energy based RL (FERL)

RL performance depends **on type of Q-function approximator**

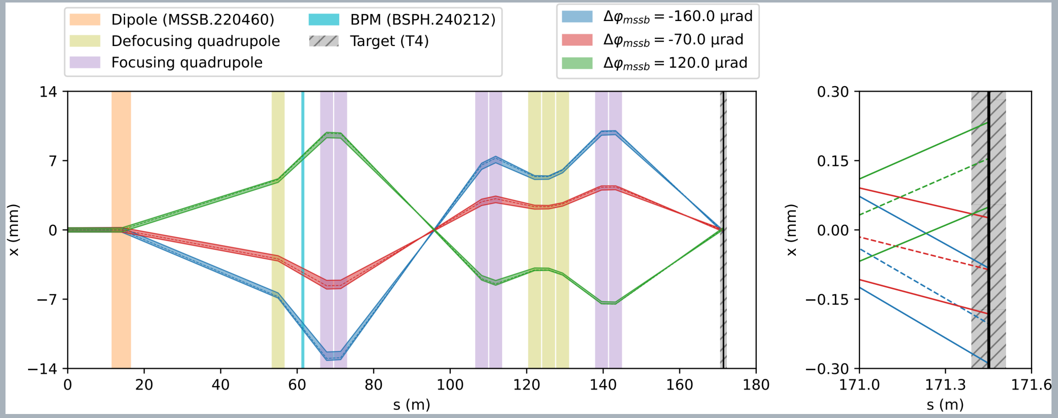
- **Classical Deep Q-learning (DQN)**
  Feed-forward neural net
- **Free-energy based RL (FERL)**
  Quantum Boltzmann machine (QBM)

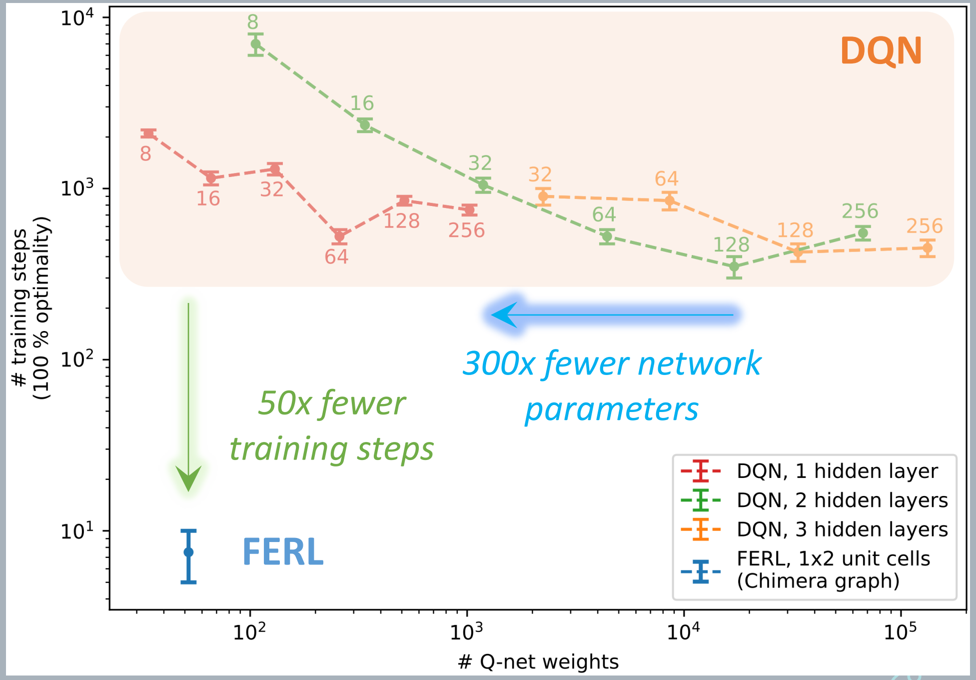
Key concept: sample-efficiency

- Relevant for **particle accelerator control** given cost of beam time *(online training)*

1st study: 1D beam steering

CERN *North Area transfer line (discrete action space)*

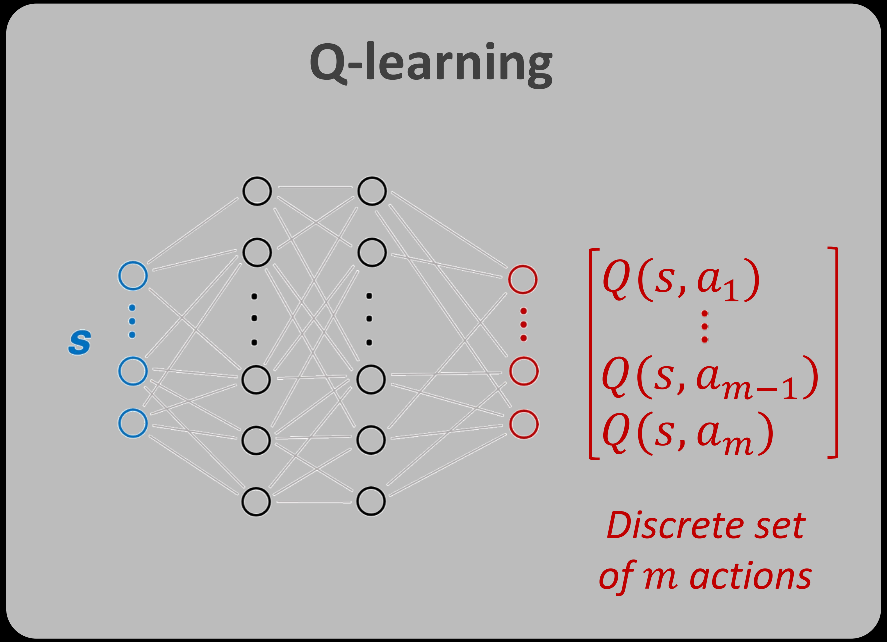


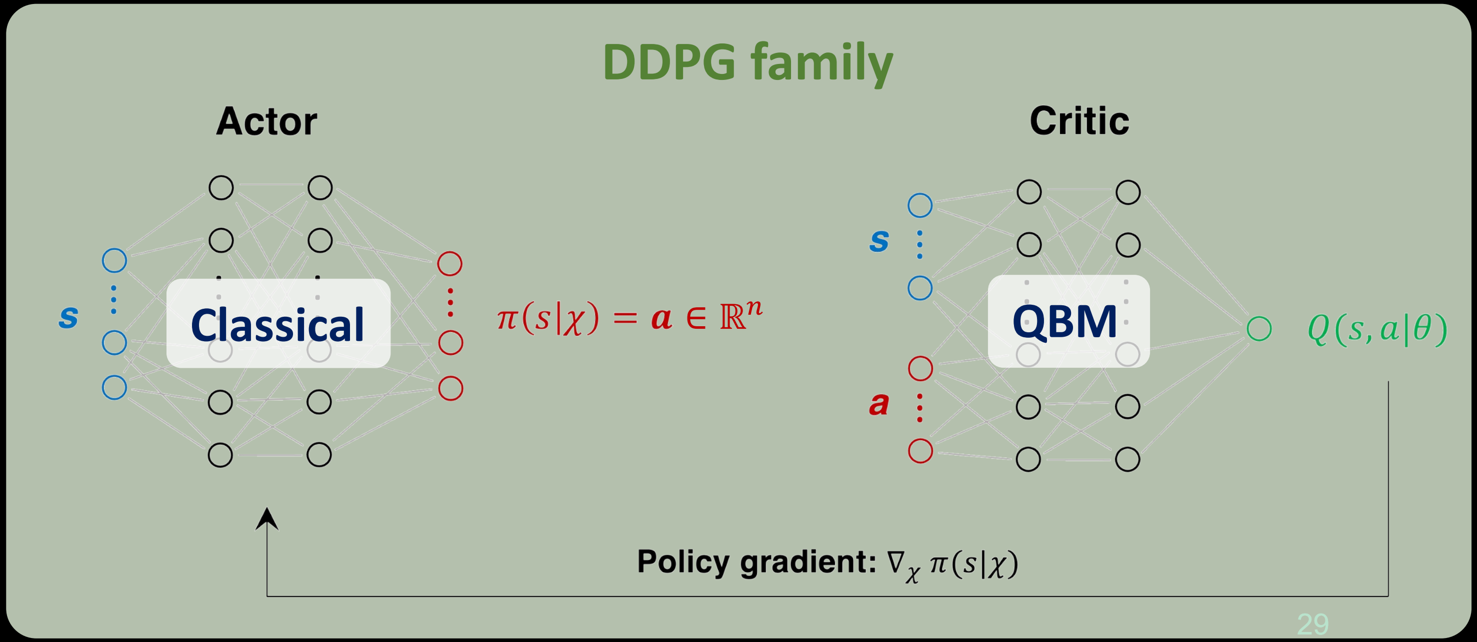


Developing a hybrid actor-critic scheme

Accelerator optimization requires **continuous action space** ➡ **develop hybrid actor-critic algorithm**

- ➢ **QBM replaces classical critic net**



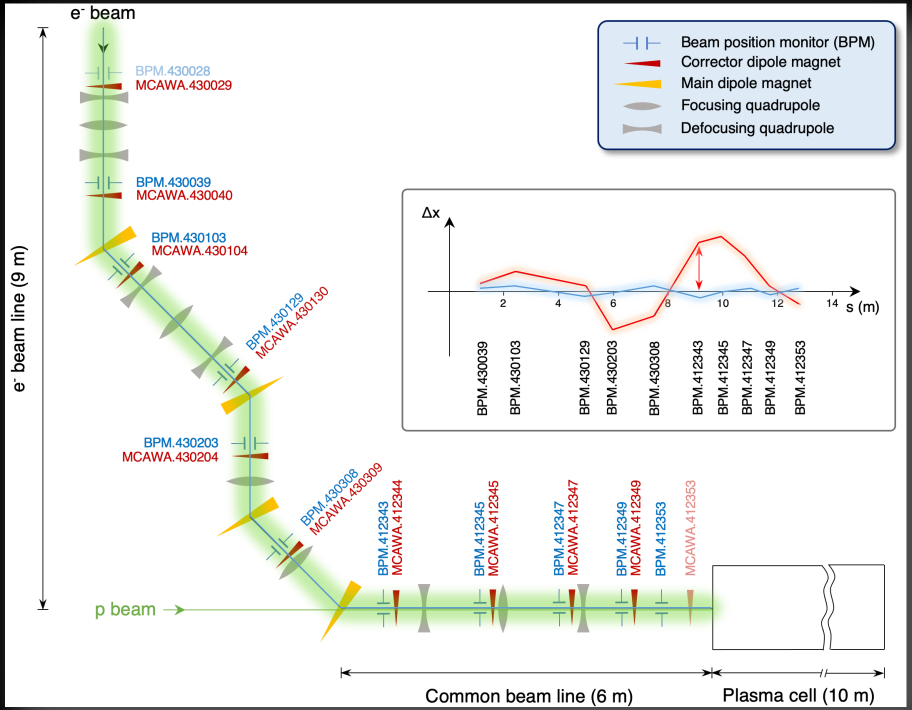


2nd study: 10D continuous beam steering

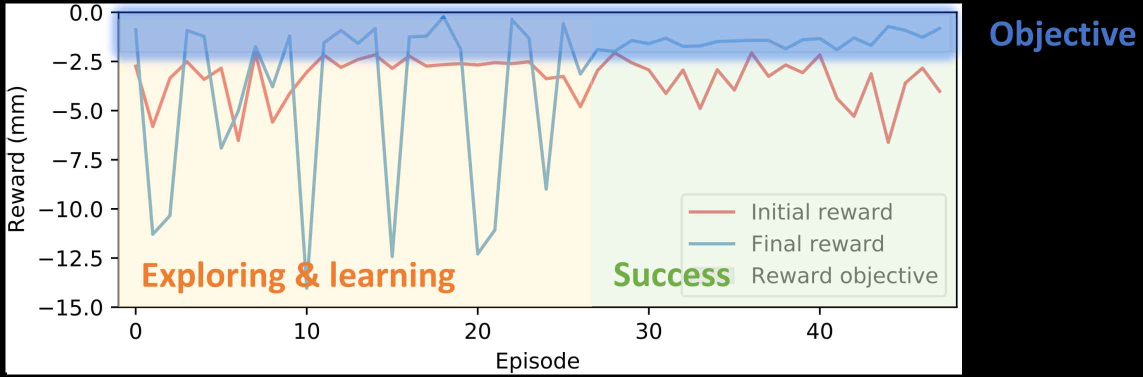
Environment: e⁻ beam line of AWAKE

- **Action:** deflection angles at 10 correctors
- **State:** beam positions at 10 BPMs
- **Objective:** minimize beam trajectory rms

