Quantum Machine Learning in High Energy Physics

Examples from CERN

QUANTUM
TECHNOLOGY
INITIATIVE

Sofia.Vallecorsa@cern.ch

Quantum Observables For Colliders – November 2023

Outline

- The CERN Quantum Technology Initiative
- Quantum Machine Learning
  - Trainability and generalisation
- Example applications at CERN
  - Anomaly Detection
  - Beam Optimisation in linear accelerators
- Summary & Outlook

Quantum potential… and computer science

Principles of quantum mechanics enhance computations

Superposition leads to parallelism → **exponential speedup?**

Entanglement → **non linear correlation and classical intractability?**

Operations (gates) are unitary transformations → **reversible computing?**

Output is the result of a measurement according to Born rule → **stochastic computation ?**

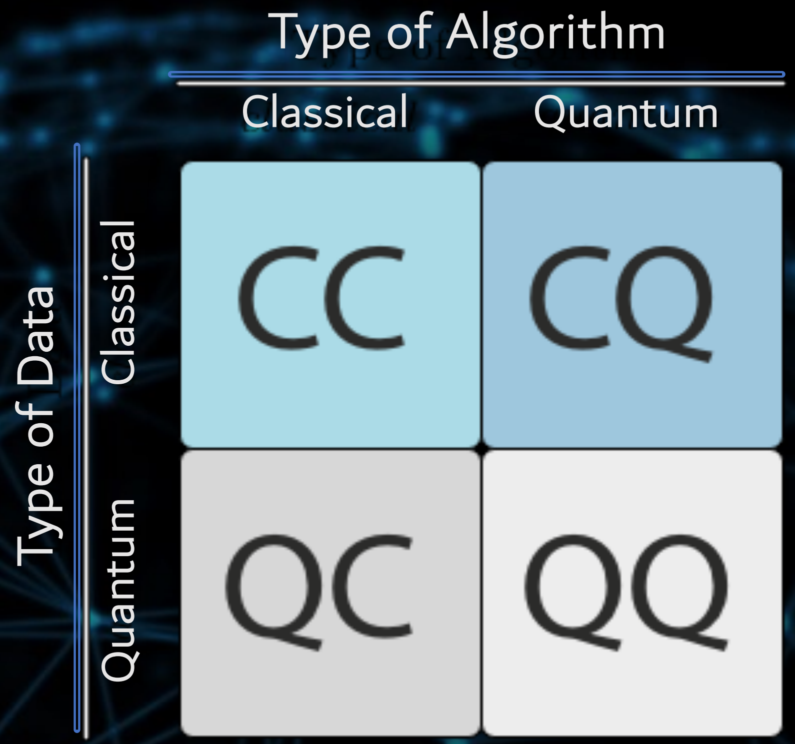
No-cloning theorem → **information security**

Quantum state coherence and isolation → **computation stability and errors**

Qubit state collapses → **reproducibility?**

QML: Quantum computing to “improve” ML

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



Type of Data \ Type of Algorithm	Classical	Quantum
Classical	CC	CQ
Quantum	QC	QQ

- Evaluate performance on realistic use cases

- QPU as accelerators within classical infrastructure?

The CERN Quantum Technology Initiative was launched in 2020

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)

Understanding the impact of quantum technologies in HEP

Quantum simulation and HEP theory applications

Quantum Computing

Quantum Sensing

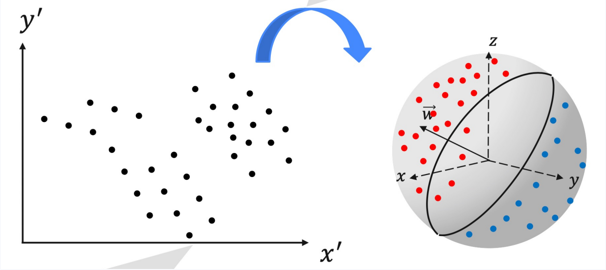
Quantum Communication

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

Quantum Machine Learning :

Some basic concepts

(Quantum) ML Lifecycle



Data Embedding

Data Preparation

Model Definition

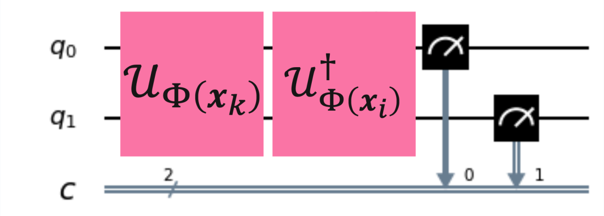
Adapt classical learning models to quantum space

Model Training

Model Testing

Model Interpretation

Readout and measurement "shots"



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

Quantum embedding for classical data

Compromise between **exponential compression and circuit depth**

Ex: **Amplitude Encoding**

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

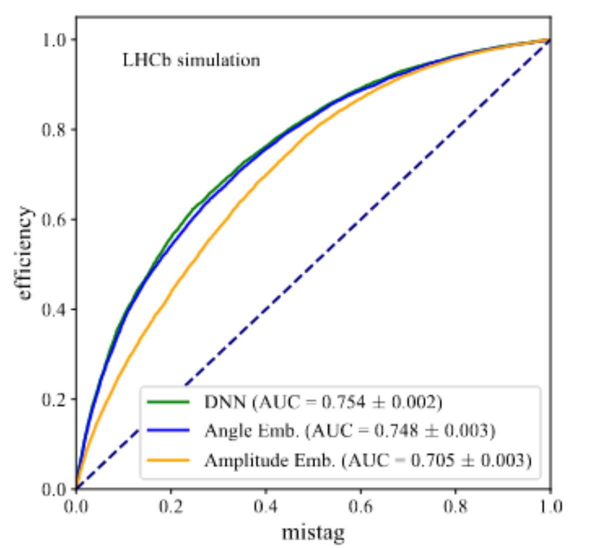
+ Exponential compression

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

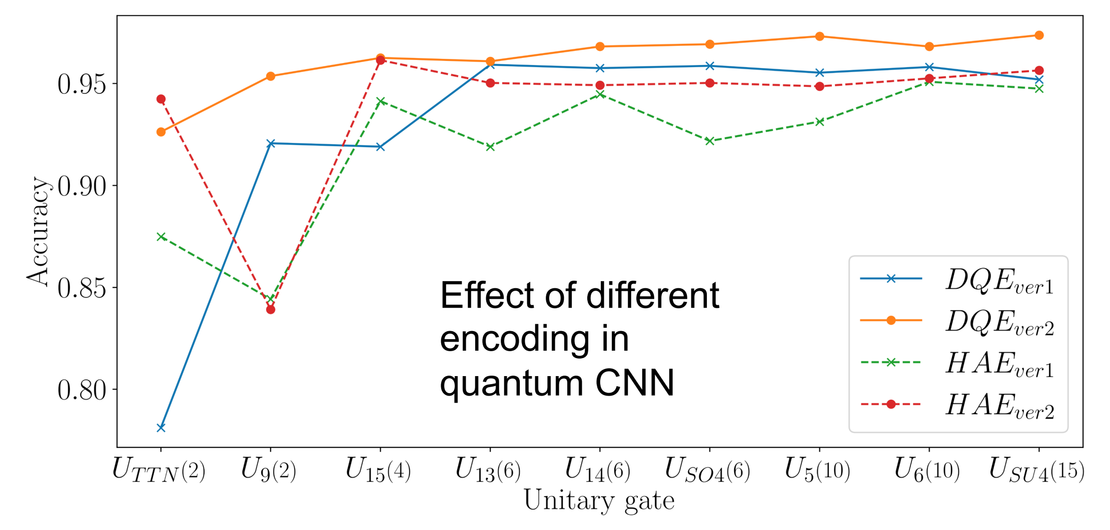
− Polynomial number of gates

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

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



Effect of different encoding in quantum CNN

Models

Variational algorithms (ex. QNN)

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

Data Embedding can be **learned**

Ansatz design can leverage data symmetries[1]

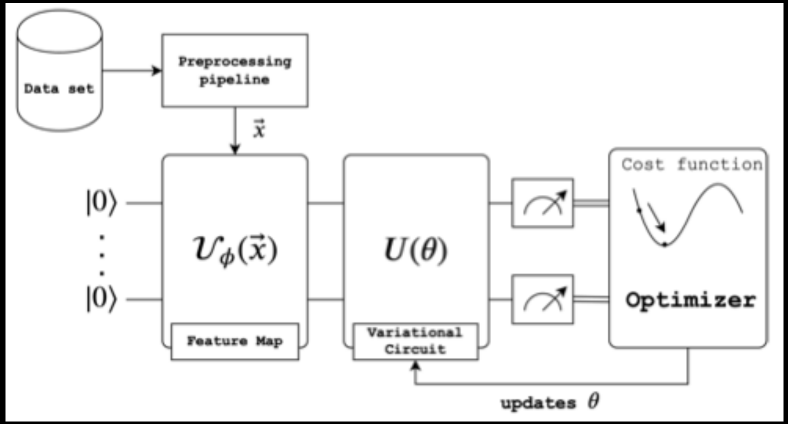


Image credit SwissQuantumHub

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]

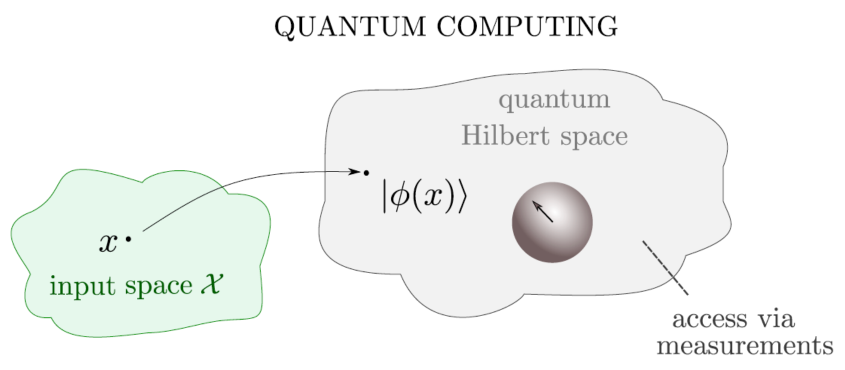


Image credit M. Schuld

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).

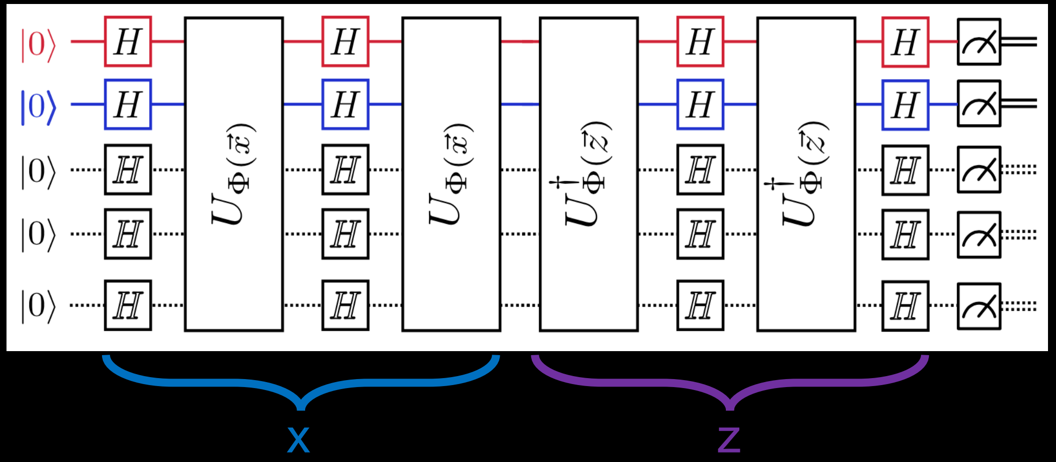
QML Convergence

Classical Intractability & expressivity vs trainability and generalization

F. Di Marcantonio *et al.* , CHEP2023

Quantum embedding and kernel methods

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



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

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

Hilbert space is exponentially larger

↓

Sparser data

↓

Loss of predictive power

Kernel trainability and kernel concentration

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

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

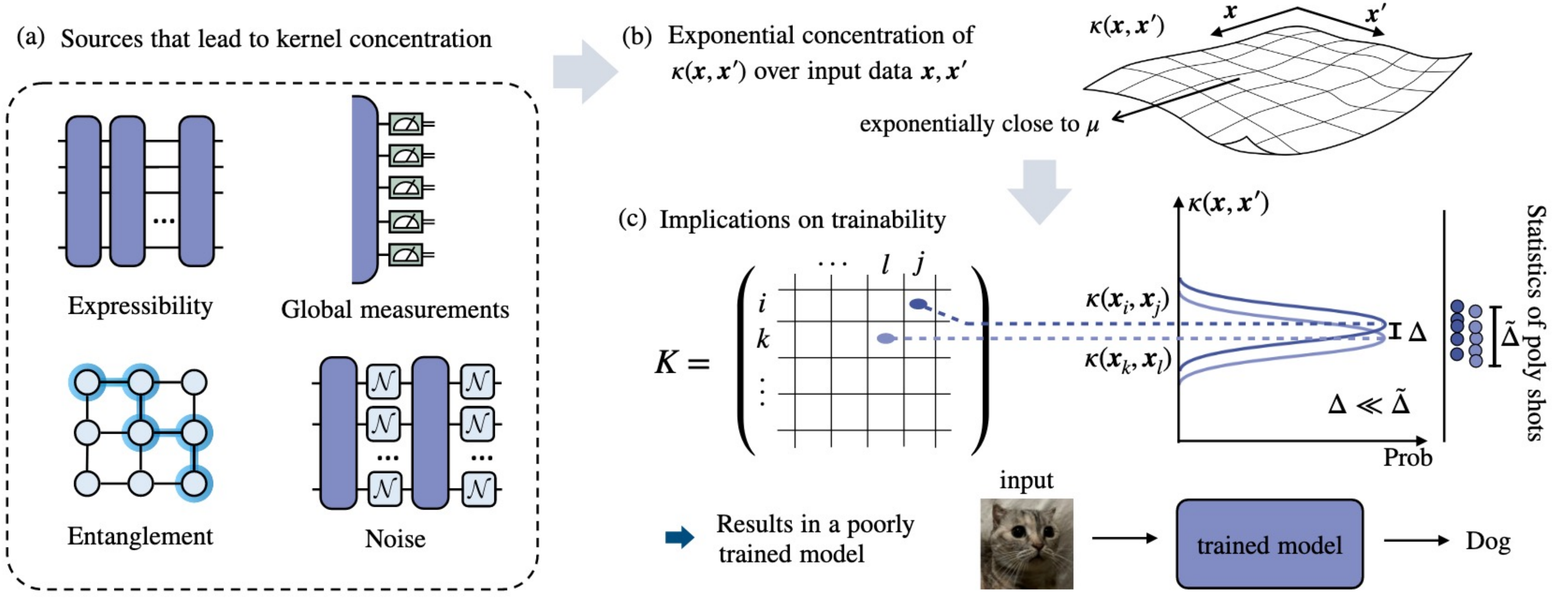


Figure 1. **Kernel concentration and its implications on trainability:** The exponential concentration (in the number of qubits n) of quantum kernels $\kappa(\boldsymbol{x}, \boldsymbol{x}')$, over all possible input data pairs $\boldsymbol{x}, \boldsymbol{x}'$, can be seen to stem from the difficulty 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)

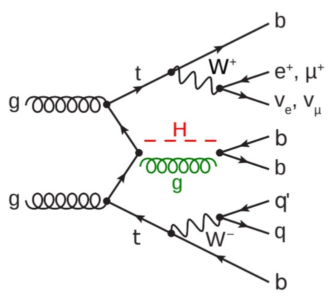
Projected Quantum Kernel

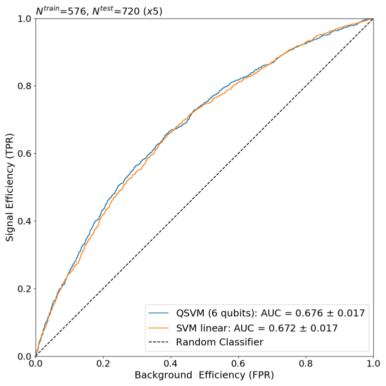
$$k^{lp}(x_i, x_j) = \sum_{k=1}^{m} \frac{Tr[\rho_k(x_i)\rho_k(x_j)]}{m}$$