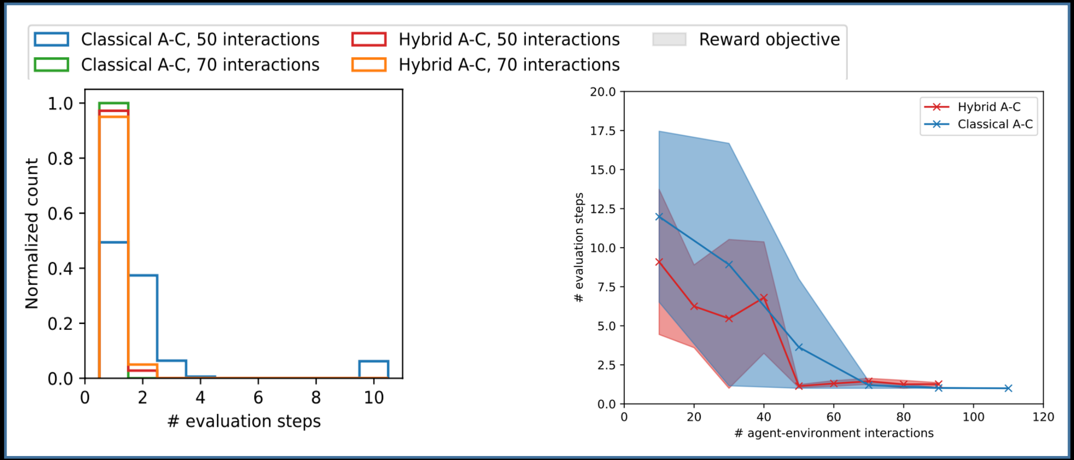
AWAKE: Advanced Proton Driven Plasma Wakefield Acceleration Experiment



Training: on D-Wave Advantage quantum annealer (QA)



Evaluation: on actual beam line

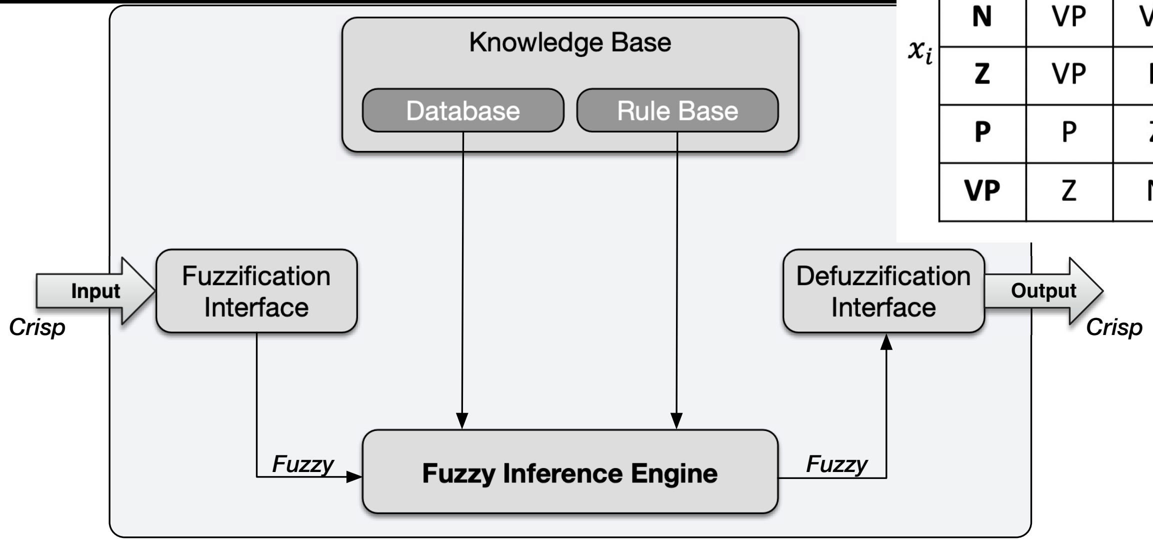


- Agent minimizes rms in **1 step in 60 % cases**
- **Minor improvement with respect to classical**
- **Linear dynamics (too simple?)**

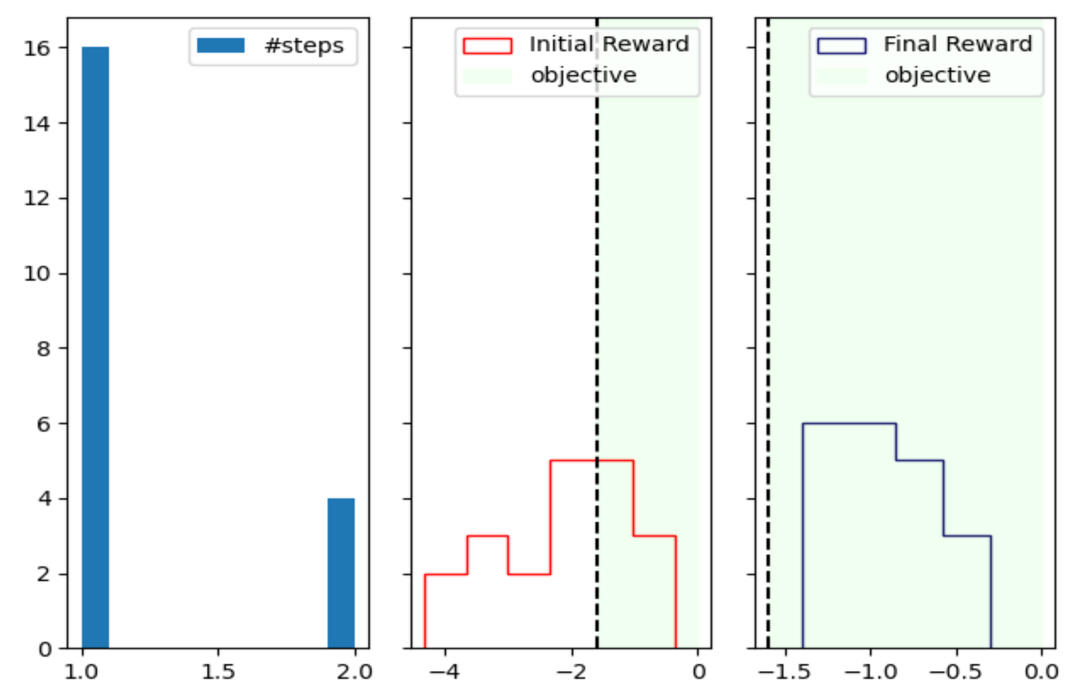
1-slide excursion: quantum fuzzy logic controller

- **Alternative control algorithm** to RL
- **Fuzzy Logic** is used to develop control systems **based on linguistic rules** ➡ **highly interpretable**
- **Quantum Fuzzy Control System** *(G. Acampora, R. Schiattarella, A. Vitiello)*
  Exploit **exponential advantage** in computing fuzzy rules on quantum computers
- **Successfully evaluated on AWAKE beam line, no training required**

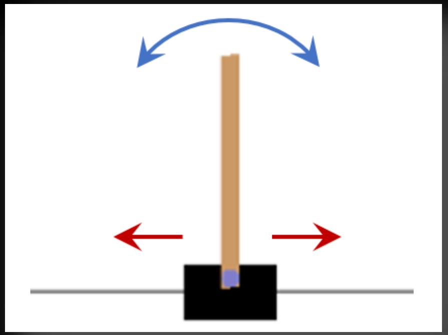
Evaluation: on AWAKE beam line
Objective reached typically in 1 step



x_i \ dk_i	**VN**	**N**	**Z**	**P**	**VP**
VN	VP	VP	VP	P	Z
N	VP	VP	P	Z	N
Z	VP	P	Z	N	VN
P	P	Z	N	VN	VN
VP	Z	N	VN	VN	VN



3rd study: Cartpole-v1

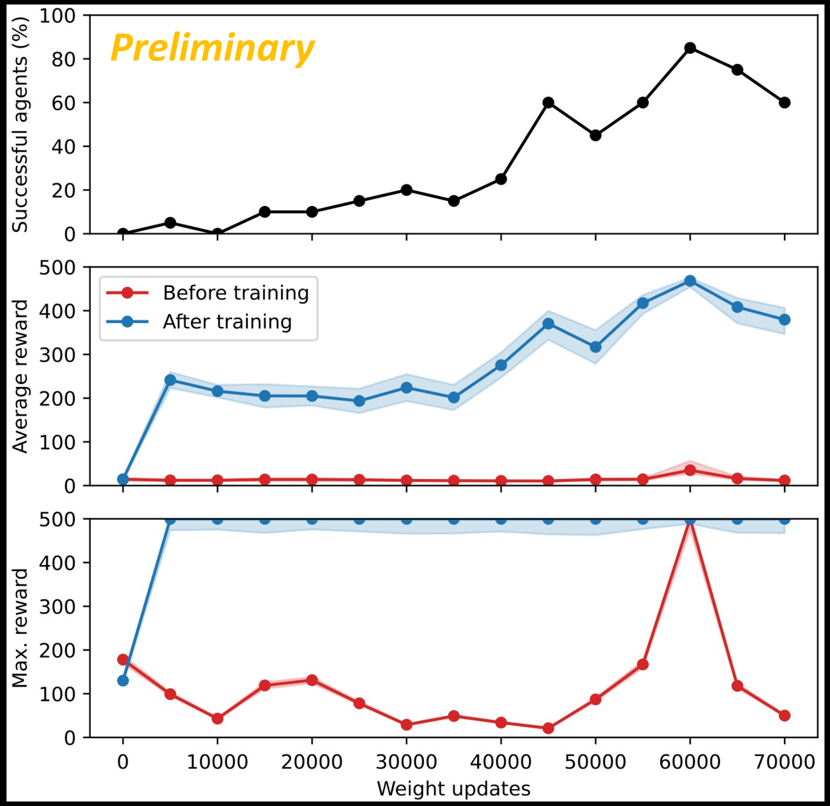


Discrete action problem, non-linear dynamics

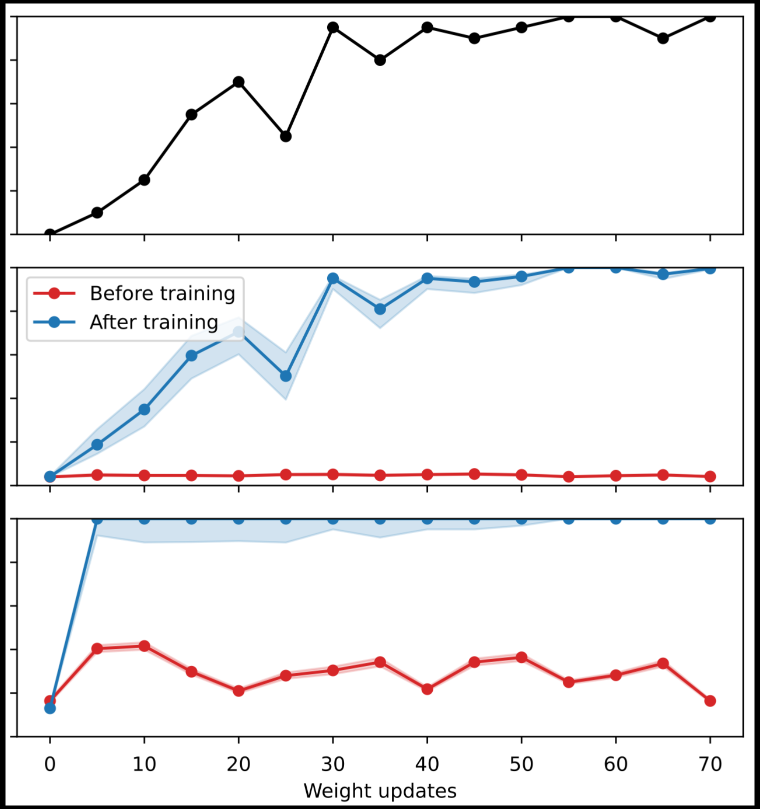
- **Cartpole-v1:** official OpenAI gym env from classic control problems domain
- Continuous state (4D), **discrete action** *(right, left)* problem with **non-linear dynamics**
- Terminate episodes after **max. 500 steps**

- **Big gain in sample-efficiency** and **robustness** for FERL vs DQN

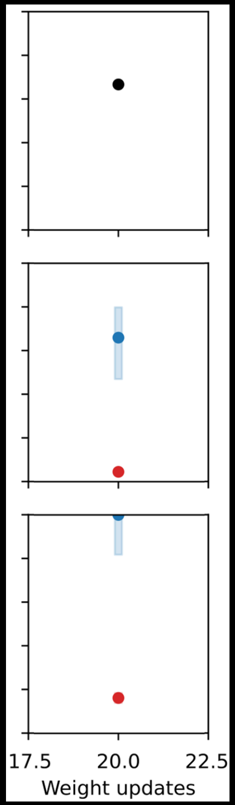
DQN

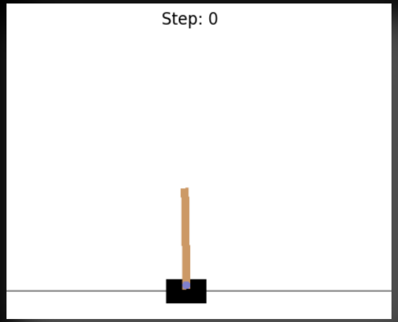


FERL *(simulated QA)*

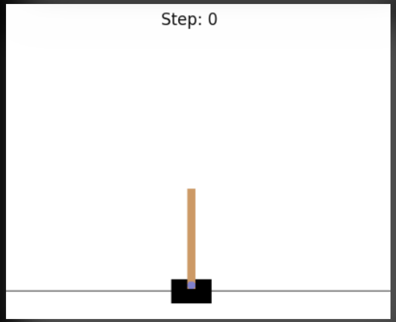


FERL *(trained on D-Wave)*





before



after

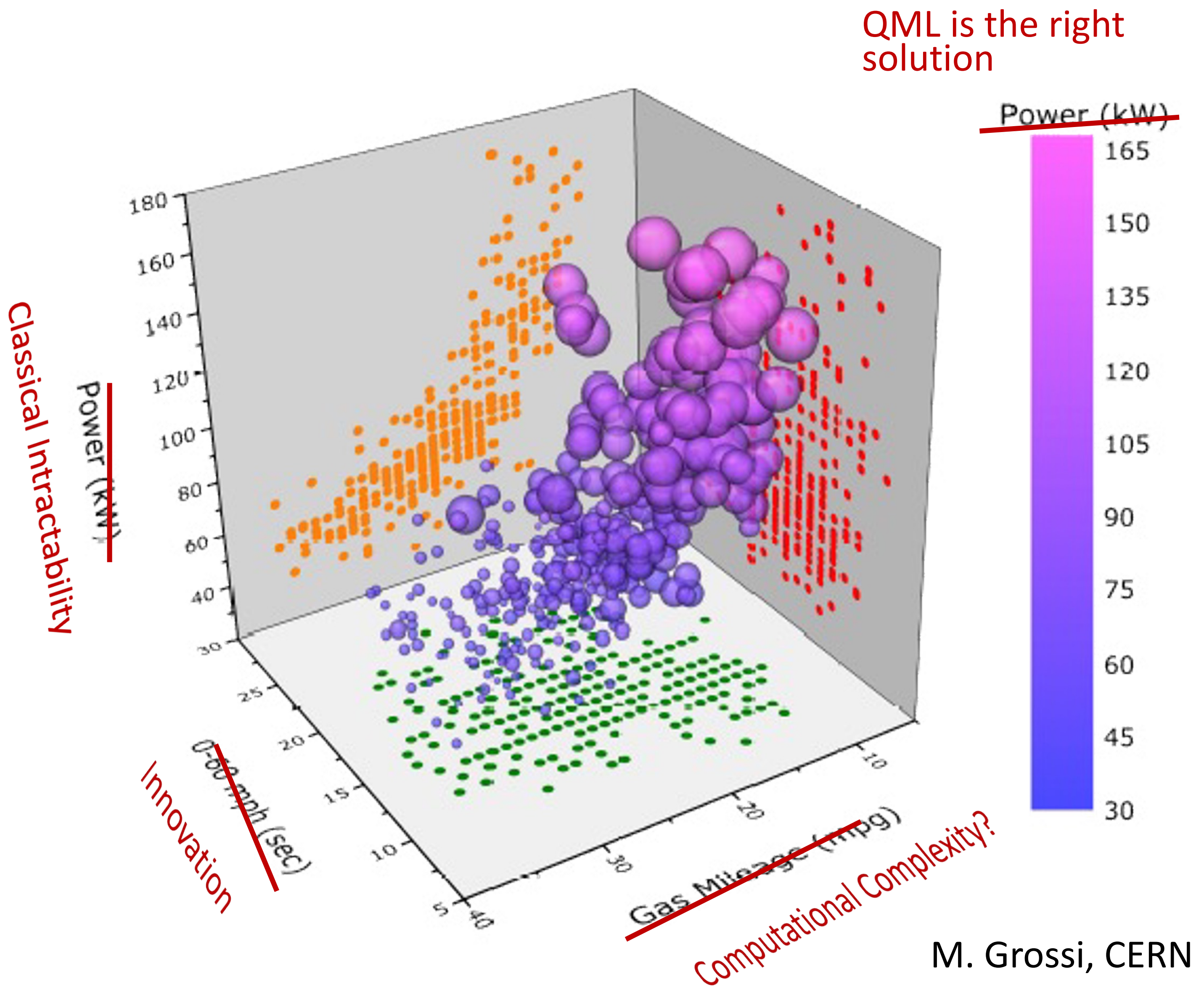
Improving Robustness of QML applications

- Understanding conditions to advantage
- Stabilizing training on NISQ (*arXiv:2212.11826, arXiv:2303.11283*)
- Trainability vs expressivity for generative models (*arXiv:2305.02881*)
- Evaluating generalisation
- Quantum vs classical data, phase transitions (*Physical Review B, 107*(8), L081105)
- Algorithms beyond QML (*Physical Review C, 106*(3), 034325.)

Outlook and open questions

- HEP provides challenges to Quantum Computing
  - **What are the most promising applications?**
  - How do we define performance and validate results on **realistic use cases**?
- Experimental data has high dimensionality
  - Can we train Quantum Machine Learning algorithms effectively?
  - Can we reduce the impact of **data reduction** techniques?
- Experimental data is shaped by physics laws
  - Can we leverage them to build better algorithms?

QML Exclusion Region in HEP?



M. Grossi, CERN

Thank you!

November 20th-24th, 2023
@CERN

QTML
Quantum Techniques in Machine Learning

Sofia.Vallecorsa@cern.ch

Model Convergence and Barren Plateau

The size of the Hilbert space requires compromises between **expressivity**, **convergence** and **generalization**

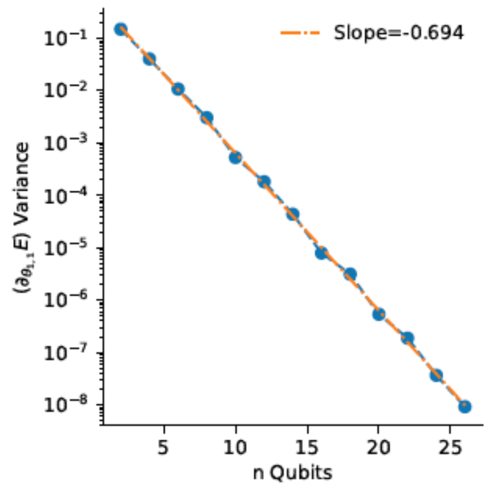
Classical gradients **vanish exponentially** with the number of layers (J. McClean *et al.*, arXiv:1803.11173)

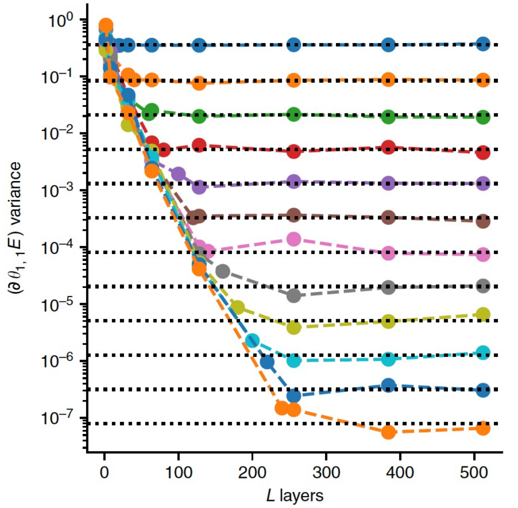
- Convergence still possible if gradients consistent between batches.

Quantum gradient decay exponentially in the number of qubits

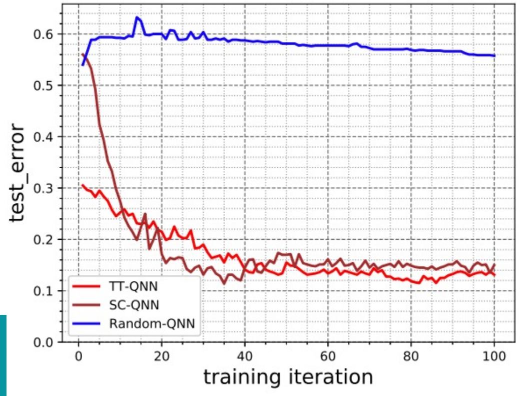
- Random circuit initialization
- Loss function locality in shallow circuits (M. Cerezo *et al.*, arXiv:2001.00550)
- Ansatz choice: TTN, CNN (Zhang *et al.,* arXiv:2011.06258, A Pesah, *et al.*, *Physical Review X* 11.4 (2021): 041011.)
- Noise induced barren plateau (Wang, S *et al.*, Nat Commun 12, 6961 (2021))