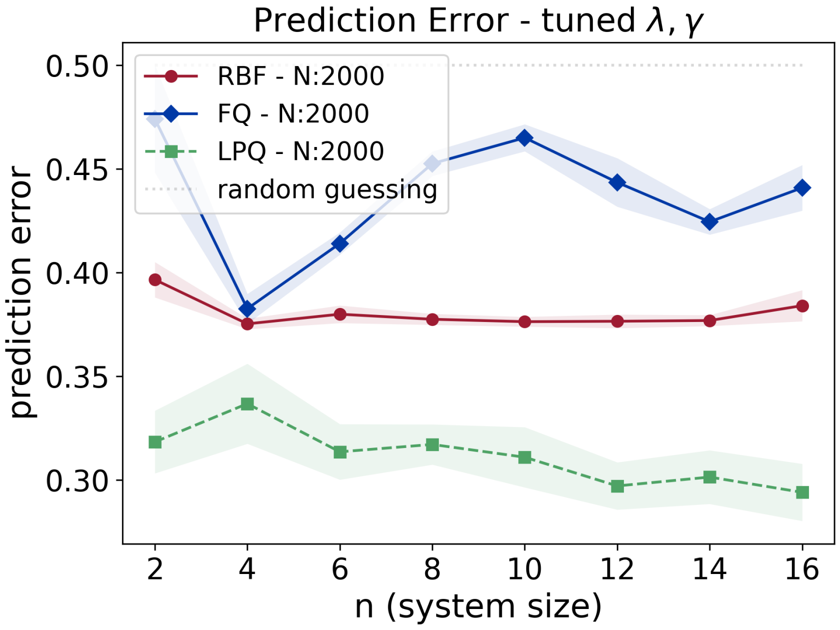
Project quantum kernels to lower dimensionality (i.e. local density matrix)[1]:

- Improved generalizion while keeping features into states classically hard

- Example: ttH(bb) binary classification[2]2







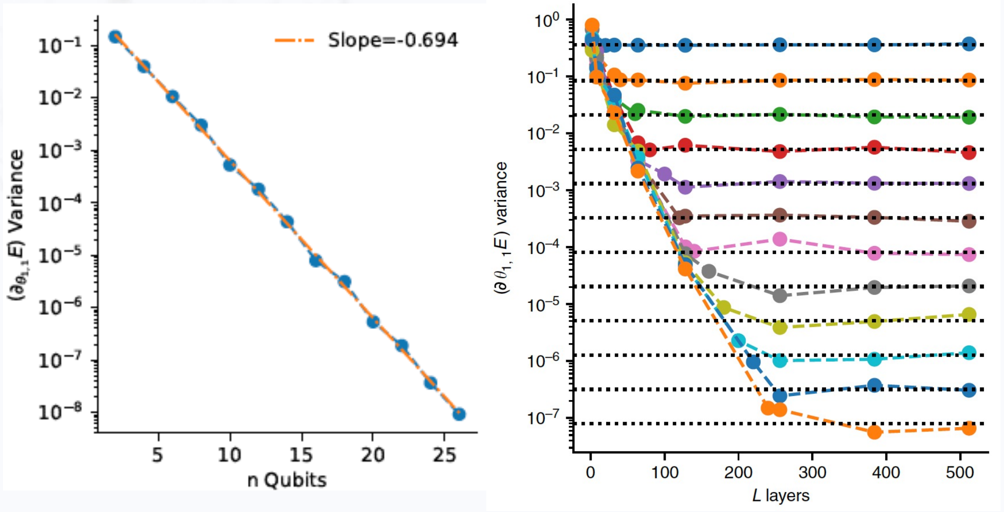
---

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

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

Gradients decay and Model Convergence

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



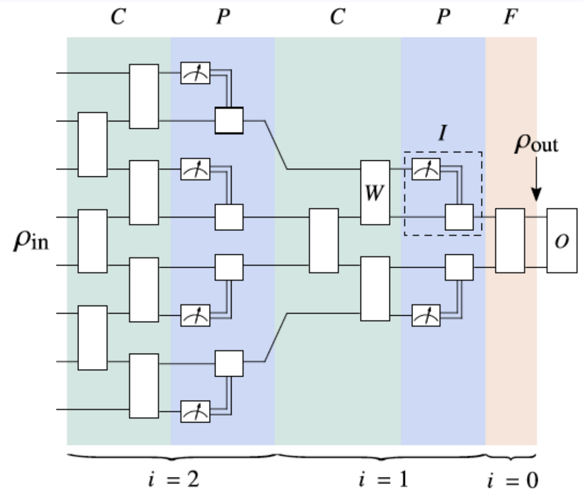
Classical gradients **vanish exponentially** with the number of layers (J.McClean *et al.*, arXiv:1803.11173)

- Convergence still possible if gradients consistent between batches.

Quantum gradient decay exponentially in the number of qubits (number of graph paths is exponential in the number of gates)

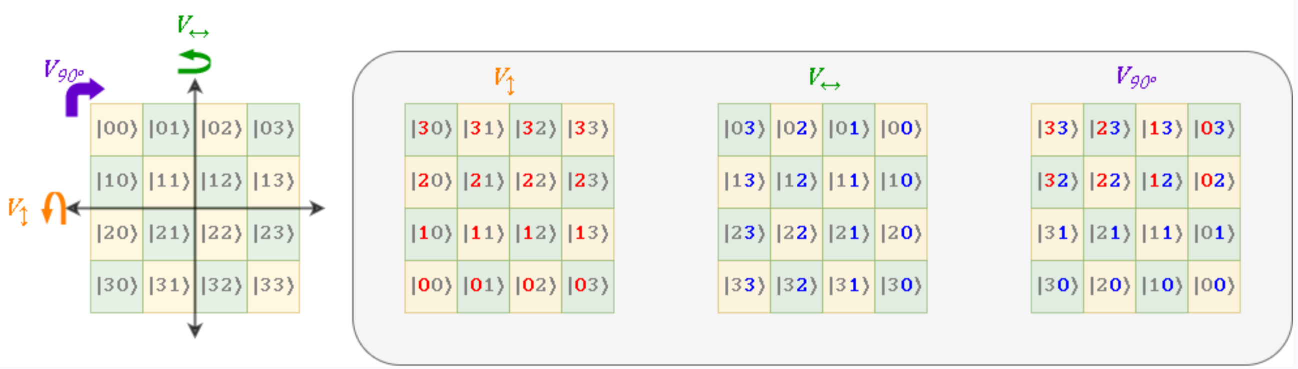
- Random circuit initialization
- Loss function locality in shallow circuits (M. Cerezo *et al.*, arXiv:2001.00550)
- Ansatz choice: TTN, CNN (Zhang *et al.,* arXiv:2011.06258, A Pesah, *et al.*, *Physical Review X* 11.4 (2021): 041011.)
- Noise induced barren plateau (Wang, S *et al.*, Nat Commun 12,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

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



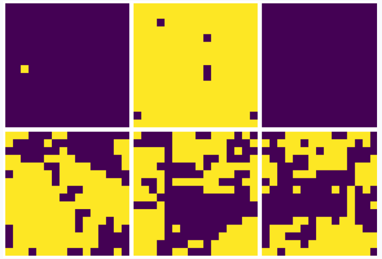
Equivariant Quantum CNN

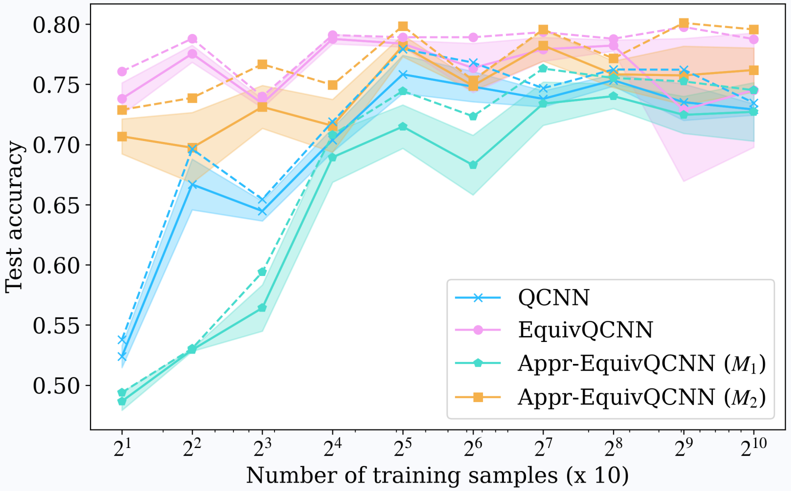
- Construct **equivariant** quantum CNN **under rotational & reflectional symmetry**
- Improved generalization power



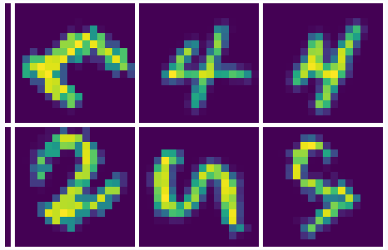
$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

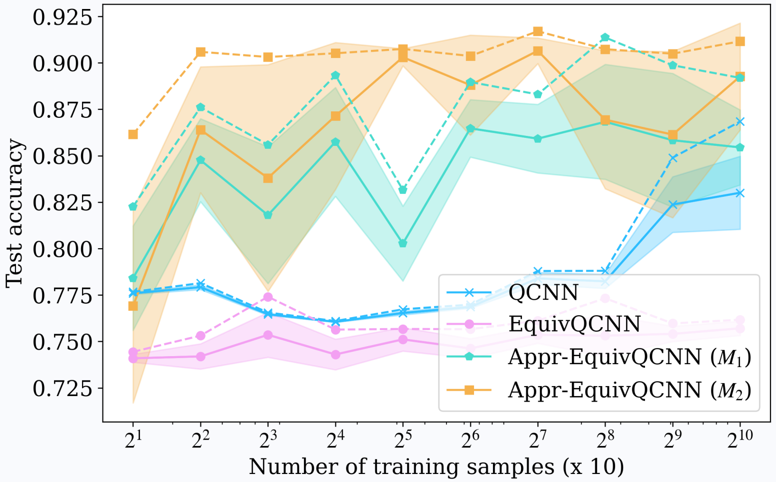
Ising spins phase classification :





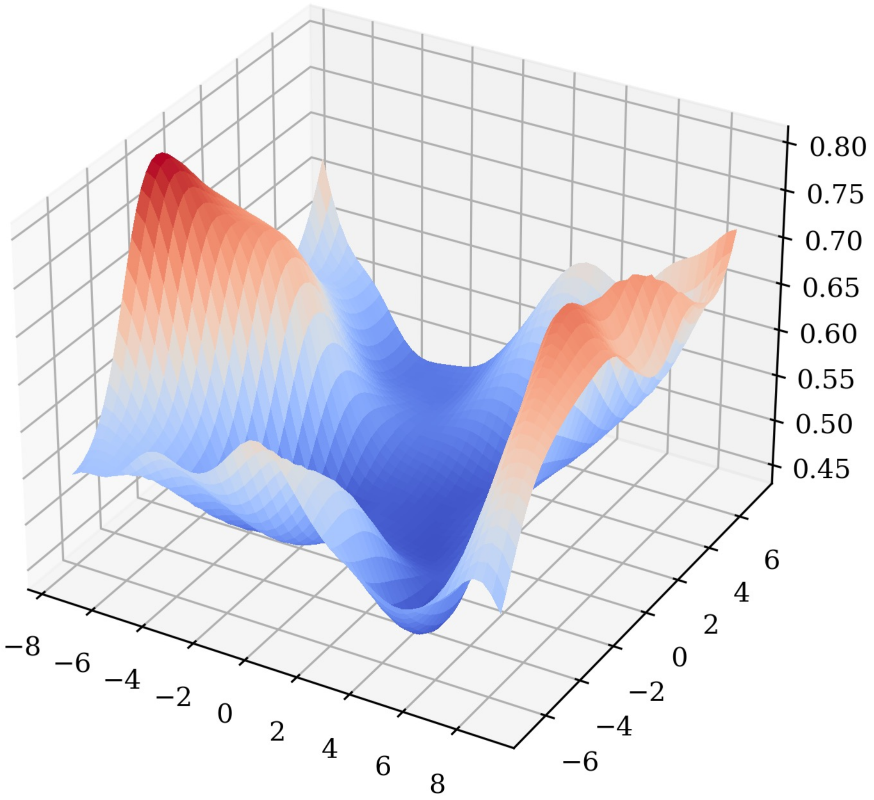
Extended MNIST
Image classification:
(digits 4,5)