J. McClean *et al.*, arXiv:1803.11173

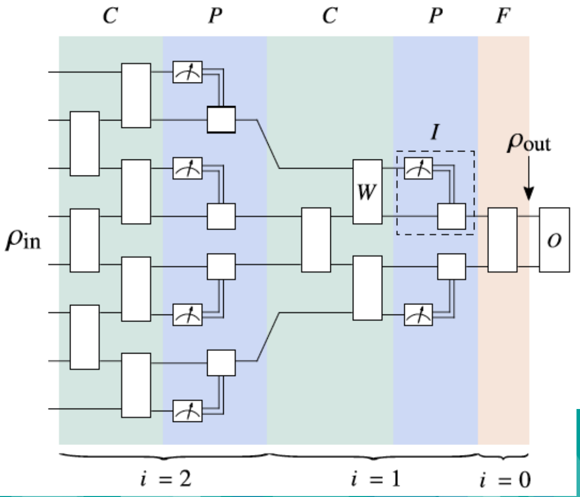




TTN for MNIST classification (8 qubits), Zhang *et al.,* arXiv:2011.06258



QCNN: A Pesah, *et al.*, *Physical Review X* 11.4 (2021): 041011



Kernel trainability and kernel concentration

Kernel values can **concentrate exponentially** around a common value

Need **exponentially larger number of measurements** to resolve

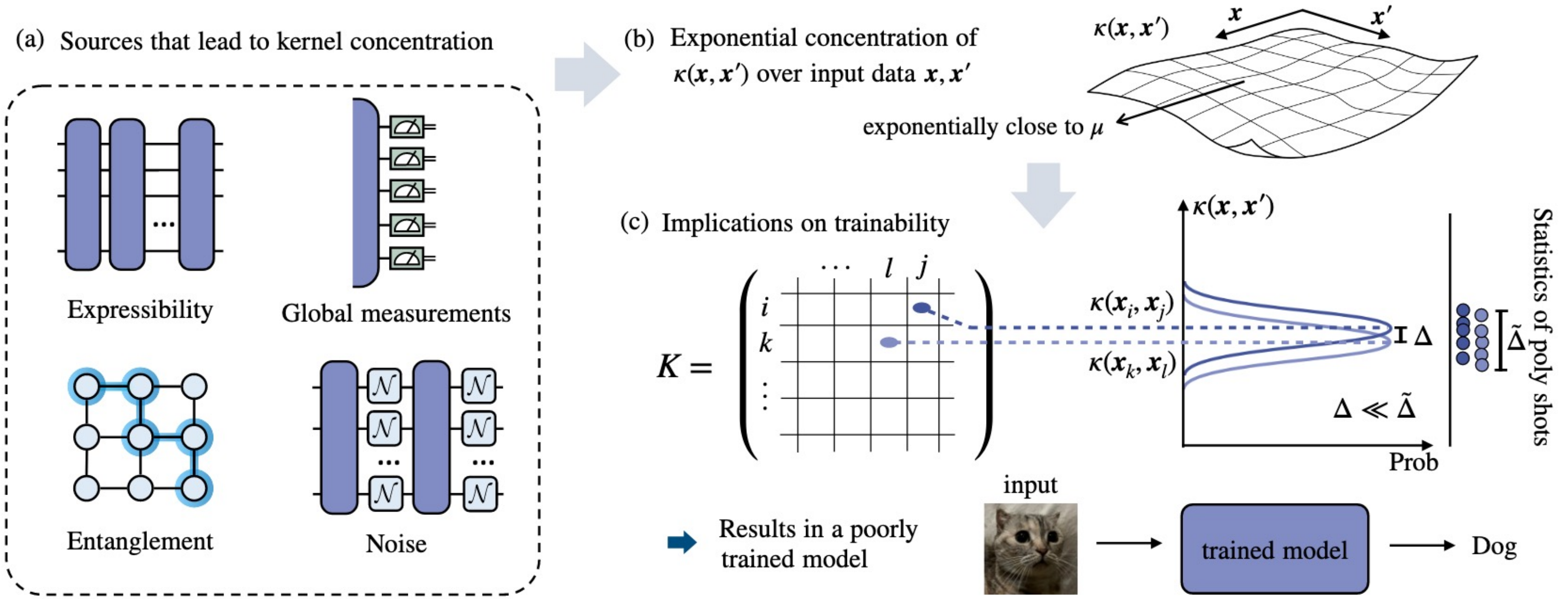
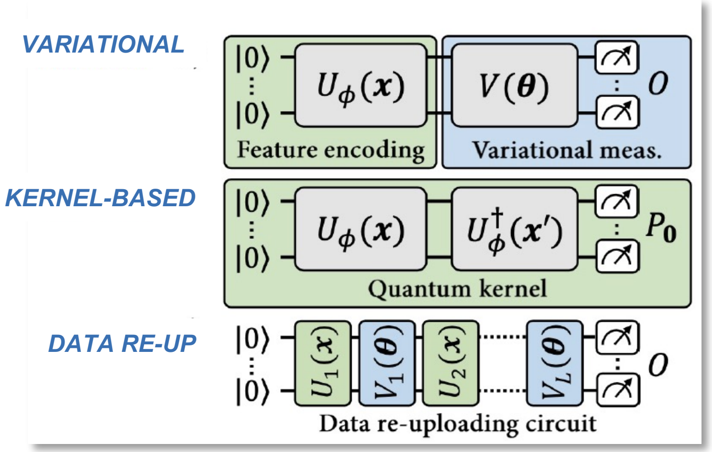


Figure 1. **Kernel concentration and its implications on trainability:** The exponential concentration (in the number of qubits n) of quantum kernels $\kappa(\boldsymbol{x}, \boldsymbol{x}')$, over all possible input data pairs $\boldsymbol{x}, \boldsymbol{x}'$, can be seen to stem from the difficulty of information extraction from data quantum states due to various sources (illustrated in panels (a) and (b)). The kernel concentration has a detrimental impact on the trainability of quantum kernel-based methods. As shown in panel (c), for a polynomial (in n) number of measurement shots, the sampling noise $\tilde{\Delta}$ dominates for large n and, as $\Delta \ll \tilde{\Delta}$, $\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j)$ cannot be resolved from some other $\kappa(\boldsymbol{x}_k, \boldsymbol{x}_l)$, leading to a poorly trained model.

Study kernel trainability in our Anomaly Detection model (arxiv:2208.11060)

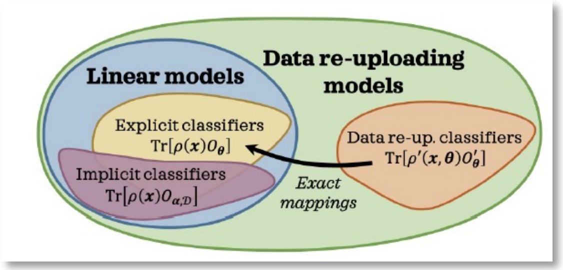
Equivalent interpretations?

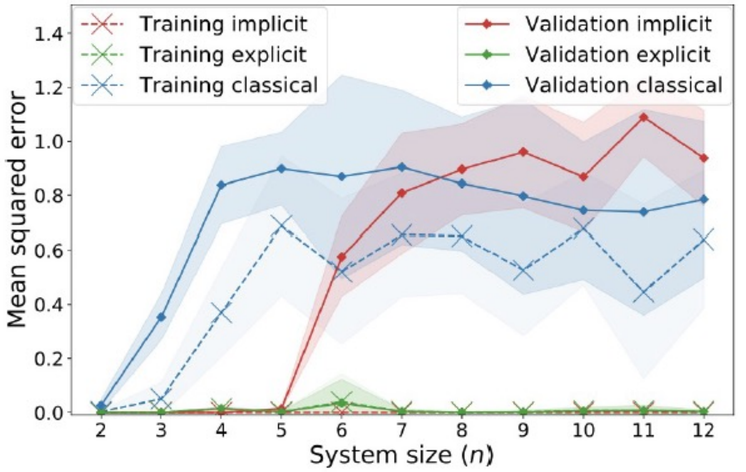


Characterize models behaviour, similarities among them and **link to data properties.**

Ex:

- Data Re-Uploading circuits: alternating data encoding and variational layers.
    - Represented as **explicit linear models** (variational) in larger feature space
    - → can be reformulated as **implicit models** (kernel)
- Representer theorem: implicit models achieve **better accuracy**
    - Explicit models exhibit **better generalization** performance





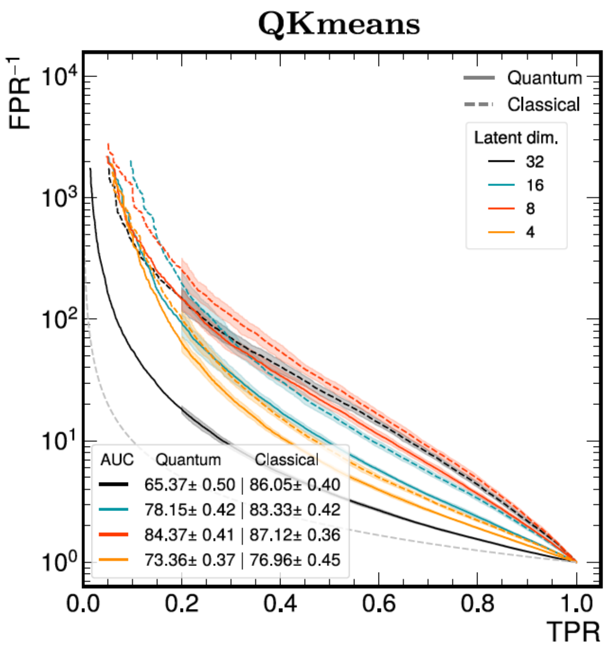
PCA on 28x28 fashion-MNIST dataset, ZZ feature encoding + hardware-efficient variational unitary

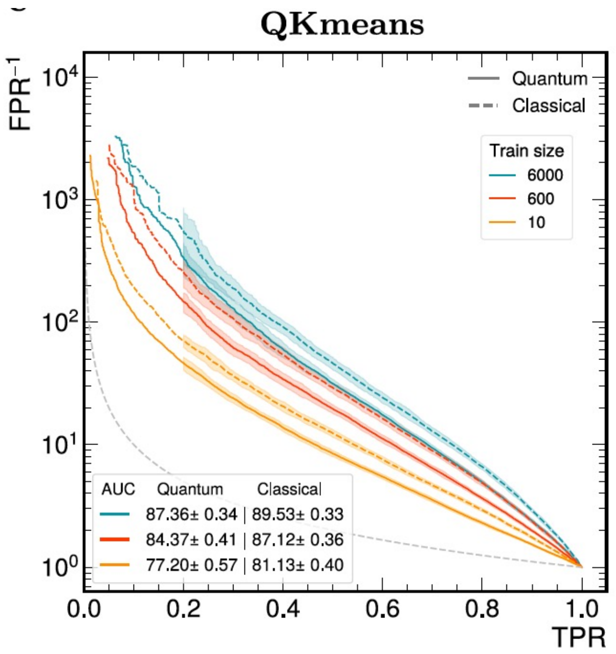
Jerbi, Sofiene, et al. "**Quantum machine learning beyond kernel methods**." *arXiv preprint arXiv:2110.13162* (2021).

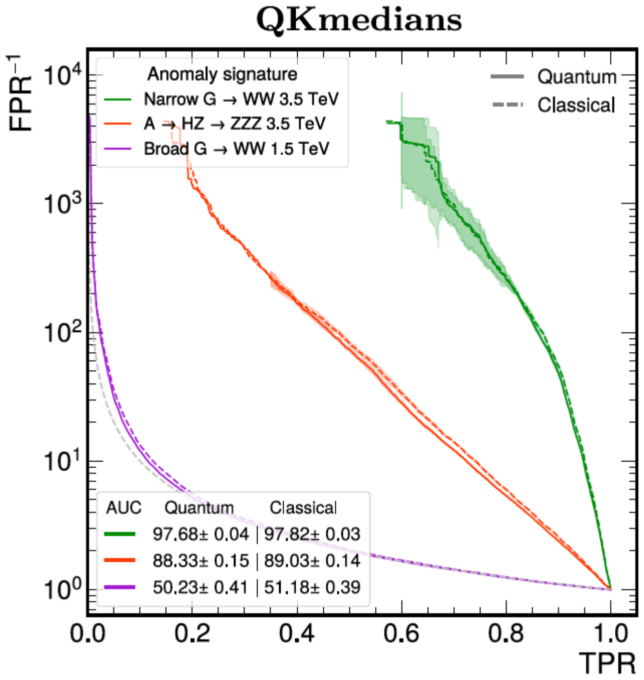
Comparison to unsupervised clustering

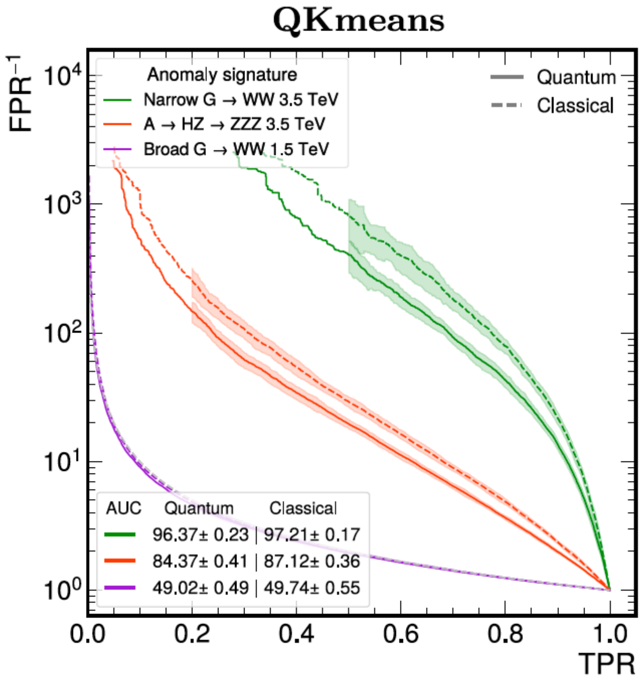
Quantum* clustering algorithms **do not outperform classical** counterpart