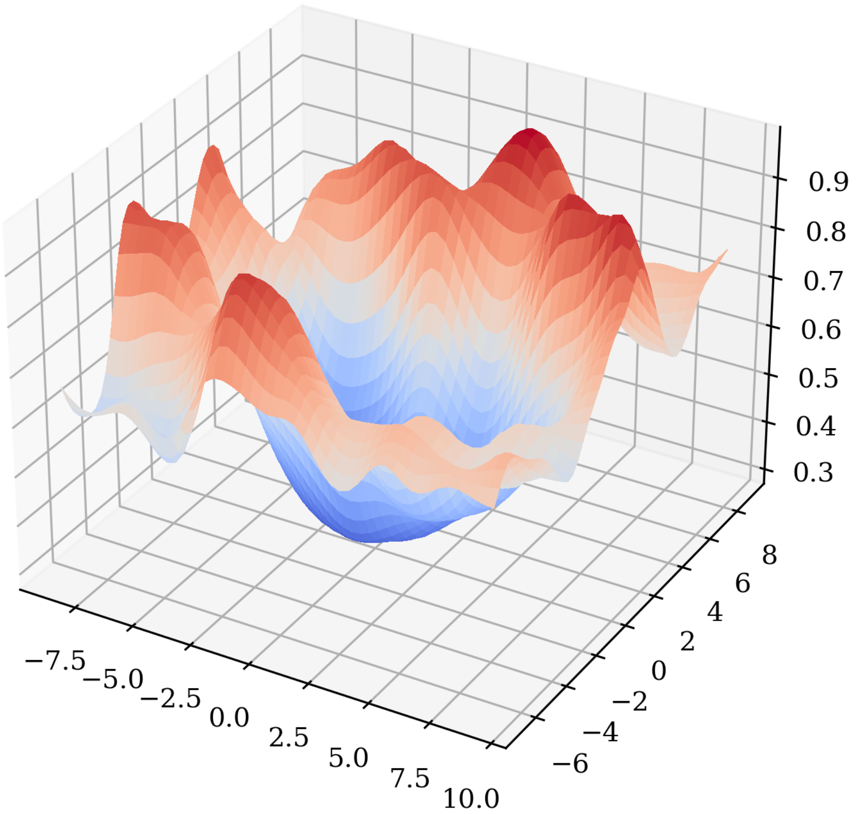


Non-convexity of loss landscape

- **Loss landscape plotted with orqviz**



Non-equivariant QCNN



ApprEquivQCNN

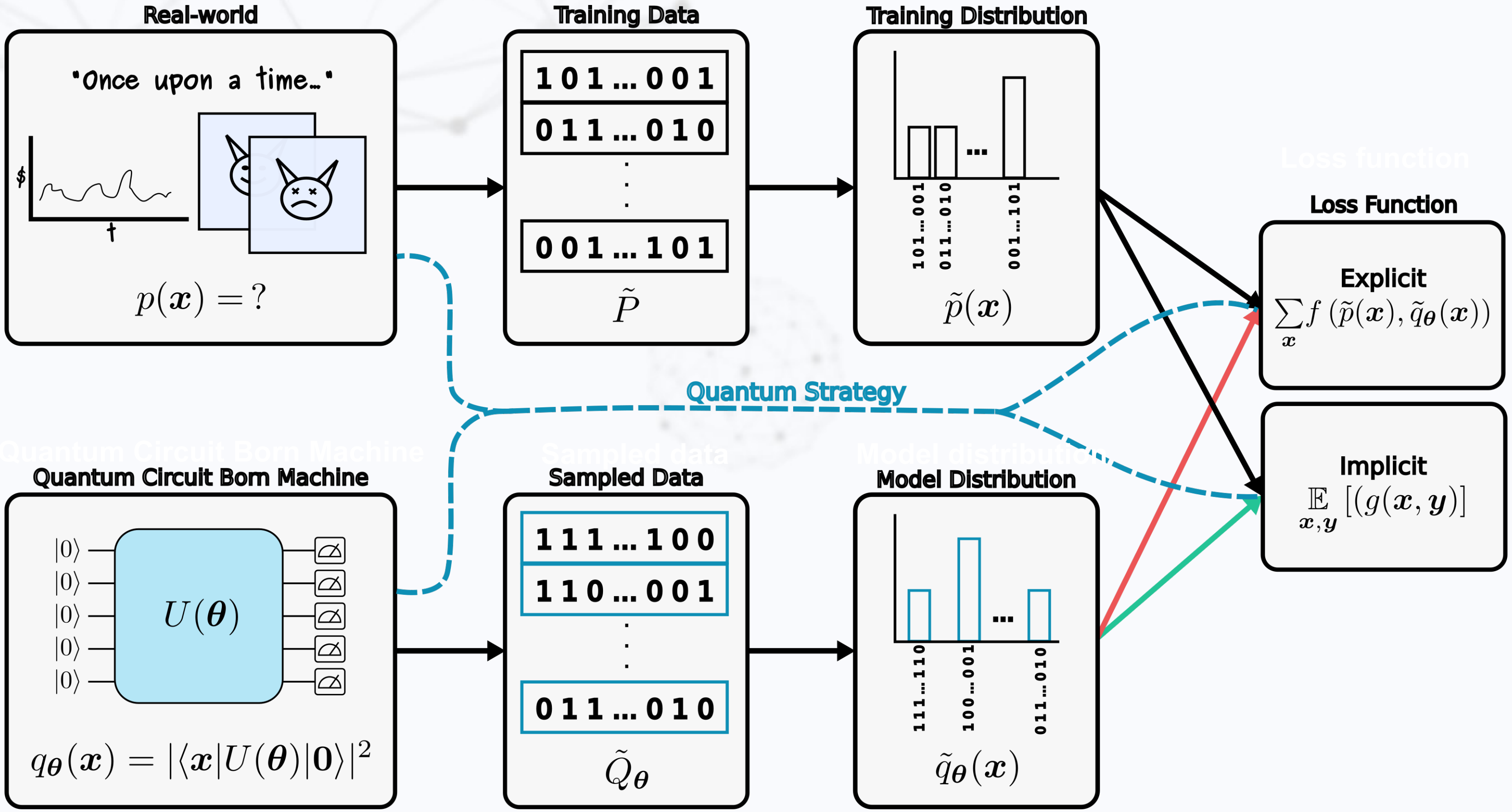
Quantum Machine Learning examples:

Generative Models

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



Real-world

'Once upon a time…'

$p(\boldsymbol{x}) = ?$

Training Data

101…001

011…010

⋮

001…101

$\tilde{P}$

Training Distribution

$\tilde{p}(\boldsymbol{x})$

Loss Function

Explicit

$\sum_{\boldsymbol{x}} f\left(\tilde{p}(\boldsymbol{x}), \tilde{q}_{\boldsymbol{\theta}}(\boldsymbol{x})\right)$

Implicit

$\underset{\boldsymbol{x},\boldsymbol{y}}{\mathbb{E}}\left[(g(\boldsymbol{x},\boldsymbol{y})\right]$

Quantum Strategy

Quantum Circuit Born Machine

$q_{\boldsymbol{\theta}}(\boldsymbol{x}) = |\langle \boldsymbol{x}|U(\boldsymbol{\theta})|\boldsymbol{0}\rangle|^2$

Sampled Data

111…100

110…001

⋮

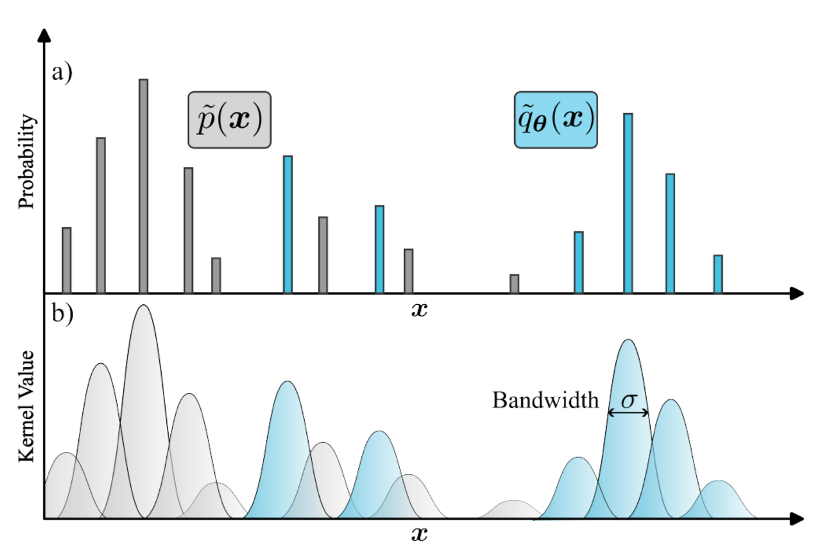
011…010

$\tilde{Q}_{\boldsymbol{\theta}}$

Model Distribution

$\tilde{q}_{\boldsymbol{\theta}}(\boldsymbol{x})$

exponentially larger number of shots is required to keep accuracy of explicit losses

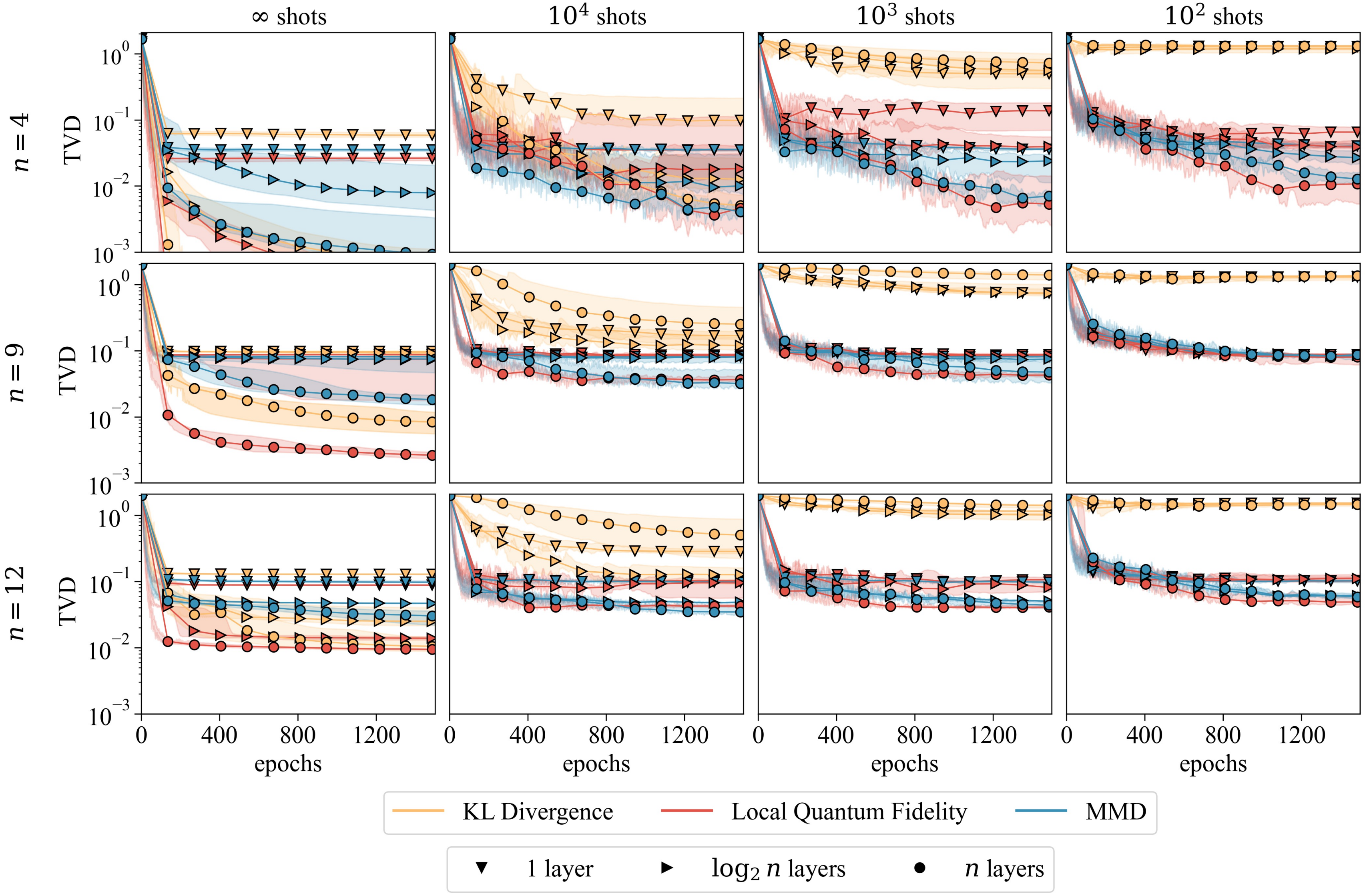


Rudolph, M. S., Lerch, S., Thanasilp, S., Kiss, O., Vallecorsa, S., Grossi, M., & Holmes, Z. (2023). **Trainability barriers and opportunities in quantum generative modeling.** *arXiv:2305.02881*.

Quantum Circuit Born Machine for HEP

QCBM

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

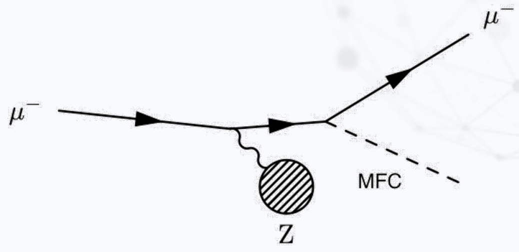


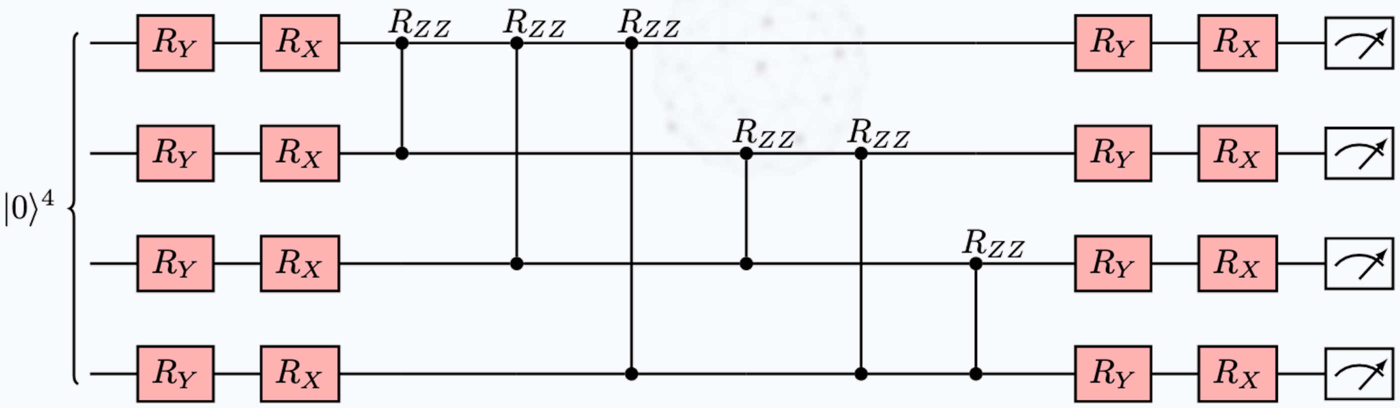
QCBM for event generation

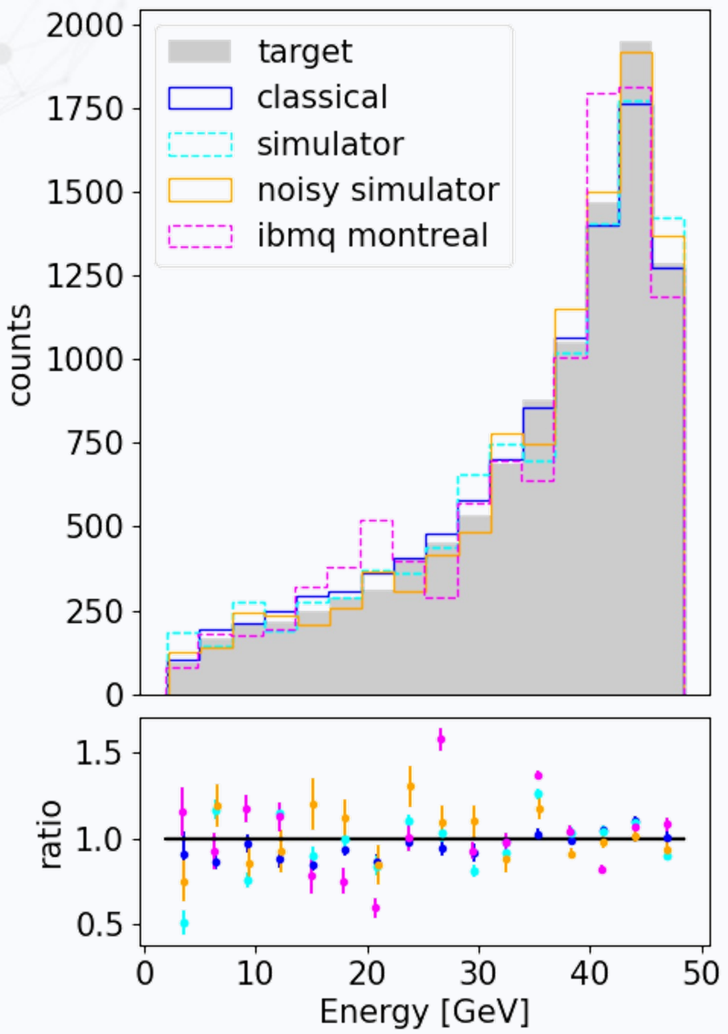
Muon Force Carriers, in muon fixed-target experiments (FASER) or muon interactions in calorimeters (ATLAS)[1].