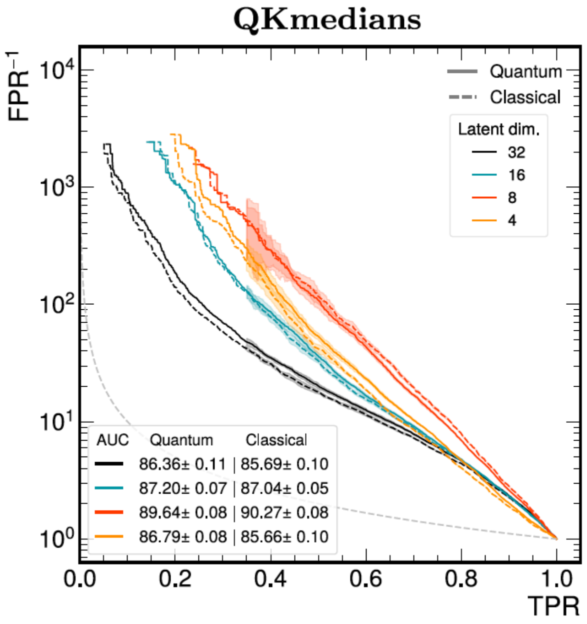
QMEANS performs worst

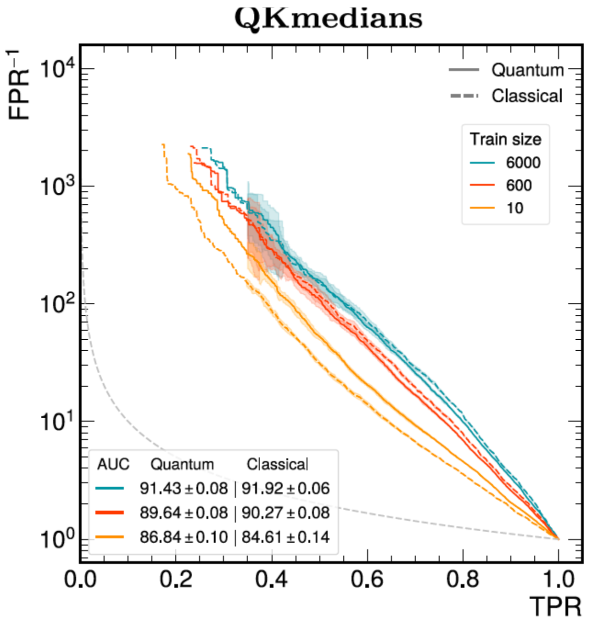








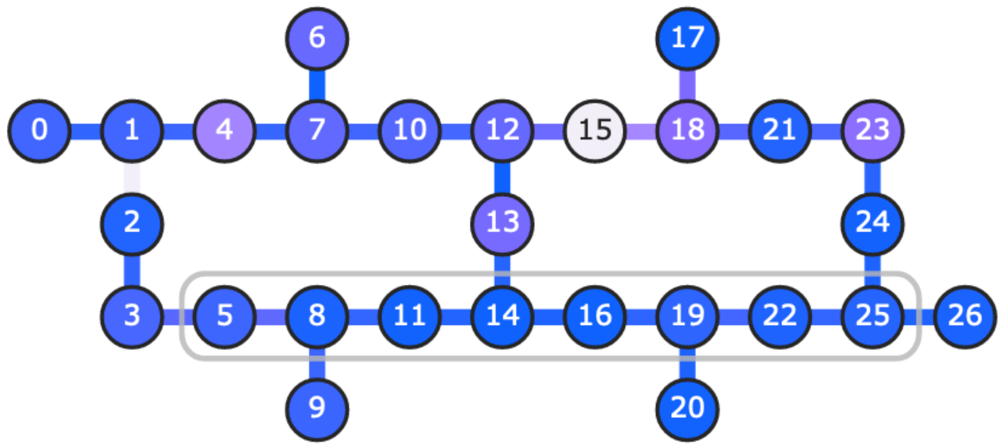




Preliminary hardware runs

Stable performance on ***ibmq_toronto*** :

- Design circuit taking **qubits topology** into account
- Use **8 qubits** and native gates
- **Reduced training set size** (100) → increased statistical uncertainty
- Use **AUC** (less affected by statistics)
- Monitor **mean purity of states** to verify state coherence during computation
  - Fully mixed state yields a purity of 0.39 10^{-2} ($1/2^n$)



Kernel Machine Run	AUC	$\langle \mathrm{tr}\rho^2 \rangle$
Hardware $L = 1$	0.844	0.271(6)
Ideal $L = 1$	0.999	1
Hardware $L = 3$	0.997	0.15(2)
Ideal $L = 3$	1.0	1
Classical	0.998	-

Incudini, M. , *et al*. "**Resource Saving via Ensemble Techniques for Quantum Neural Networks**." *arXiv:2303.11283* (2023).

Ensembles of quantum neural networks

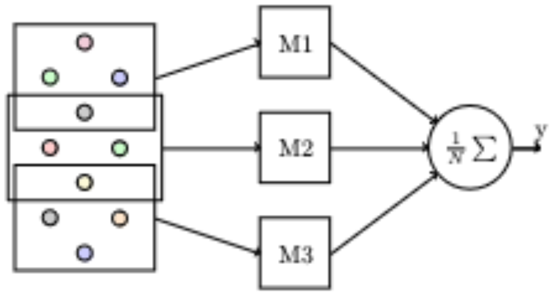
Bagging: best for **high variance**; reduces BPs by keeping the feature space limited

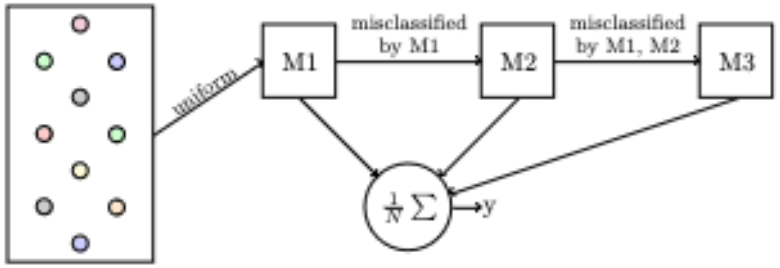
- 10 independently trained instances
- r_f :% of samples, r_n:% features

Boosting: **high bias** models (little sensitivity to subsampling)

- AdaBoost, 10 repetitions

Study **regression** and **classification** tasks in toy and realistic datasets

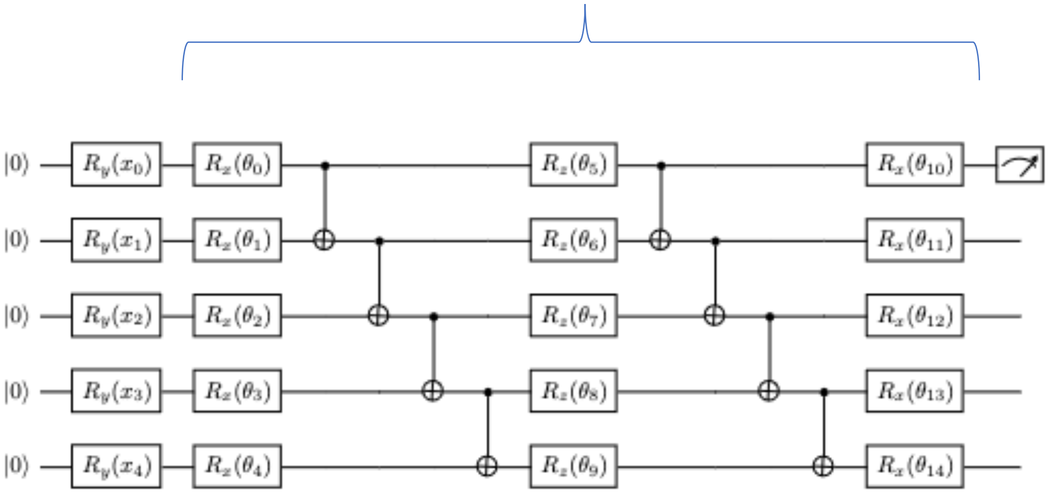




Dataset	Source	Nature	# Features	# Samples	Task
Linear	-	Synthetic	5	250	Regression
Concrete	UCI	Real-world	8	1030	Regression
Diabetes	Scikit-Learn	Real-world	10	442	Regression
Wine	UCI	Real-world	13	178	Classification

QNN setup and simulated results

1 layer



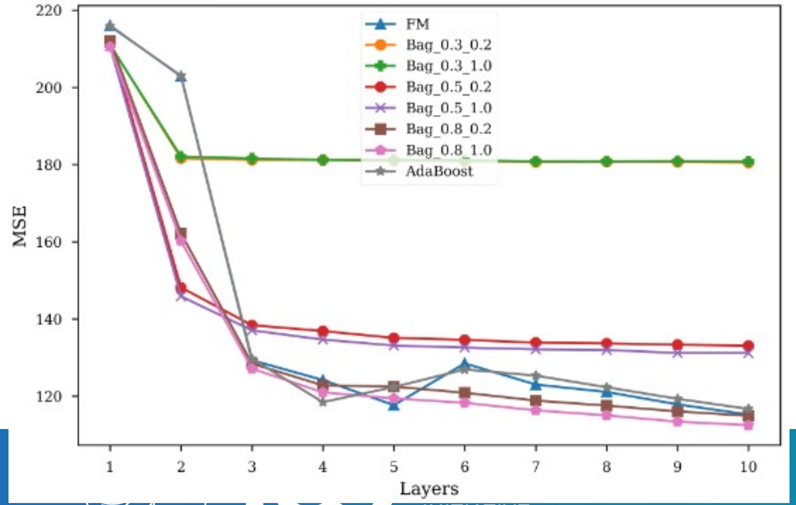
Choose **relatively simple QNN**:

- n qubits = n features
- Ry single rotation gates
- CNOT in linear entanglement
- Local observable (σ_z)

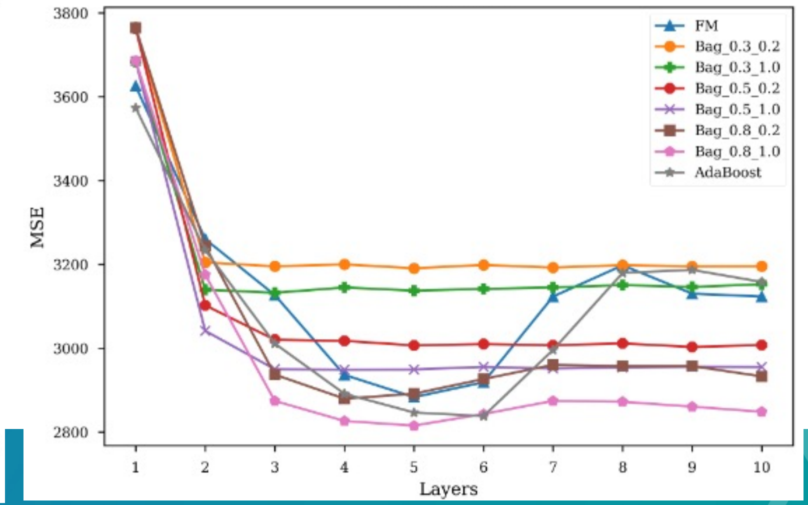
Measure the generalisation error on **test sample** (20 %)

Bagging methods outperform full model and Boosting: **shallower networks**, **fewer input features**

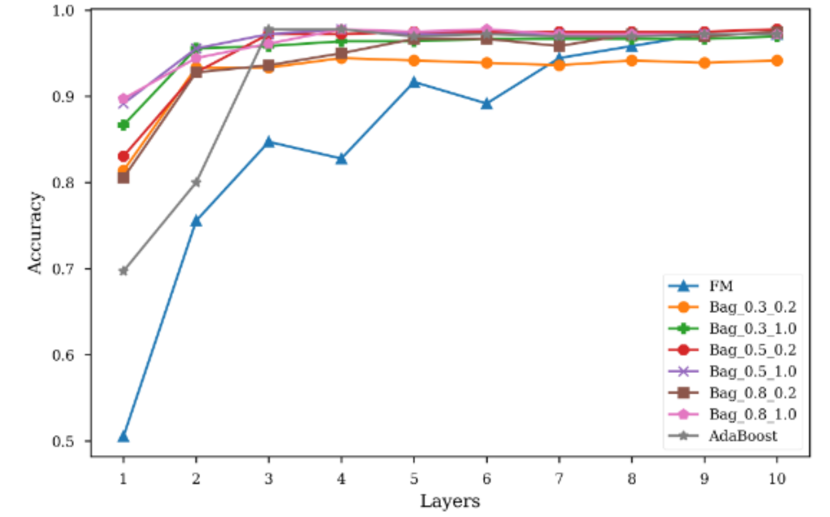
Concrete (MSE)



Diabetes (MSE)