Generate multivariate distribution (E, p_t, η) using a QCBM

Maximum Mean Discrepancy for training





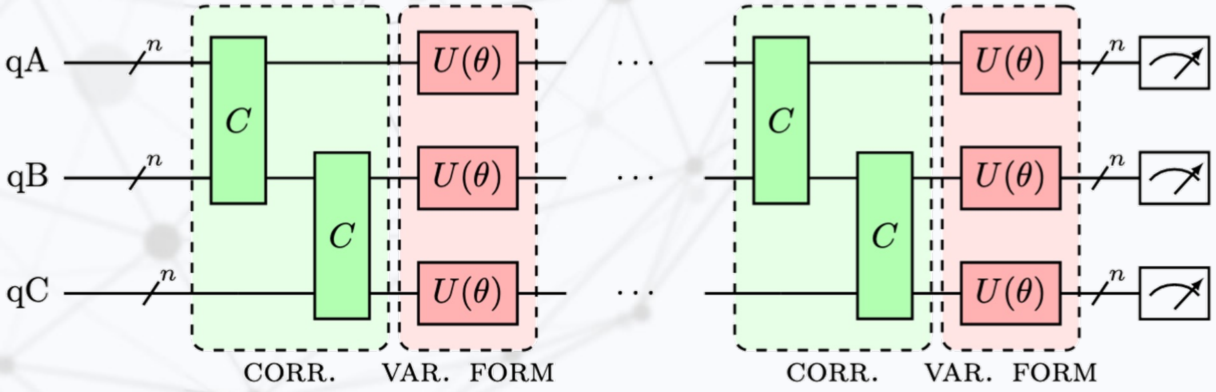


---

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

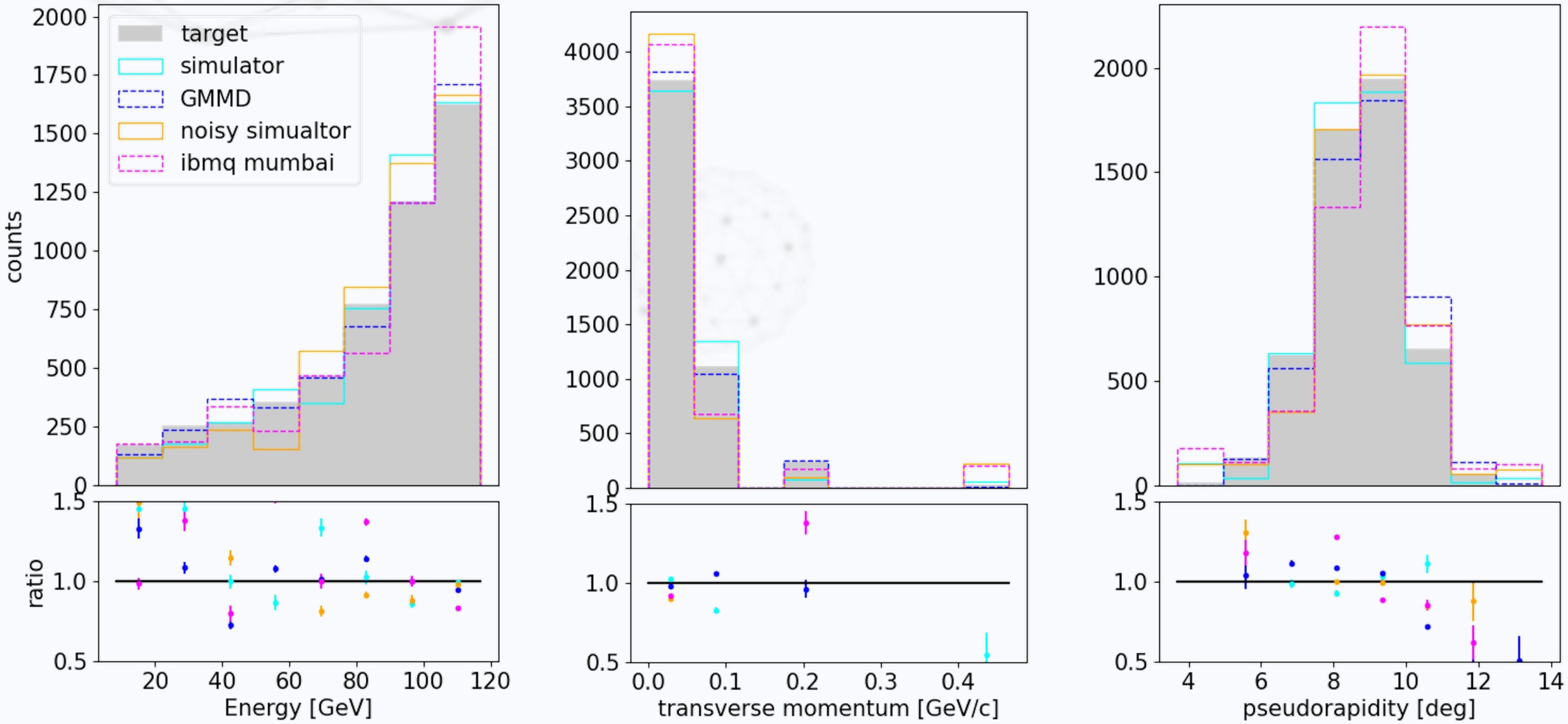
Kiss, Grossi, *et al.*, *Phys. Rev. A* **106**, 022612 (2022)

Multivariate PDFs

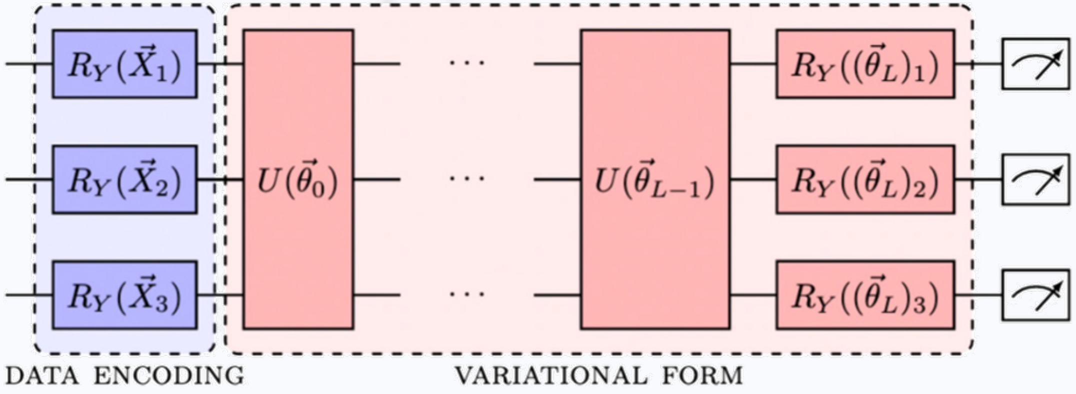


Simulation	Noisy simulation	IBMQ Mumbai	Classical
0.12	0.06	0.06	0.01

Mean difference between the correlations in the MC and generated samples

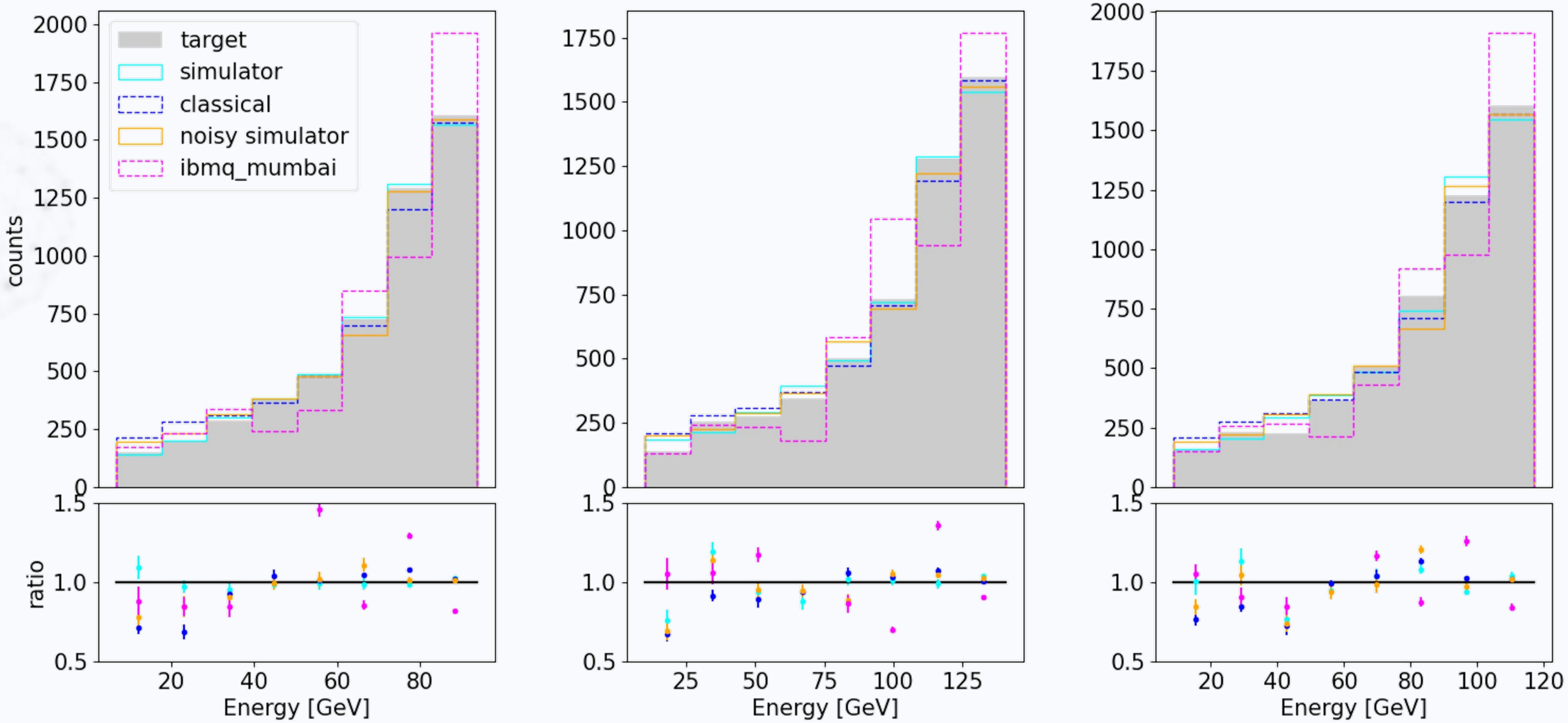


Conditional probability distribution



We want to **modelize** $\boldsymbol{p(y|x)}$ where x is the incoming energy E_{in}.

1. Data re-uploading does not improve the sampling.
2. Training on hardware is important to assimilate the noise.



a) train: 100 GeV

b) train 150 GeV

c) test 125 GeV

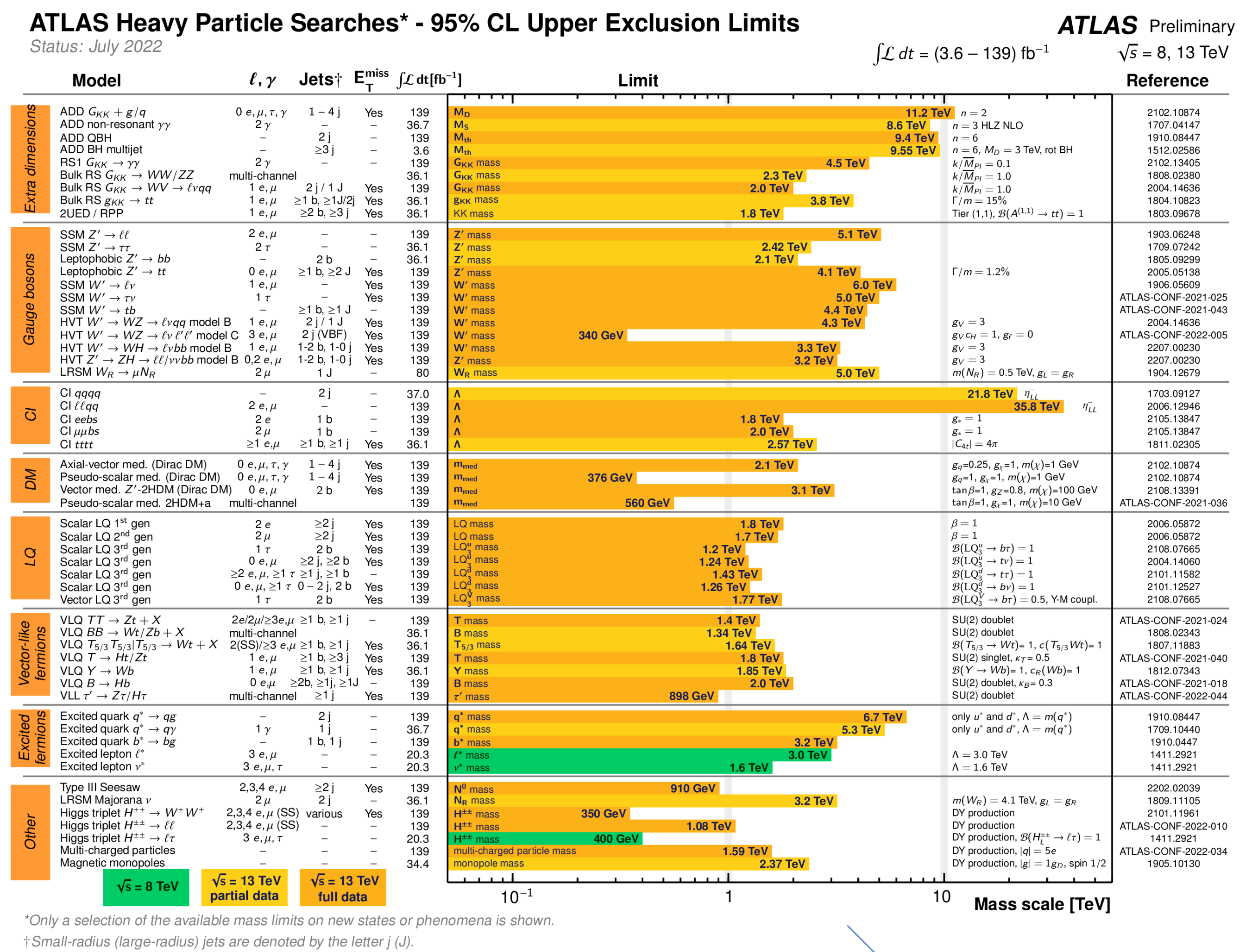
Quantum Machine Learning examples:

Anomaly Detection

Quantum anomaly detection in the latent space of proton collision events at the LHC
arXiv:2301.10780.

New Physics at the LHC

So far only **negative results** in **direct** (model dependent) searches

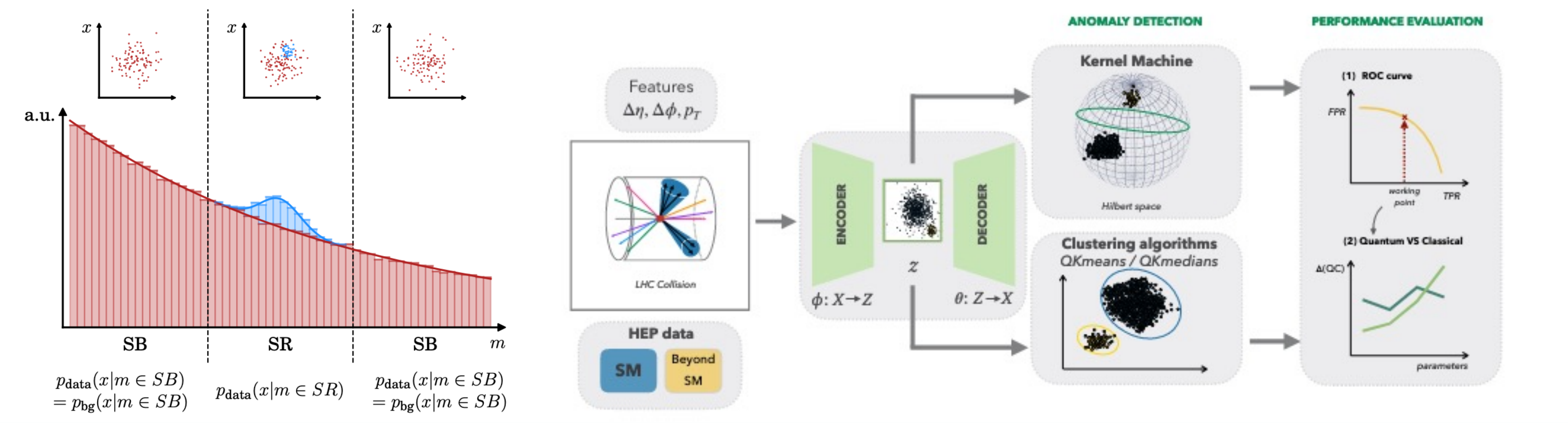


How to insure we do not miss potential discoveries?

We can design model agnostic searches!

Unsupervised learning for Anomaly Detection

A typical hybrid QML workflow



Standard Model jets

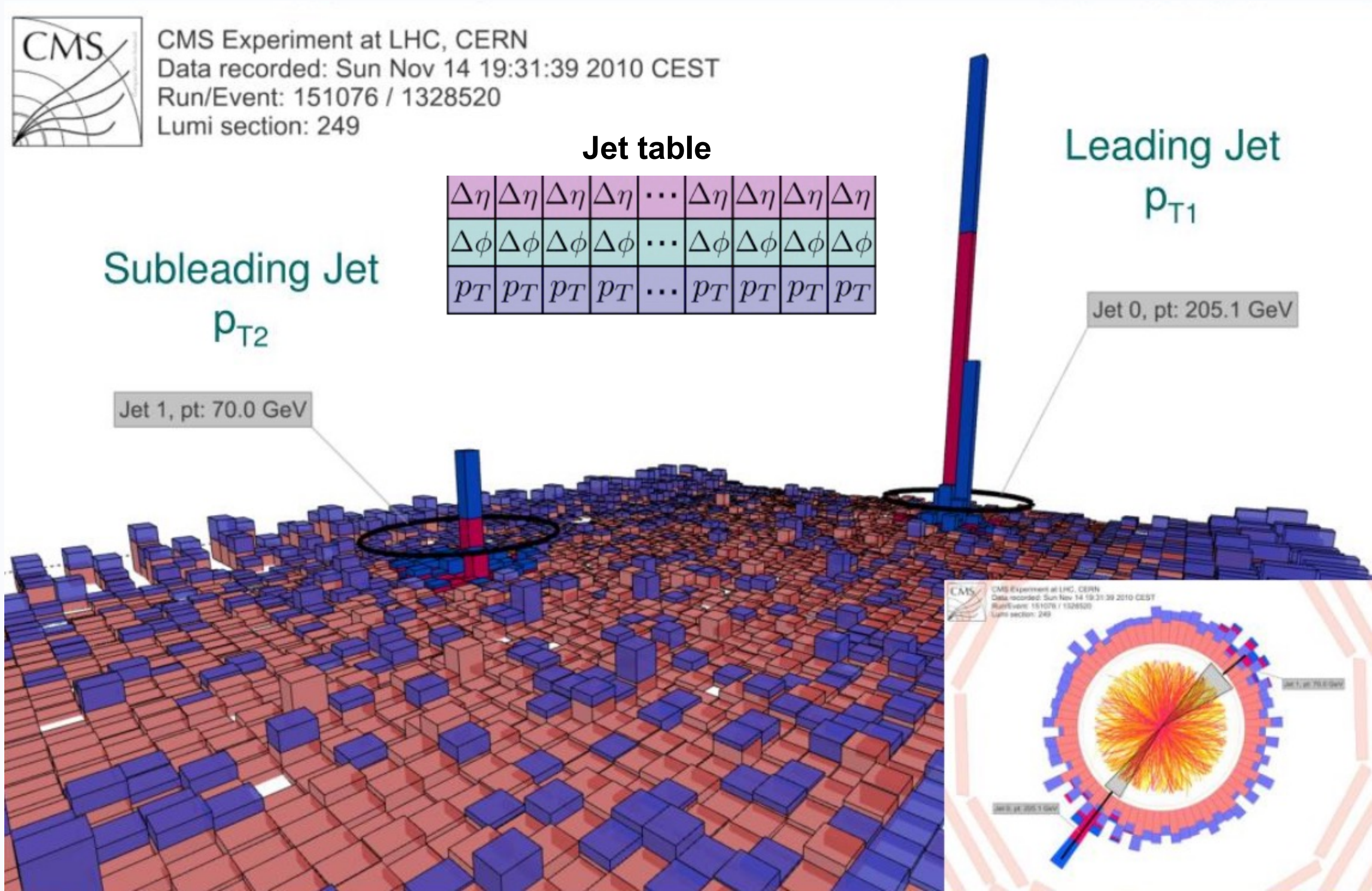
Simulate QCD multi-jets at the LHC

Build jet from 100 highest pt particles

Apply realistic event selection

Convolutional AutoEncoder
learns the jet **internal structure**

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



Jet table

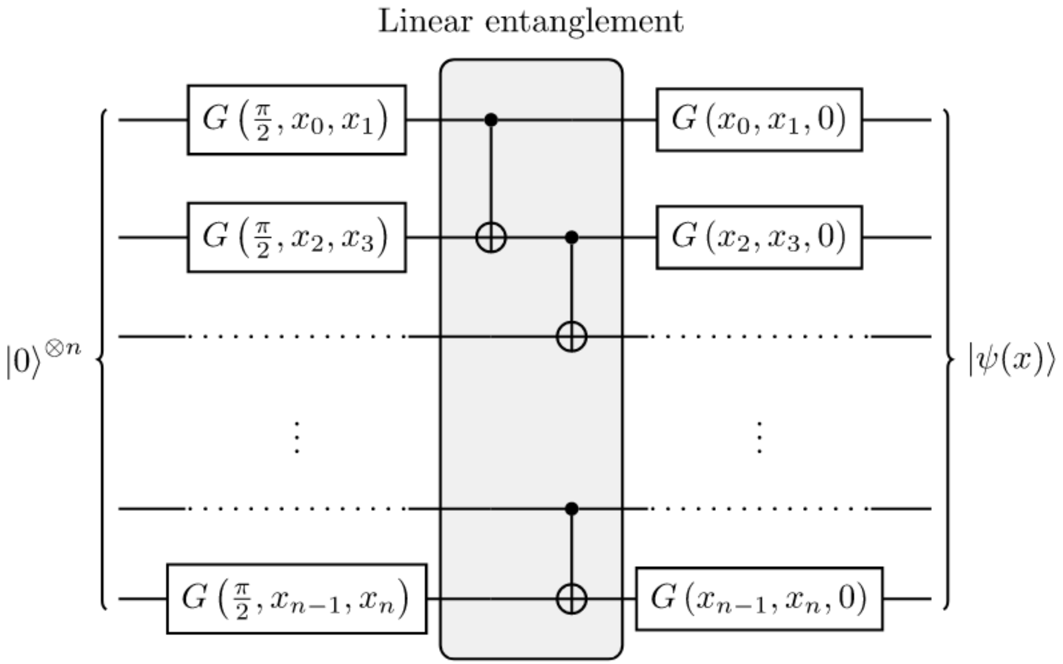
$\Delta\eta$	$\Delta\eta$	$\Delta\eta$	$\Delta\eta$	$\cdots$	$\Delta\eta$	$\Delta\eta$	$\Delta\eta$	$\Delta\eta$
$\Delta\phi$	$\Delta\phi$	$\Delta\phi$	$\Delta\phi$	$\cdots$	$\Delta\phi$	$\Delta\phi$	$\Delta\phi$	$\Delta\phi$
p_T	p_T	p_T	p_T	$\cdots$	p_T	p_T	p_T	p_T

Unsupervised kernel machine

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

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

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

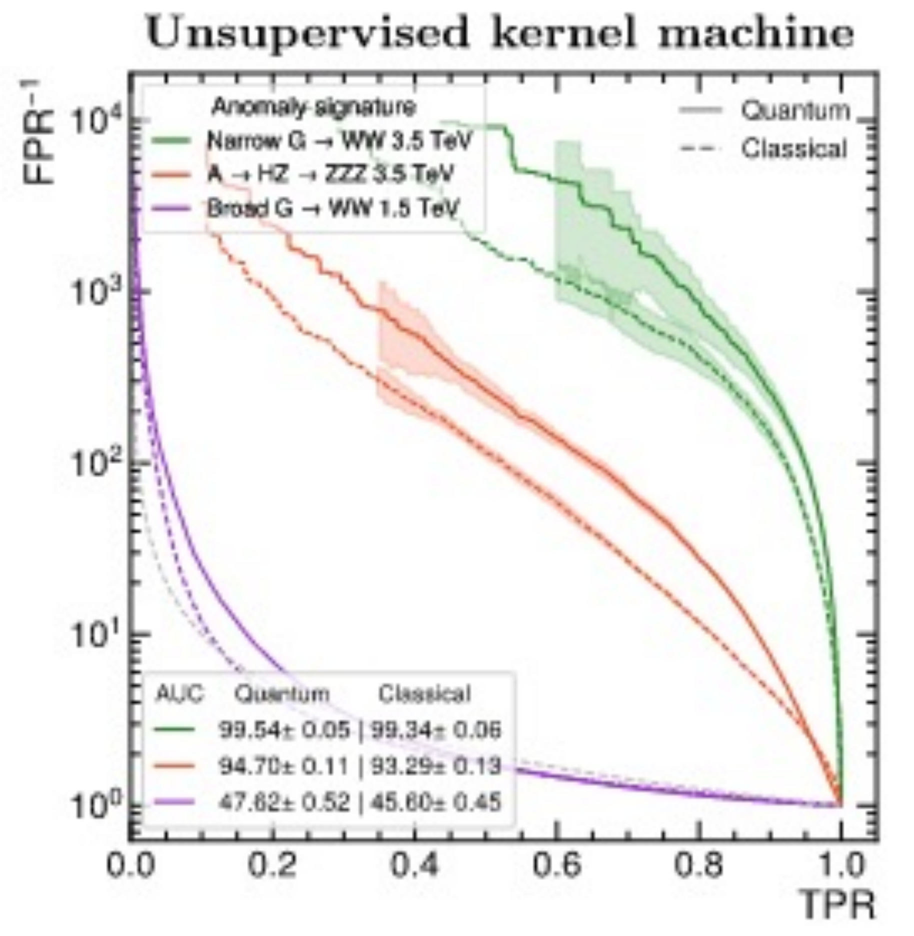


$$\min_{w\in\mathcal{F},\,\xi\in\mathbb{R}^\ell,\,\rho\in\mathbb{R}} \quad \frac{1}{2}||w||^2 + \frac{1}{\nu\ell}\sum_i \xi_i - \rho$$

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

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

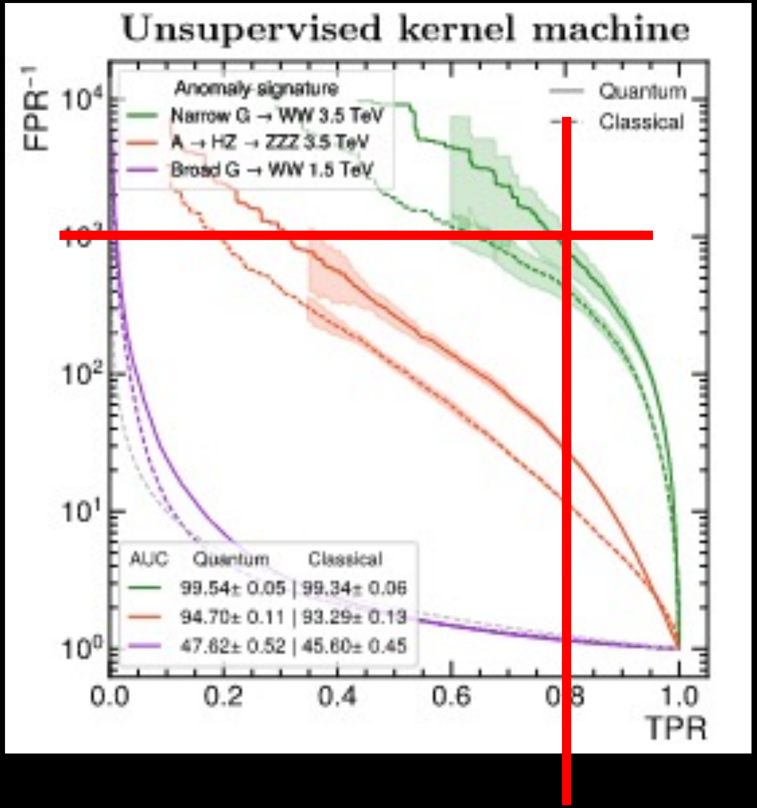
Results



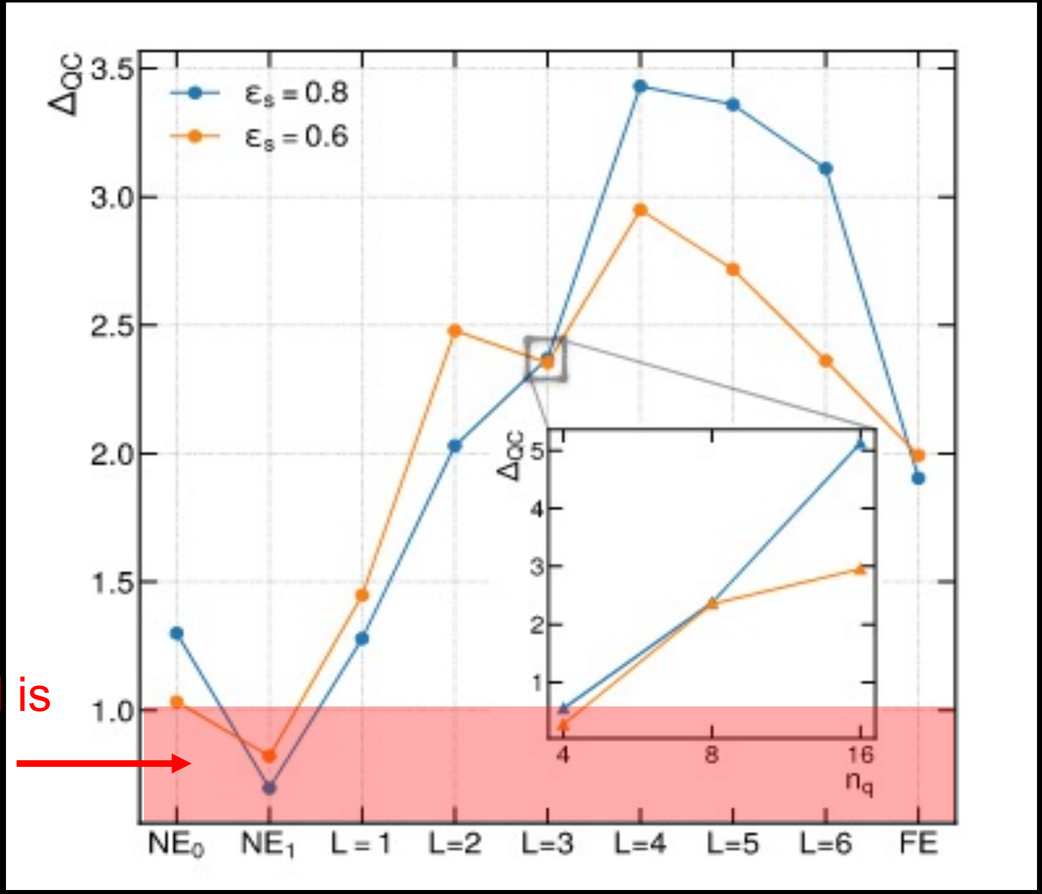
Is this an «advantage» we can use?

Quantum anomaly detection in the latent space of proton collision events at the LHC
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
arXiv:2301.10780.