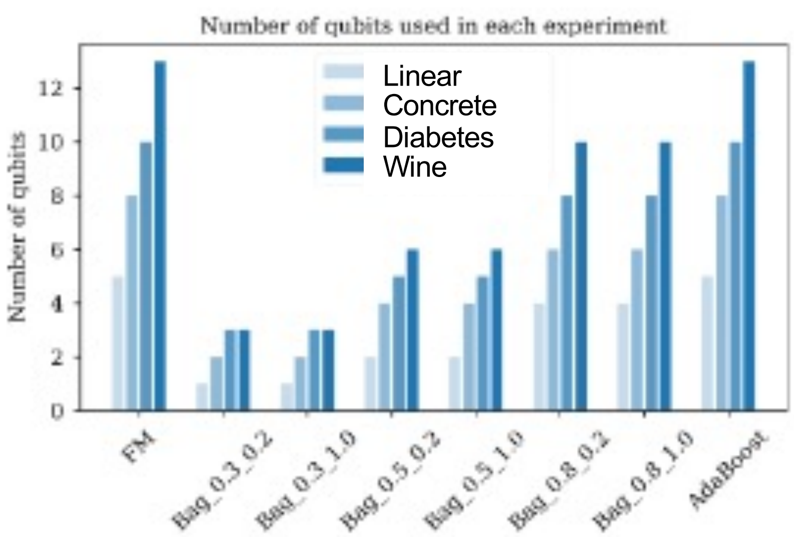


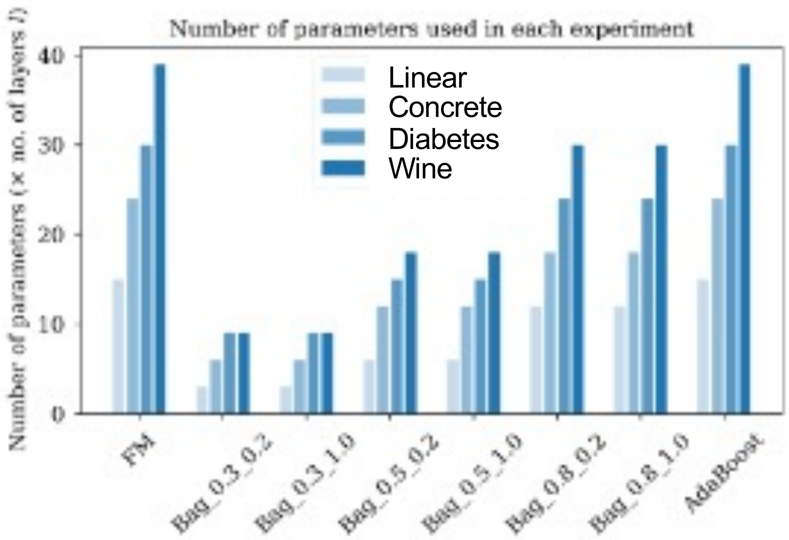
Diabetes (CCE)



Bagging brings significant advantage

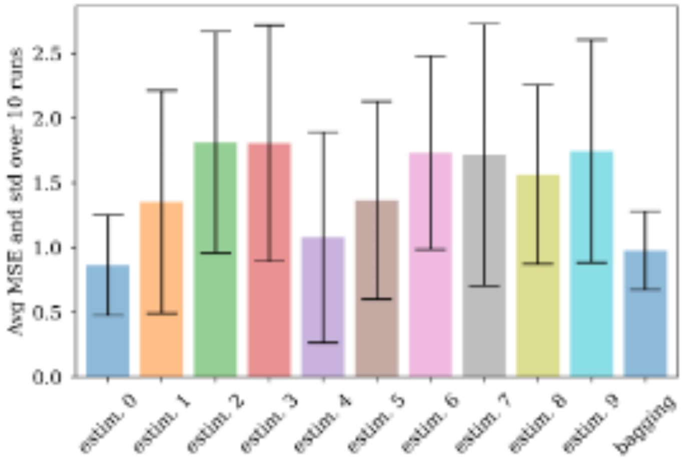
Reducing resources:
Best performance for low dimensionality

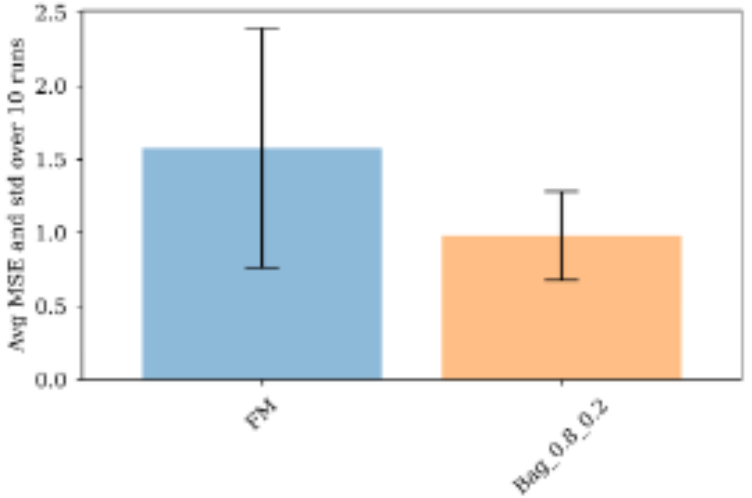




Robustness against noise:
Linear regression task on **IBM QPU** (ibm_lagos):

> **Bagging:** 80% features, 20% samples
> **QNN**: 4 qubit, 1 layer



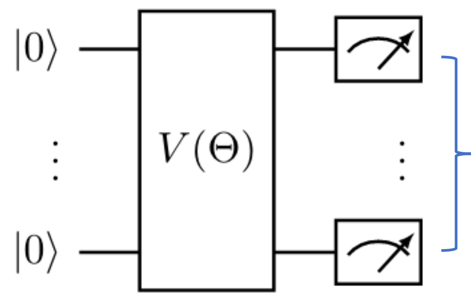


Quantum Generative Models

Delgado and Hamilton, arXiv:2203.03578 (2022)
Zoufal, et al., *npj Quantum Inf* **5**, 103 (2019)
Leadbeater et al., *Entropy* **2021**, *23*, 1281.
Amin, et al. *Physical Review X* 8.2 (2018): 021050.

QCBM

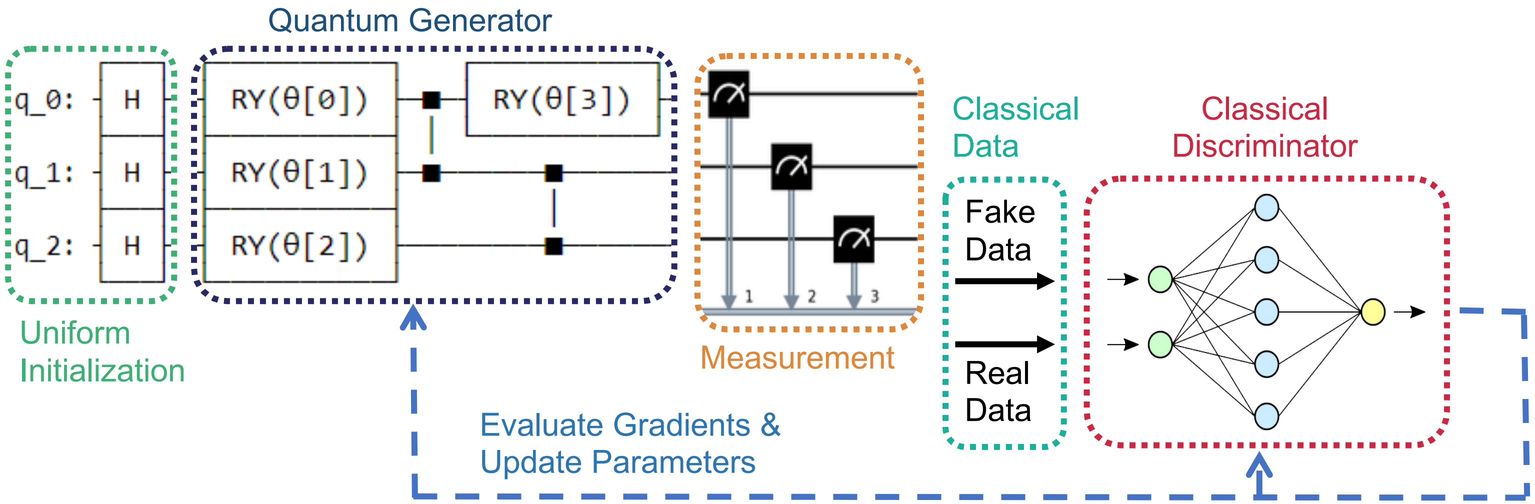
Sample variational pure state $|\psi(\theta)\rangle$ by projective measurement through **Born rule**: $\mathbf{p}_{\boldsymbol{\theta}}(\mathbf{x}) = |\langle\mathbf{x}|\boldsymbol{\psi}(\boldsymbol{\theta})\rangle|^2$.



n dimensional binary strings map to 2^n bins of the discretized dataset.

QGAN

Multiple implementations, mostly classical-quantum hybrid



QBM

Network of stochastic binary units with a quadratic energy function that follows the Boltzman distribution (Ising Hamiltonian)

$$H = -\sum_a b_a \sigma_a^z - \sum_{a,b} w_{ab}\sigma_a^z \sigma_b^z$$

Typical metrics:

$$D_{\mathrm{KL}}(P\|Q) = \sum_i P(i)\,\log\left(\frac{P(i)}{Q(i)}\right)$$

$$\mathrm{MMD}(\mathbb{P}_r, \mathbb{P}_g) = \left(\mathbb{E}_{\substack{\mathbf{x}_r, \mathbf{x}_r' \sim \mathbb{P}_r, \\ \mathbf{x}_g, \mathbf{x}_g' \sim \mathbb{P}_g}}\left[k(\mathbf{x}_r, \mathbf{x}_r') - 2k(\mathbf{x}_r, \mathbf{x}_g) + k(\mathbf{x}_g, \mathbf{x}_g')\right]\right)^{\frac{1}{2}}$$

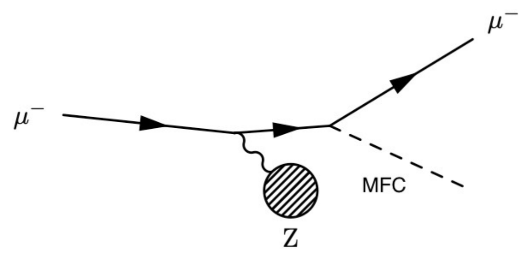
QCBM for event generation

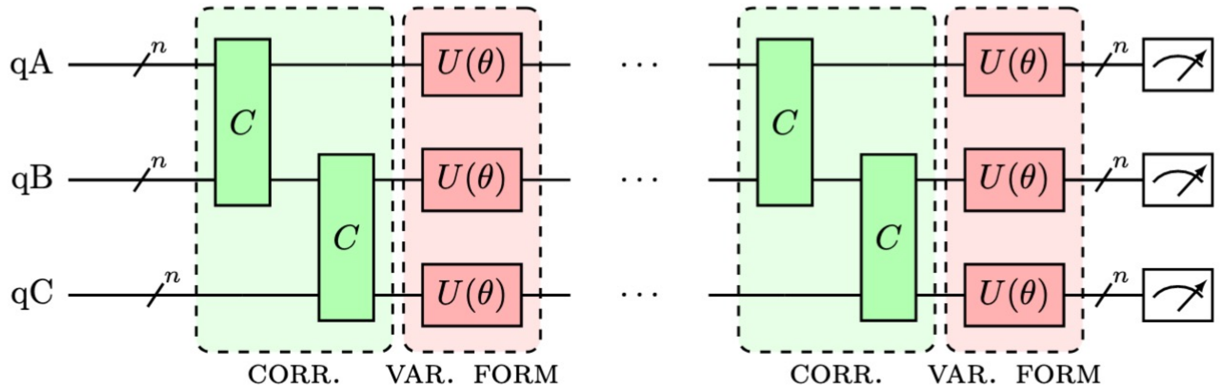
UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