Quantum Machine Learning examples:

Phase Transitions identification

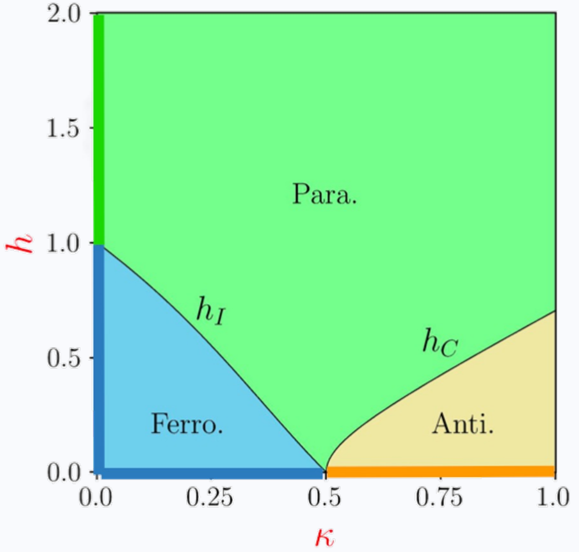
QML for quantum data: drawing phase diagrams

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)

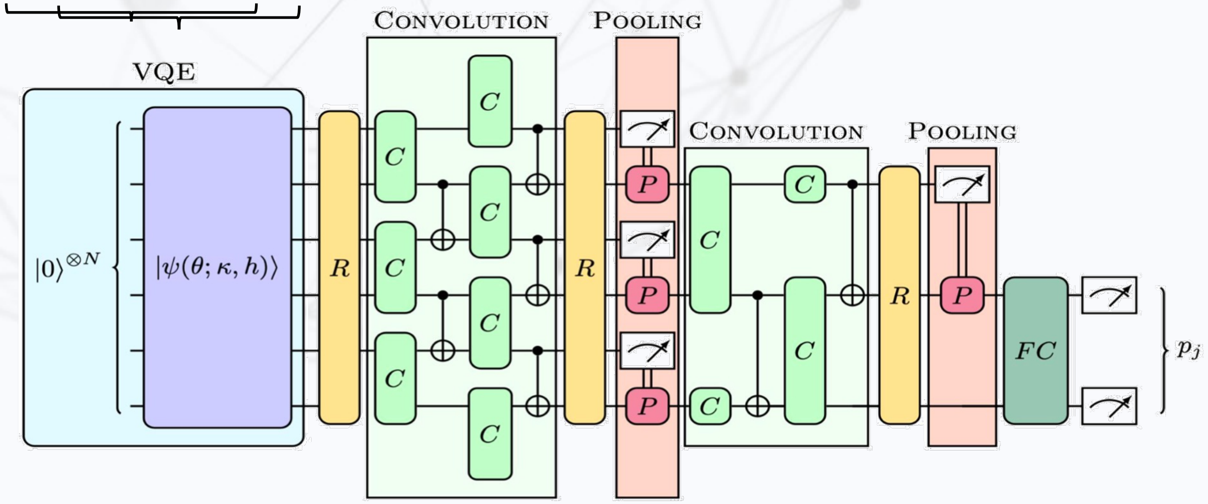
Integrable for $\kappa = 0$ or $h = 0$.



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]

Results

Variational quantum data



Binary Cross-entropy

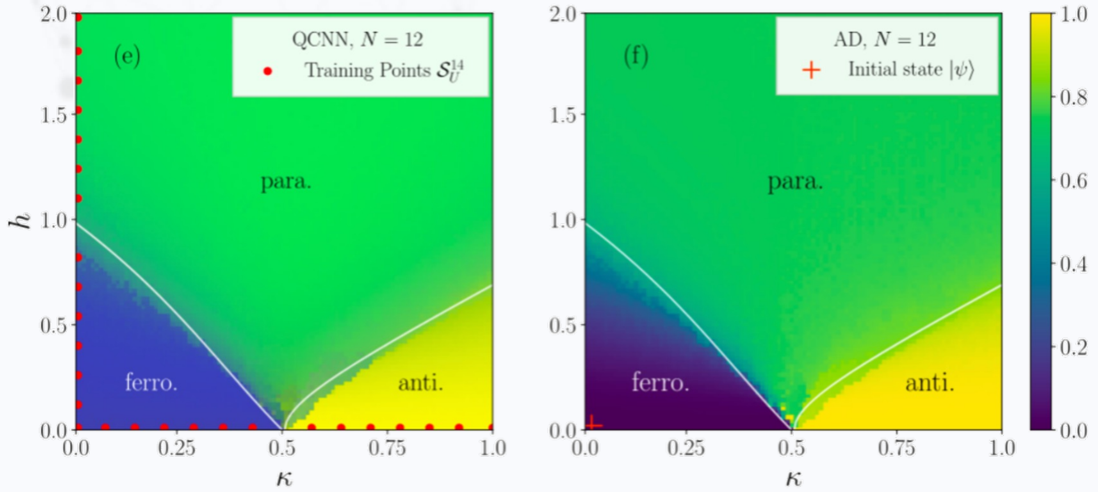
Loss: $$\mathcal{L} = -\frac{1}{|\mathcal{S}_X^n|}\sum_{(\kappa,h)\in\mathcal{S}_X^n}\sum_{j=1}^{K} y_j(\kappa,h)\log\left(p_j(\kappa,h)\right)$$

Labels:

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

QCNN (95%)

Autoencoder[1]



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

[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)

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

Quantum Machine Learning examples:

QNN for Quntum Monte Carlo Integration

Monte Carlo integration

Monte Carlo Integration is widely used in multiple applications.

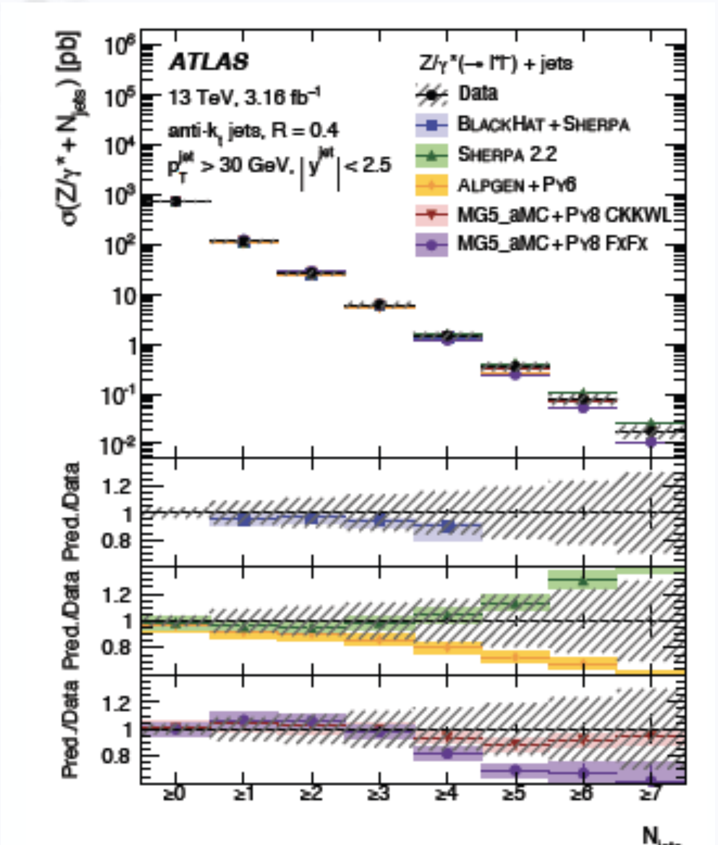
Computationally challenging

In HEP:

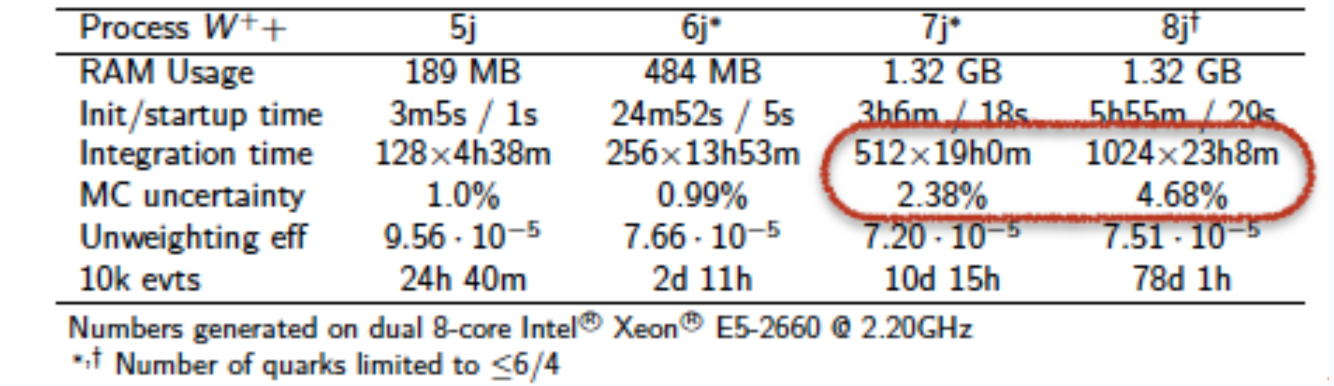
Phase space sampling scales exponentially with number of final state particles[1]

- HL-LHC @ 3 10^{-3} fb^{-1} will have percent-level precision @N_{jet} = 9
- Need comparable (higher-order) MC
- N_{jet} increases with center-of-mass energy

Agliardi, Gabriele, et al. "**Quantum integration of elementary particle processes**." *arXiv preprint arXiv:2201.01547* (2022)



Time and memory usage (Sherpa 3.x.y + HDF5) (H. Schulz 2018)



Process $W^+ +$	5j	6j*	7j*	8j†
RAM Usage	189 MB	484 MB	1.32 GB	1.32 GB
Init/startup time	3m5s / 1s	24m52s / 5s	3h6m / 18s	5h55m / 20s
Integration time	128×4h38m	256×13h53m	512×19h0m	1024×23h8m
MC uncertainty	1.0%	0.99%	2.38%	4.68%
Unweighting eff	$9.56 \cdot 10^{-5}$	$7.66 \cdot 10^{-5}$	$7.20 \cdot 10^{-5}$	$7.51 \cdot 10^{-5}$
10k evts	24h 40m	2d 11h	10d 15h	78d 1h

Numbers generated on dual 8-core Intel® Xeon® E5-2660 @ 2.20GHz

*,† Number of quarks limited to ≤6/4

Recent estimates give ~3 billion CPU-hours per year!

---

[1] arxiv:1905.05120

See also 1908.00167, 2004.13687

Agliardi, Gabriele, et al. "**Quantum integration of elementary particle processes**." *arXiv preprint arXiv:2201.01547* (2022)

Monte Carlo integration for HEP

integrand

$$I = \int dx\, f(x) g(x)$$

probability distributions

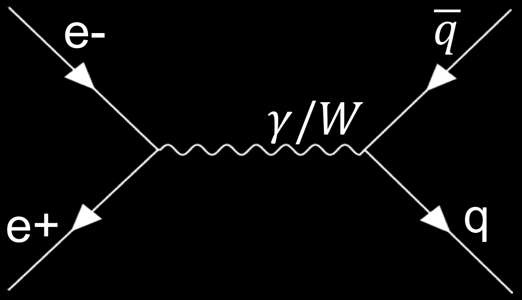
phase-space factor [possibly including parton-distribution function (PDF)]

$$\sigma = \frac{1}{F}\int d\Phi |M|^2 \Theta(\Phi - \Phi_c)$$

matrix element squared which encodes the quantum mechanical process

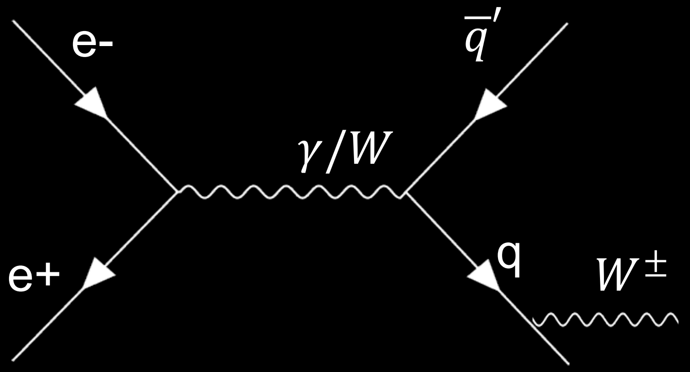
phase-space cuts

QED



$$\sigma \sim \int_{-1}^{1}\int_{0}^{2\pi} d\cos\theta d\phi(1 + cos^2\theta)$$

QED - EW



$$\sigma \sim \int_{M_W^2}^{s}\int_{0}^{s_1^{Max}}\int_{-1}^{1}\int_{0}^{2\pi}\int_{0}^{2\pi} d\phi_3 \left|\mathcal{M}_{e^+e^-\to q\bar{q}'W}\right|^2$$

where $S_1^{Max} = (s_2 - M_W)(s - s_2)/s_2$ and $d\phi_3 = ds_2 ds_1 d\cos\theta_1 d\phi_1 d\phi_2$

Agliardi, Gabriele, et al. "**Quantum integration of elementary particle processes**." *arXiv preprint arXiv:2201.01547* (2022)

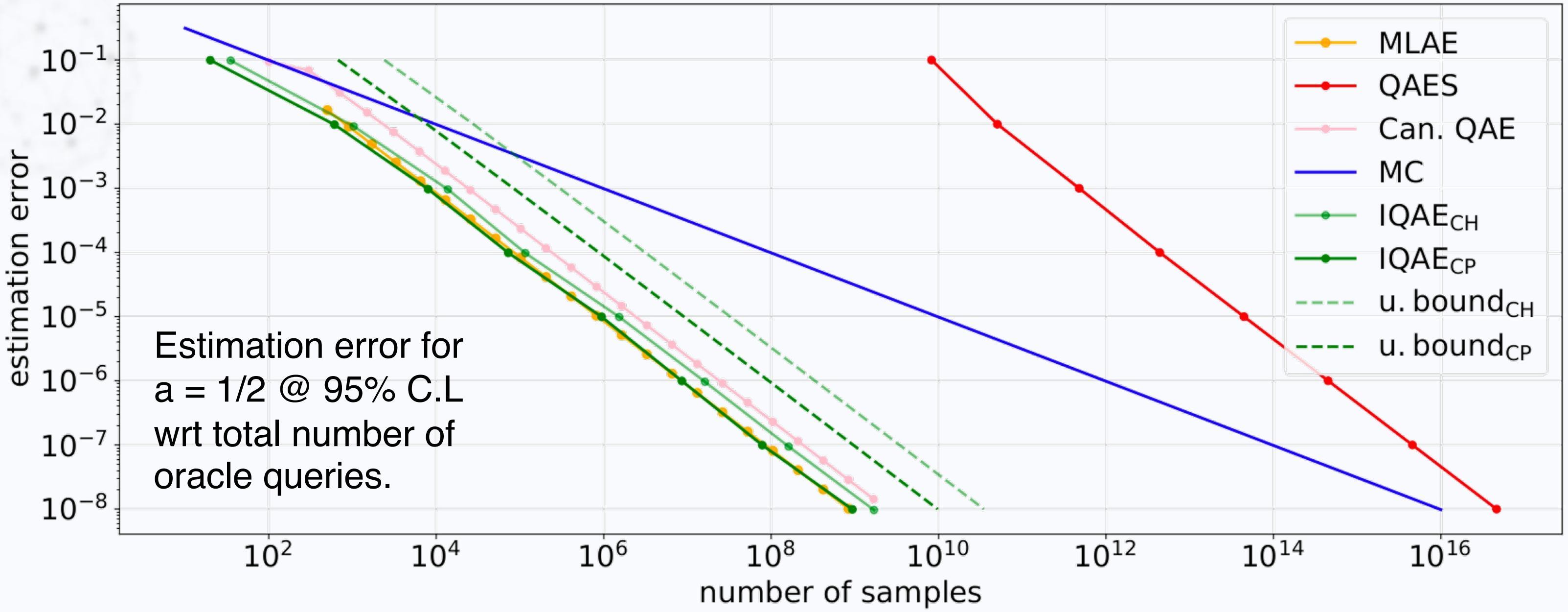
Quantum Monte Carlo

Quantum Monte Carlo is based on **Quantum Amplitude Estimation** (*Brassard, Hoyer, Mosca, Tapp, arxiv:0005055*)
→ quadratic speedup from Grover's algorithm

$$\mathcal{A}|0\rangle = \sqrt{1-a}|\Psi_0\rangle + \sqrt{a}|\Psi_1\rangle$$

QAE estimates **a** with high probability such that the estimation error scales as **O(1/M)** instead of **O(1/√M)**

Different implemetations available. (Grinko, Gacon, Zoufal, Woerner; arxiv: 1912.05559)

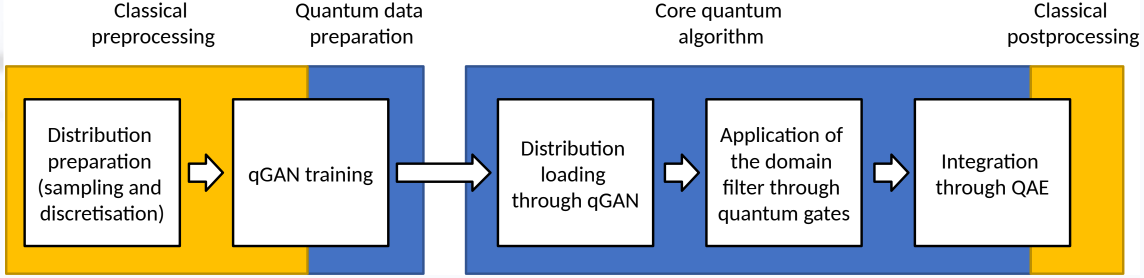


Estimation error for a = 1/2 @ 95% C.L wrt total number of oracle queries.