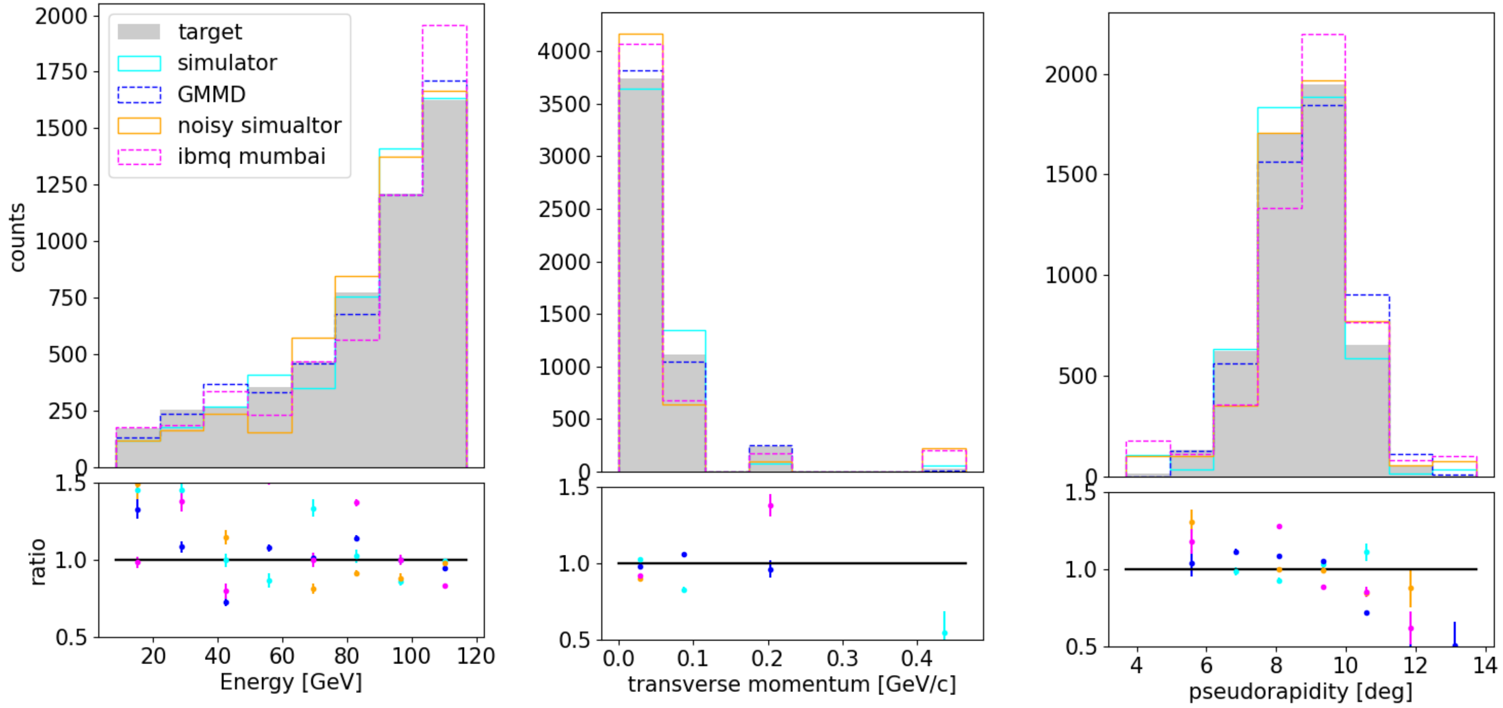
Muon Force Carriers, in muon fixed-target experiments (FASER) or muon interactions in calorimeters (ATLAS)[1].

Generate multivariate distribution (E, p_t, η)

Maximum Mean Discrepancy for training







1 Galon, I, Kajamovitz, E et al. "**Searching for muonic forces with the ATLAS detector**". In: Phys. Rev. D 101, 011701 (2020)

qGAN for event generation

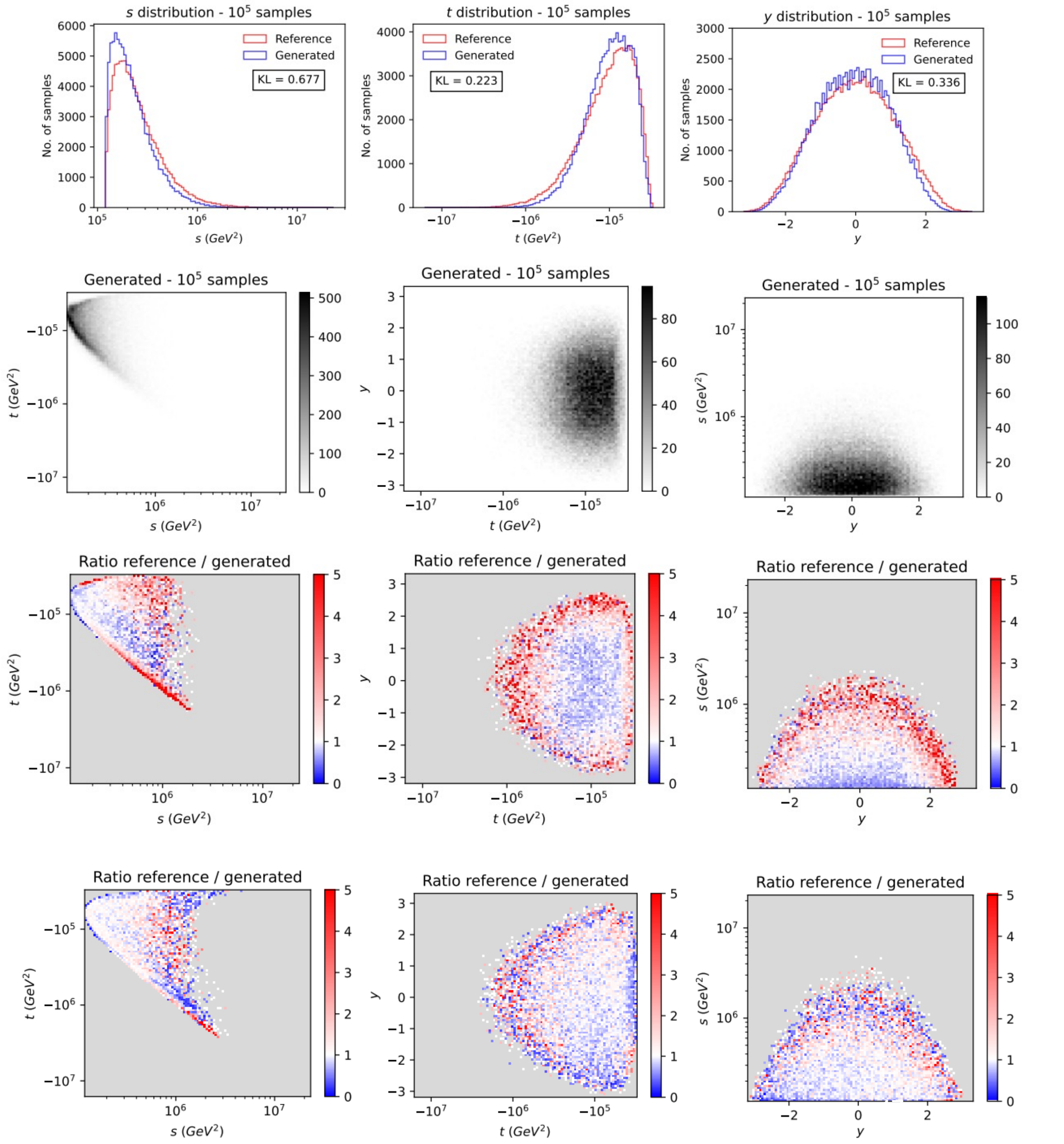
Generate Mandelstam (s,t) + y variables for **t-tbar production**

Introduce a **style-based** approach

	$pp \rightarrow t\bar{t}$ **LHC events**
Qubits	3
D_{latent}	5
Layers	2
Epochs	3×10^4
Training set	10^4
Batch size	128
Parameters	62
U_{ent}	2 sequential CR_y gates

IBM Q Santiago

Quantum simulator

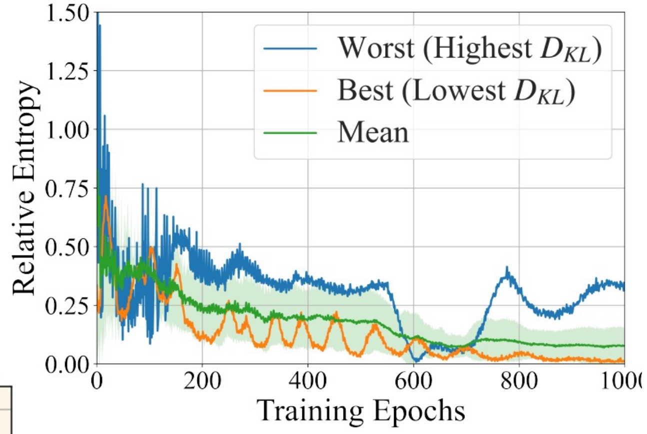


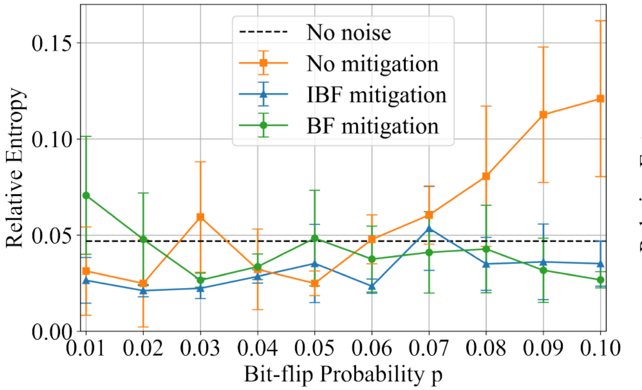
Bravo-Prieto et al. "**Style-based quantum generative adversarial networks for Monte Carlo events.**" Quantum 6, 777 (2022) , *arXiv preprint arXiv:2110.06933* (2021).

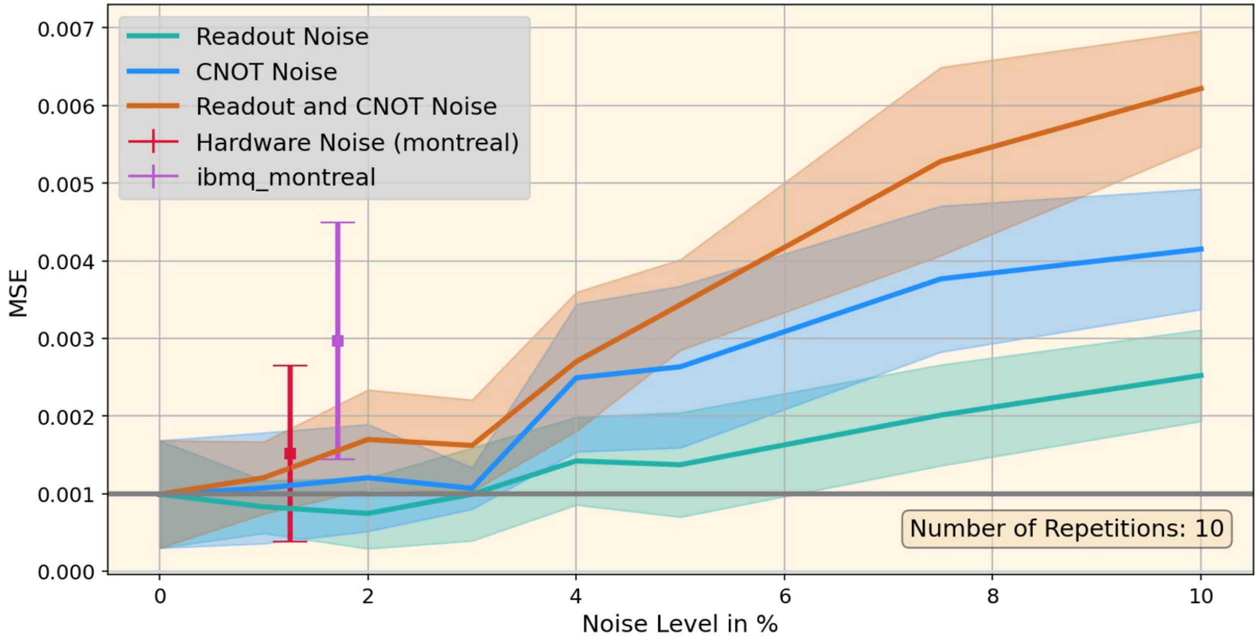
Borras, Kerstin, et al. "**Impact of quantum noise on the training of quantum Generative Adversarial Networks**." ACAT2021, *arXiv preprint arXiv:2203.01007* (2022).

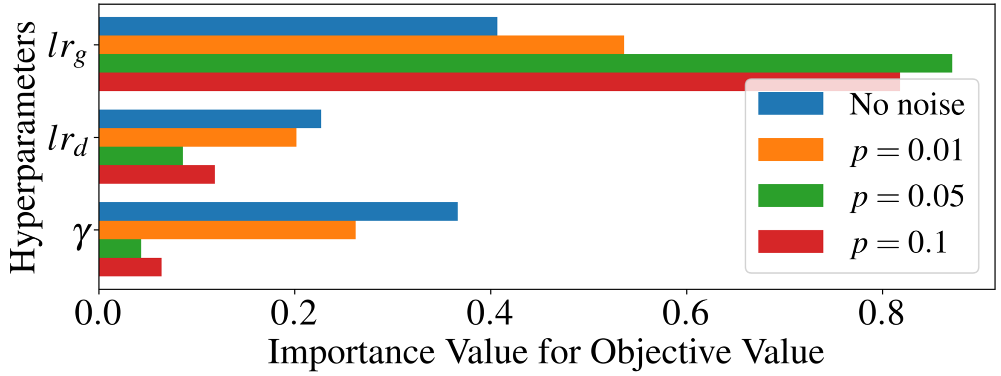
Robustness against noise

QML training process seems **robust against noise** (error mitigation is needed in extreme cases)









Chang S.Y. *et al.*, Running the Dual-PQC GAN on Noisy Simulators and Real Quantum Hardware, QTML2021, ACAT21

qGAN Benchmarks on hardware

Train models using **noisy simulator** and test the inferen **trapped-ion (IONQ) quantum hardware**

- For IBMQ machines, choose the qubits with the lowest

Device	Readout error CX error	$D_{KL}/D_{KL,ind}$ $(\times 10^{-2})$
ibmq_jakarta	0.028 $1.367 \cdot 10^{-2}$	0.14 ± 0.14 6.49 ± 0.54
ibm_lagos	0.01 $5.582 \cdot 10^{-3}$	0.26 ± 0.11 6.92 ± 0.71
ibmq_casablanca	0.026 $4.58 \cdot 10^{-2}$	4.03 ± 1.08 6.58 ± 0.81
IONQ	NULL $1.59 \cdot 10^{-2}$	1.24 ± 0.74 10.1 ± 5.6

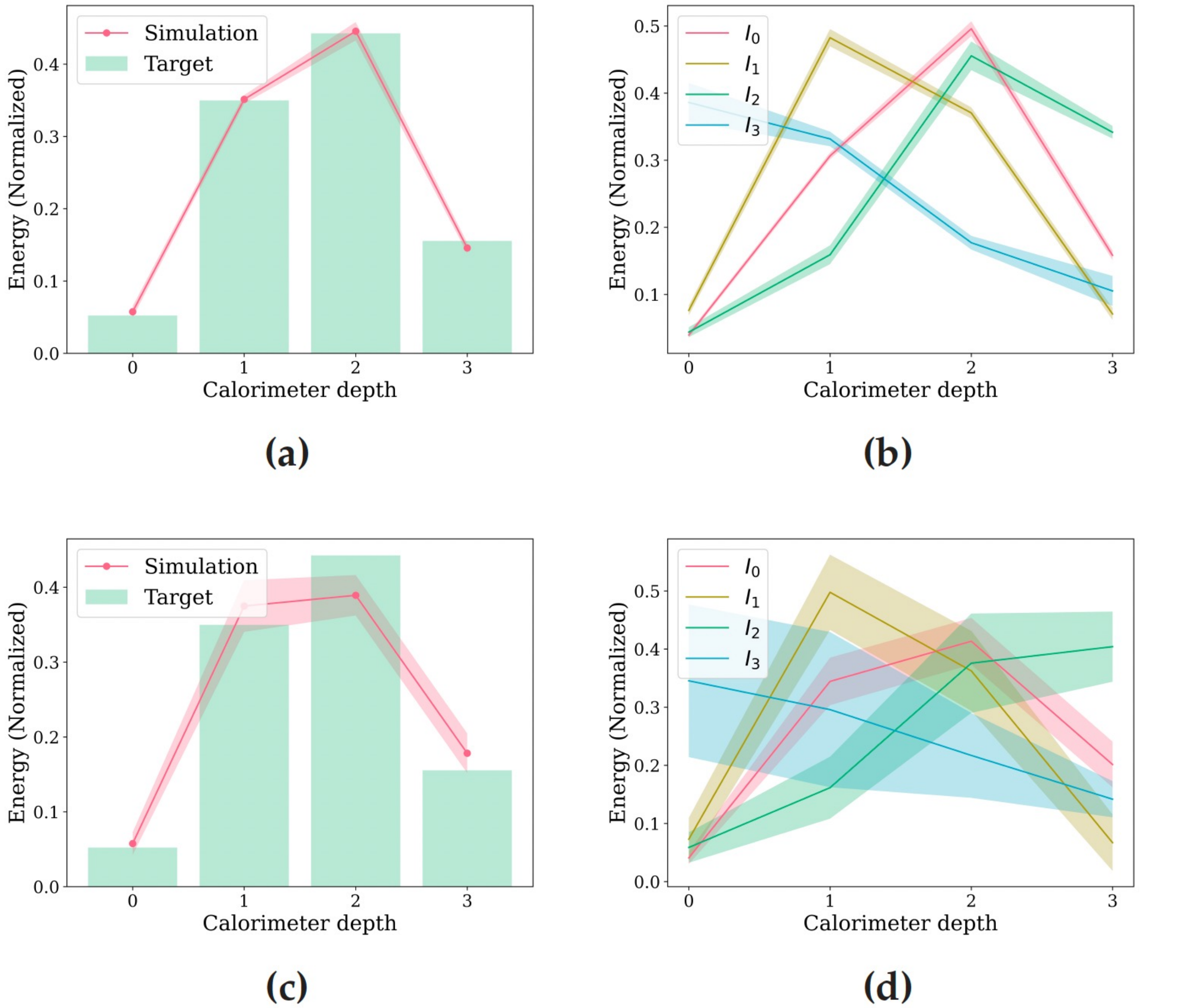


Figure 4: Mean (a,c) and individual images (b,d) obtained by inference test on ibmq_jakarta (a,b) and IONQ (c,d).

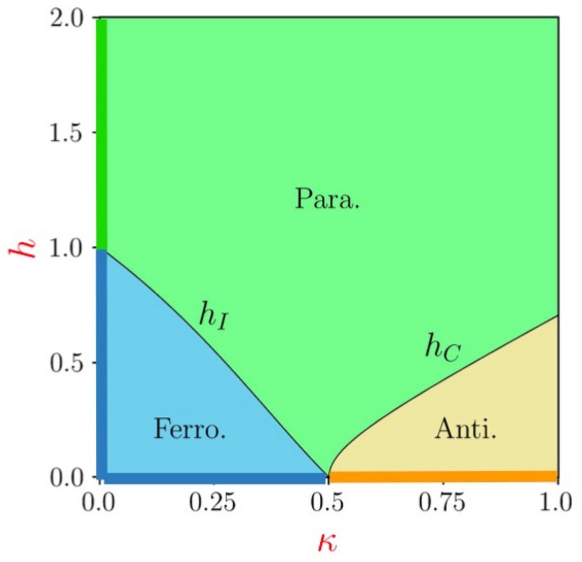
QML for quantum data: drawing phase diagrams

Model: Axial Next Nearest Neighbor Ising (ANNNI) Hamiltonian:

$$H = J\sum_{i=1}^{N} \sigma_x^i \sigma_x^{i+1} - \kappa \sigma_x^i \sigma_x^{i+2} + h \sigma_z^i,$$

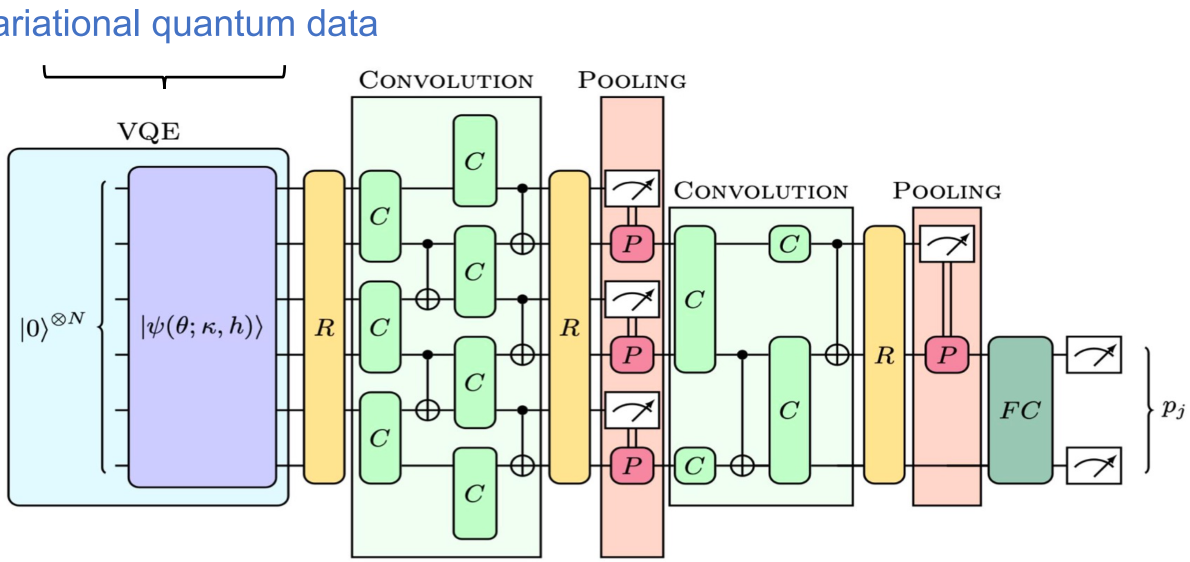
Senk, *Physics Reports*, **170**, 4 (1988)

Integrable for $\kappa = 0$ or $h = 0$.

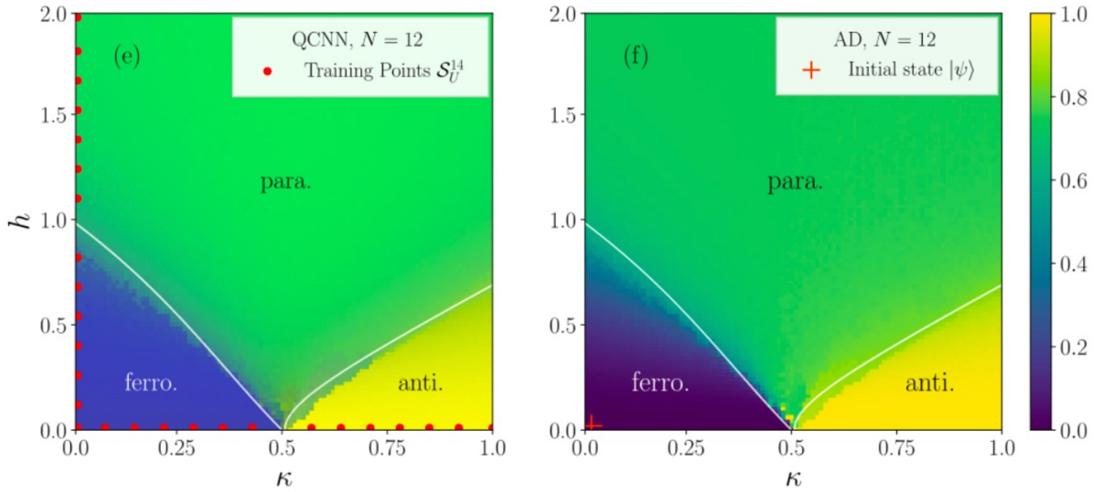


1. **Supervised classification of the ground state** using a convolutional QNN
2. Quantum states are **exponentially hard to save classically**.
3. **Bottleneck** from access to classical training labels (Interpolation does not work)
   - Train in integrable subregions
   - Generalize to a full model[1]

Setting the stage



QCNN (95%) / Autoencoder[1]



Binary Cross-entropy

Loss: $$\mathcal{L} = -\frac{1}{|\mathcal{S}_X^n|} \sum_{(\kappa,h)\in\mathcal{S}_X^n} \sum_{j=1}^{K} y_j(\kappa,h)\log\left(p_j(\kappa,h)\right)$$

Labels:

- [0,1] ferromagnetic
- [1,0] antiphase
- [1,1] paramagnetic
- [0,0] trash label

1. **Out of Distribution** Generalization[2]?
2. Performance increases with the system's size N=6 → N=12).
3. QCNN gives quantitative predictions

Monaco, at al. arXiv: 2208.08748 (2022), accepted PRB

[1]Kottman, *et al., Phys. Rev. Research* **3**, 043184 (2021)
[2]M..Caro et al., arxiv:2204.10268, Banchi et all., PRX QUANTUM 2, 040321 (2021)