Agliardi, Gabriele, et al. "**Quantum integration of elementary particle processes.**" *arXiv preprint arXiv:2201.01547* (2022)

qGAN for data encoding

Classical *f* encoding affects the quality and speed of integration

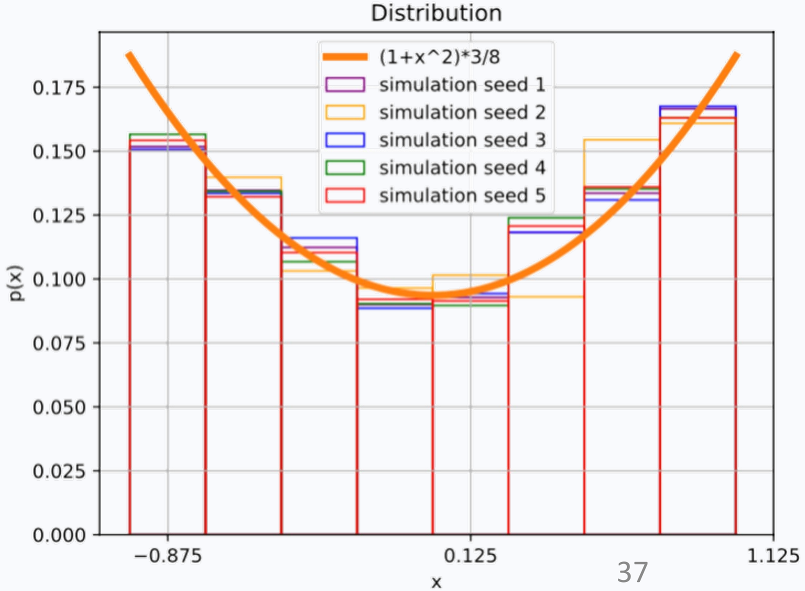


integrand

$$\sigma = \frac{1}{F}\int d\Phi |M|^2 \Theta(\Phi - \Phi_c)$$

probability distributions

$$G(\phi)\left|\psi_{in}\right\rangle = \left|g(\phi)\right\rangle = \sum_{i=0}^{N-1}\sqrt{p_g^i(\phi)}\left|i\right\rangle$$



Use a **quantum GAN:**

- **Efficient learning** of probabilities over discrete values
- Control **resolution through number of qubits**

Test on $1 + x^2$ distribution:

- 10k events, 3 qubits, circular entanglement

Loading	Difference per bin [%]			σ_x
	Min.	Max.	Average	
Direct	+0.207	−1.88	1.35	1.80×10^{-3}
qGAN default	+2.36	−21.1	8.51	0.0118
qGAN optimised	−0.995	−12.4	4.65	7.00×10^{-3}

Jorge J. Martínez de Lejarza, 2305.01686

Fourier series for QMC

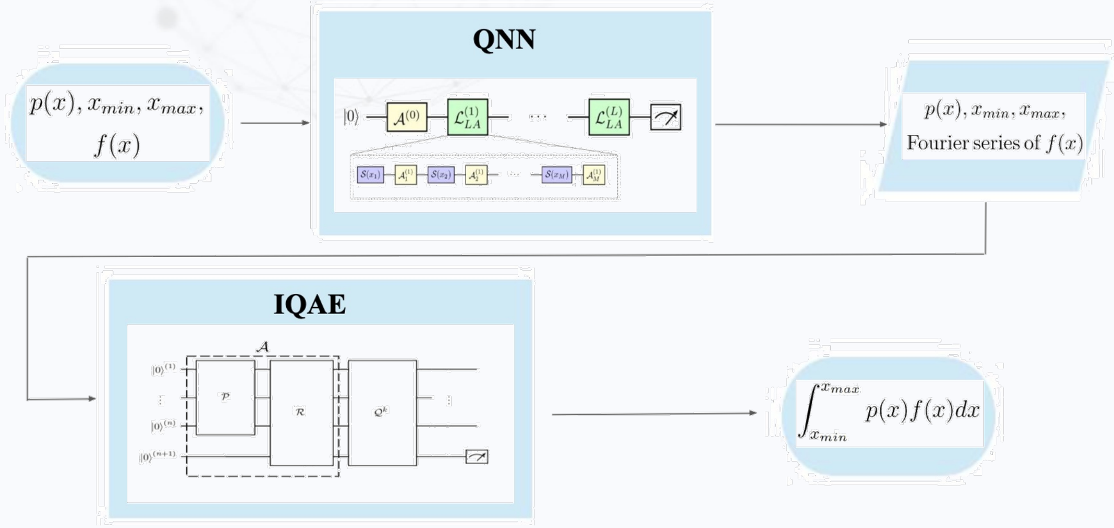
Encode Fourier expansions coefficients and integrate trigonometric functions
FQMCI: **Herbert [Q6, 823 (2022)](#)**

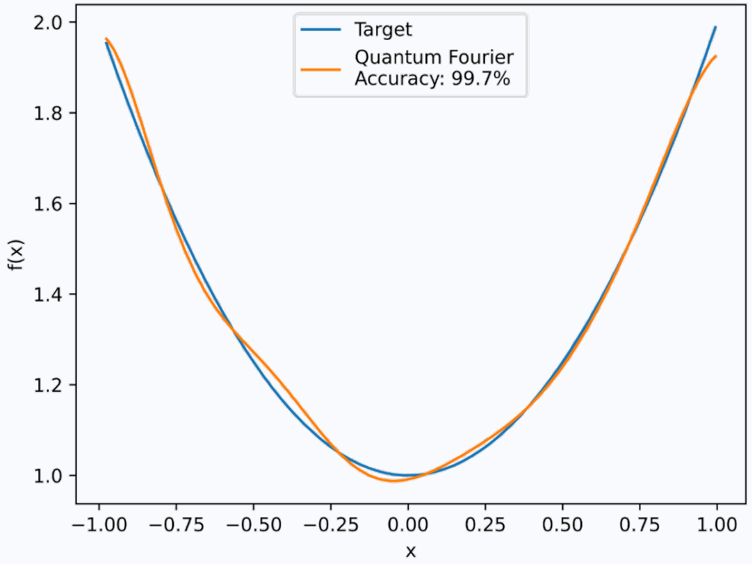
- Assumes Fourier coefficients are analytically computed

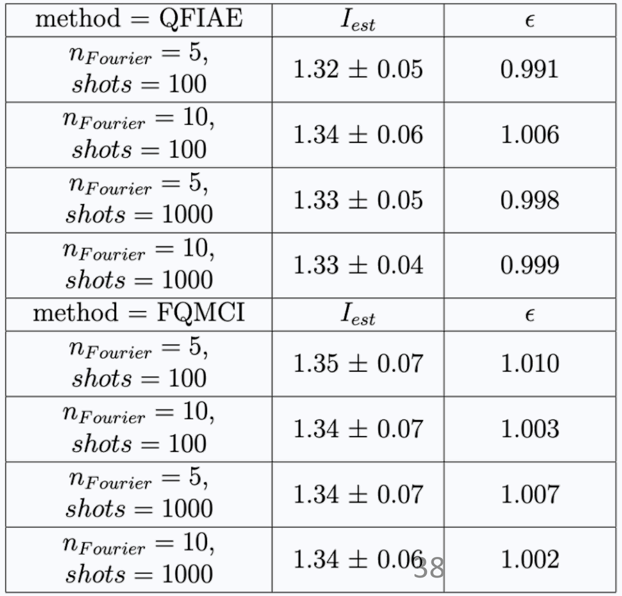
New approach provides a end-to-end quantum implementation

QFIAE (Quantum Fourier Iterative Amplitude Estimation)

- Fourier series decomposition through a QNN
- Integrate each trigonometric component using QAE

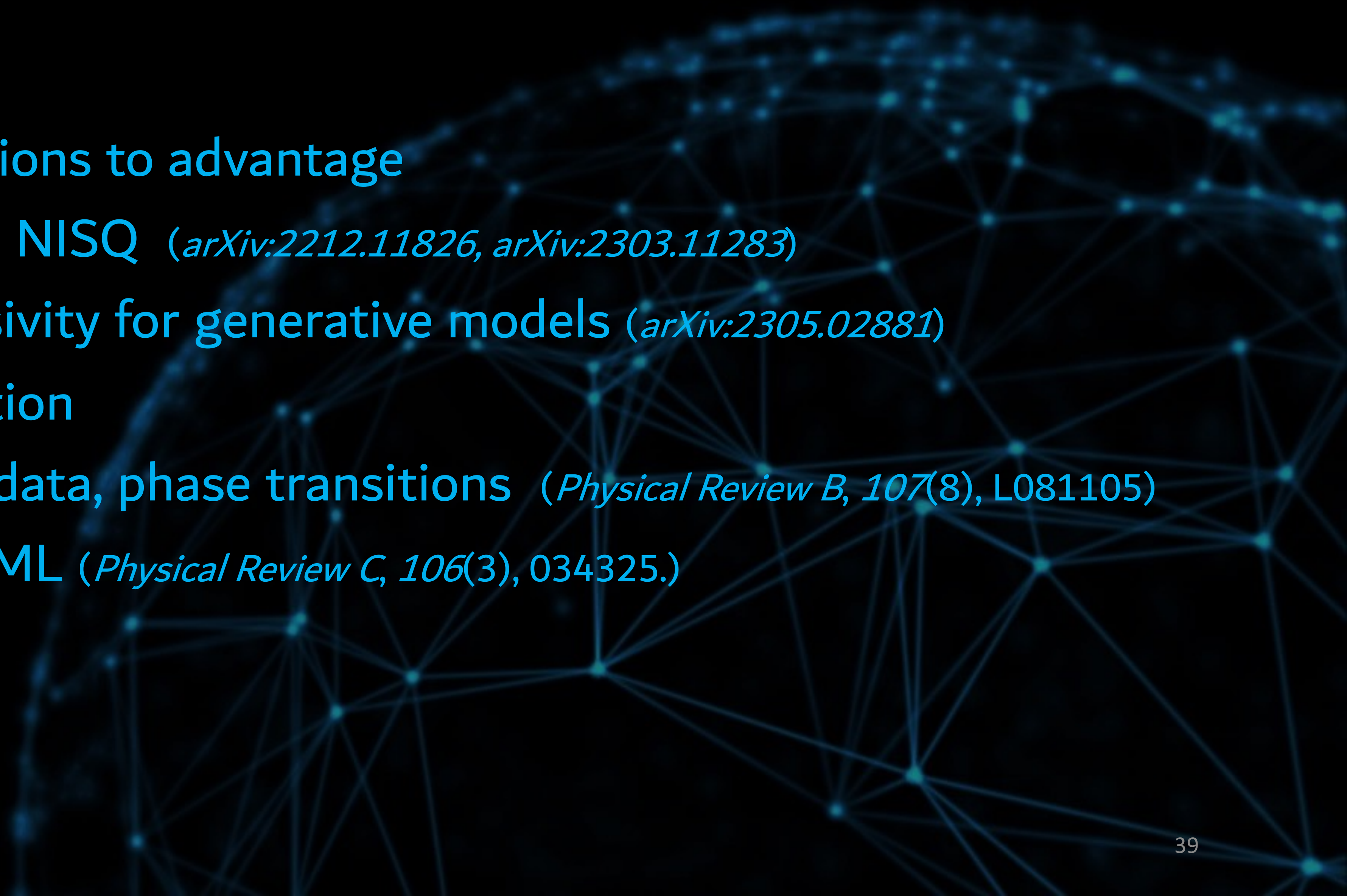






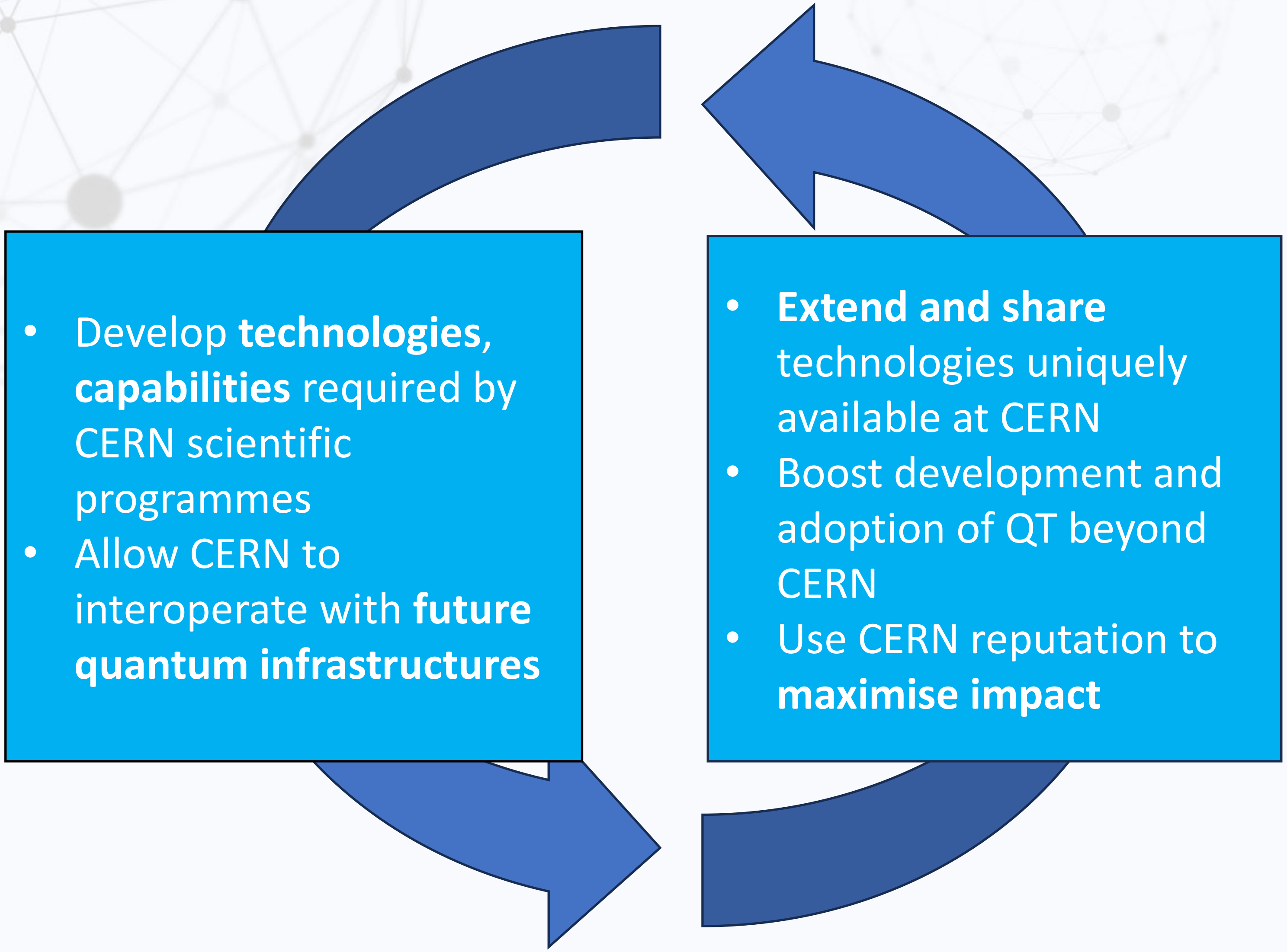
method = QFIAE	I_{est}	ϵ
$n_{Fourier} = 5$, $shots = 100$	1.32 ± 0.05	0.991
$n_{Fourier} = 10$, $shots = 100$	1.34 ± 0.06	1.006
$n_{Fourier} = 5$, $shots = 1000$	1.33 ± 0.05	0.998
$n_{Fourier} = 10$, $shots = 1000$	1.33 ± 0.04	0.999
method = FQMCI	I_{est}	ϵ
$n_{Fourier} = 5$, $shots = 100$	1.35 ± 0.07	1.010
$n_{Fourier} = 10$, $shots = 100$	1.34 ± 0.07	1.003
$n_{Fourier} = 5$, $shots = 1000$	1.34 ± 0.07	1.007
$n_{Fourier} = 10$, $shots = 1000$	1.34 ± 0.06	1.002

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.)

Why does CERN engage in Quantum Technologies?



QT4HEP

Can CERN stay out of quantum technologies?

- Develop **technologies, capabilities** required by CERN scientific programmes
- Allow CERN to interoperate with **future quantum infrastructures**

- **Extend and share** technologies uniquely available at CERN
- Boost development and adoption of QT beyond CERN
- Use CERN reputation to **maximise impact**

HEP4QT

How can CERN contribute to quantum technologies?

CERN QTI Phase 2

CERN QUANTUM TECHNOLOGY PLATFORMS

HYBRID QUANTUM COMPUTING AND ALGORITHMS

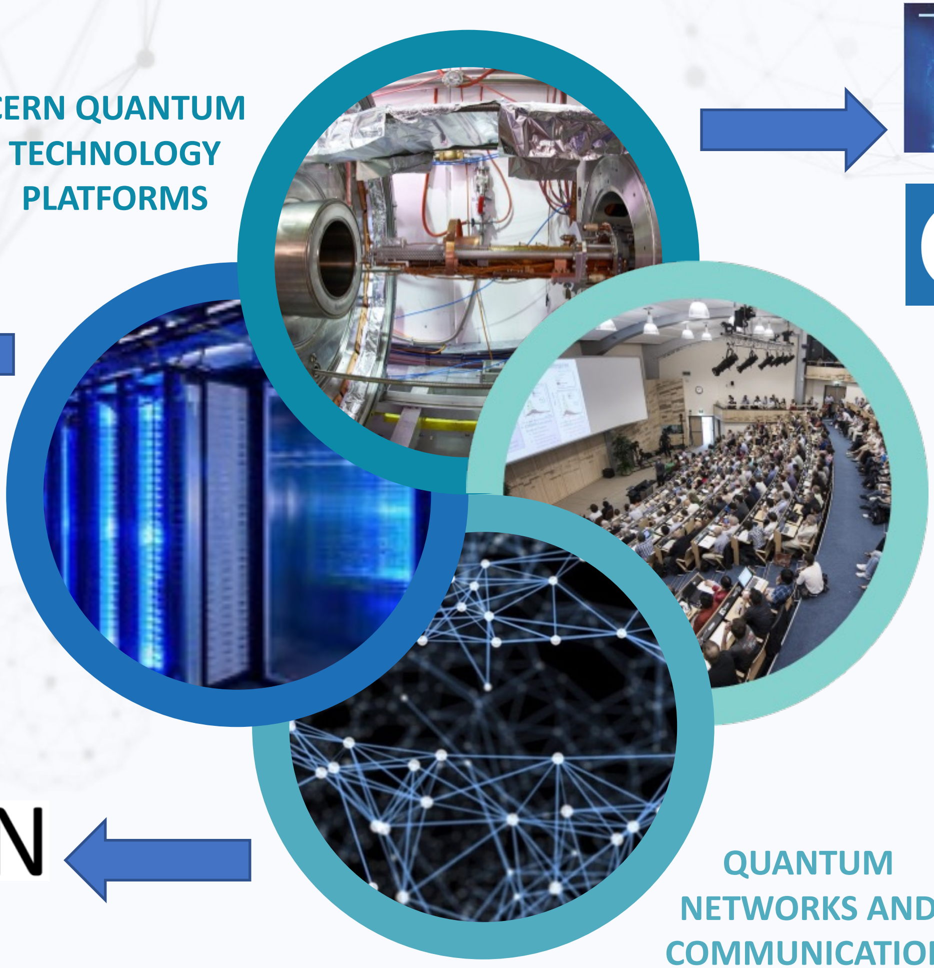
COLLABORATION FOR IMPACT

QUANTUM NETWORKS AND COMMUNICATIONS

CERN

QUANTUM
TECHNOLOGY
INITIATIVE

CERN QTI Phase 2 Impact



CERN QUANTUM TECHNOLOGY PLATFORMS

HYBRID QUANTUM COMPUTING AND ALGORITHMS

COLLABORATION FOR IMPACT

QUANTUM NETWORKS AND COMMUNICATIONS

Physics Beyond Colliders

FCC hh ee he

HiLumi HL-LHC PROJECT

WLCG
Worldwide LHC Computing Grid

LHCOPN

LHC 1

OPEN QUANTUM INSTITUTE

CERN | QUANTUM TECHNOLOGY INITIATIVE

Impact: Bringing together HEP and QC experts

CERN DESY. IBM Quantum

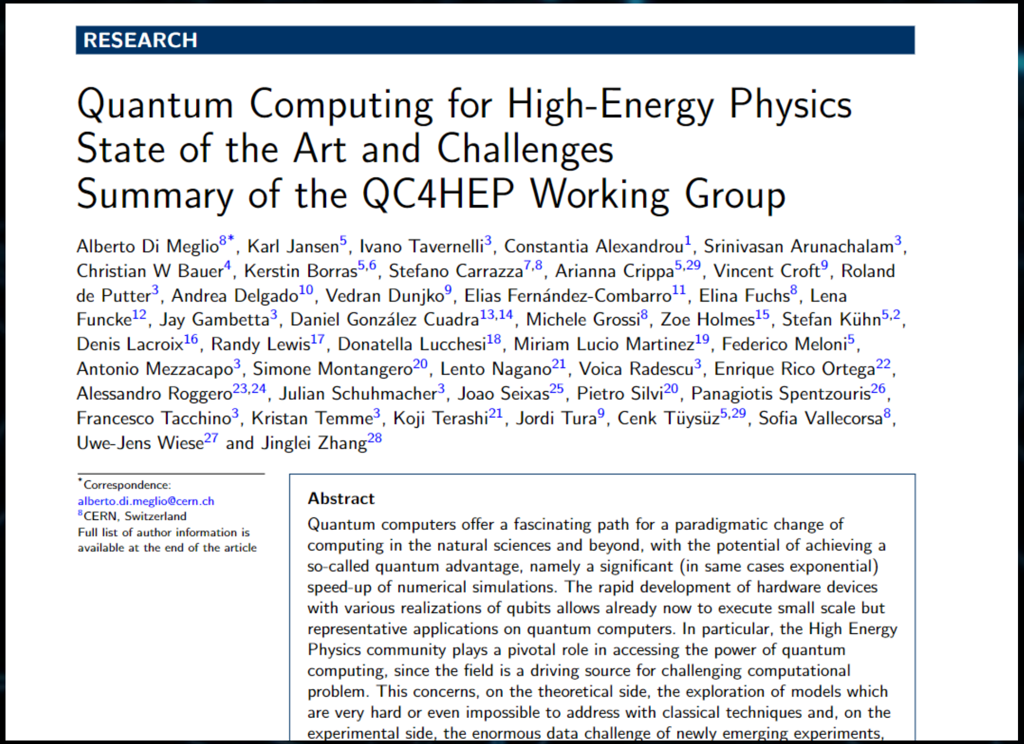
QC4HEP: High Energy Physics Working Group

https://doi.org/10.48550/arXiv.2307.03236

Kick-off meeting last November

When: Nov 3-4, 2022, within the QT4HEP Conference week
Where: CERN, Geneva
Who: organized by **CERN/DESY/IBM Quantum** with invited HEP and QC experts from IBM Quantum Network and beyond (by invitation only)

45+	19	3
Participants	Institutes	Continents



RESEARCH

Quantum Computing for High-Energy Physics
State of the Art and Challenges
Summary of the QC4HEP Working Group

Alberto Di Meglio[8*], Karl Jansen[5], Ivano Tavernelli[3], Constantia Alexandrou[1], Srinivasan Arunachalam[3], Christian W Bauer[4], Kerstin Borras[5,6], Stefano Carrazza[7,8], Arianna Crippa[5,29], Vincent Croft[9], Roland de Putter[3], Andrea Delgado[10], Vedran Dunjko[9], Elias Fernández-Combarro[11], Elina Fuchs[8], Lena Funcke[12], Jay Gambetta[3], Daniel González Cuadra[13,14], Michele Grossi[8], Zoe Holmes[15], Stefan Kühn[5,2], Denis Lacroix[16], Randy Lewis[17], Donatella Lucchesi[18], Miriam Lucio Martinez[19], Federico Meloni[5], Antonio Mezzacapo[3], Simone Montangero[20], Lento Nagano[21], Voica Radescu[3], Enrique Rico Ortega[22], Alessandro Roggero[23,24], Julian Schuhmacher[3], Joao Seixas[25], Pietro Silvi[20], Panagiotis Spentzouris[26], Francesco Tacchino[3], Kristan Temme[3], Koji Terashi[21], Jordi Tura[9], Cenk Tüysüz[5,29], Sofia Vallecorsa[8], Uwe-Jens Wiese[27] and Jinglei Zhang[28]

---

[*] Correspondence:
alberto.di.meglio@cern.ch
[8] CERN, Switzerland
Full list of author information is available at the end of the article

Abstract

Quantum computers offer a fascinating path for a paradigmatic change of computing in the natural sciences and beyond, with the potential of achieving a so-called quantum advantage, namely a significant (in same cases exponential) speed-up of numerical simulations. The rapid development of hardware devices with various realizations of qubits allows already now to execute small scale but representative applications on quantum computers. In particular, the High Energy Physics community plays a pivotal role in accessing the power of quantum computing, since the field is a driving source for challenging computational problem. This concerns, on the theoretical side, the exploration of models which are very hard or even impossible to address with classical techniques and, on the experimental side, the enormous data challenge of newly emerging experiments,

Next meeting at CERN on November 16-17, 2023

Outlook and open questions

- Quantum computing offers great opportunties while HEP provides challenging problems
  - **What are the most promising applications?**
  - How do we define performance and validate results on **realistic use cases**?

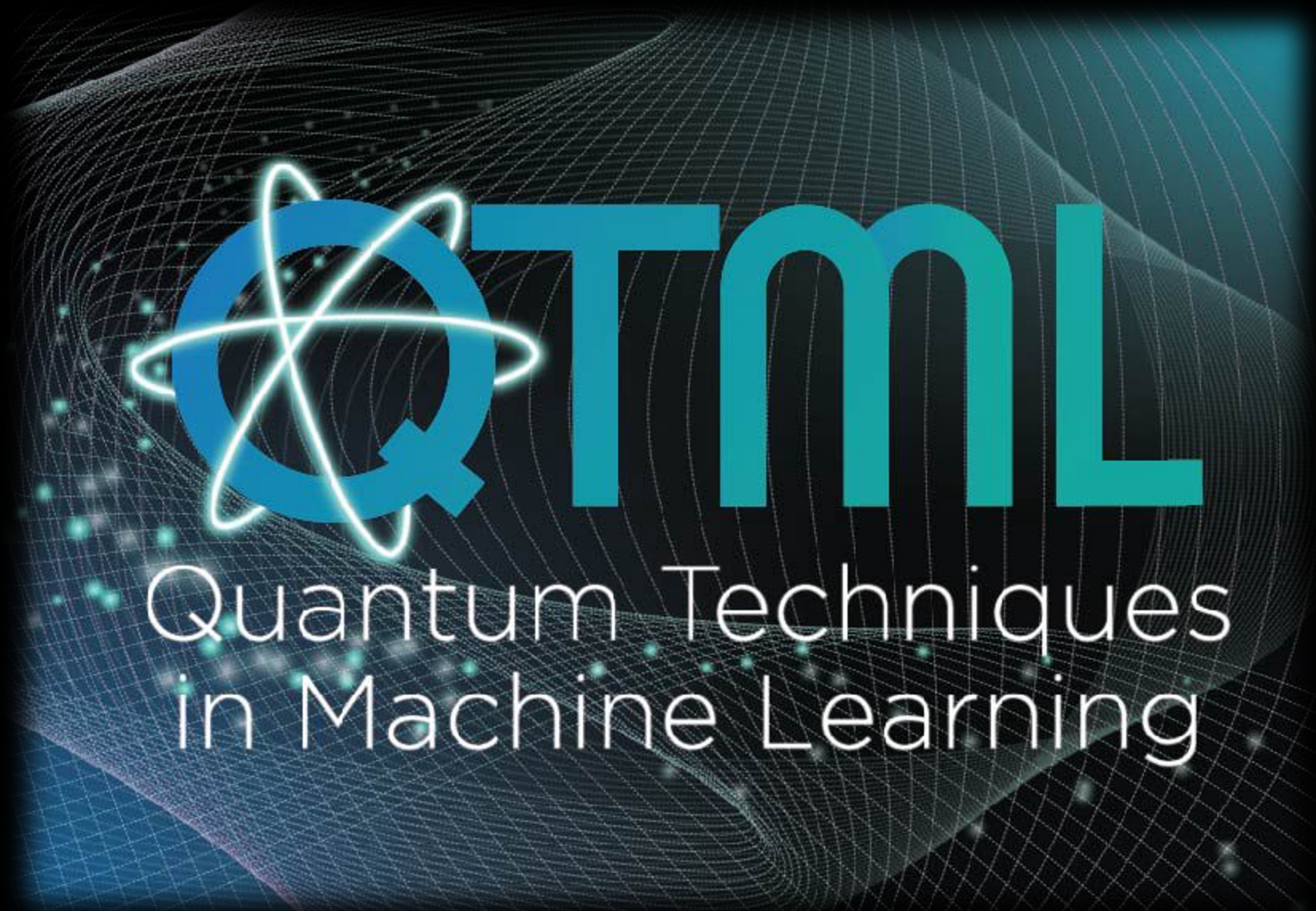
- Experimental data has high dimensionality
  - Can we train Quantum Machine Learning algorithms effectively?
  - Can we reduce the impact of **data reduction** techniques?
- Experimental data is shaped by physics laws
  - Can we leverage them to build better algorithms?
- CERN is committed to creating impact on QT research in the coming years

Thank you!

November 20th-24th, 2023
@CERN

https://qtml-2023.web.cern.ch/

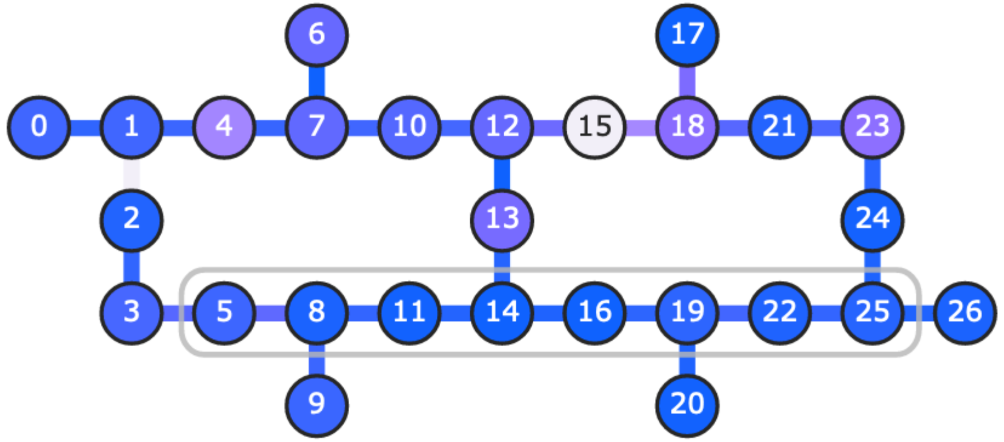
Sofia.Vallecorsa@cern.ch

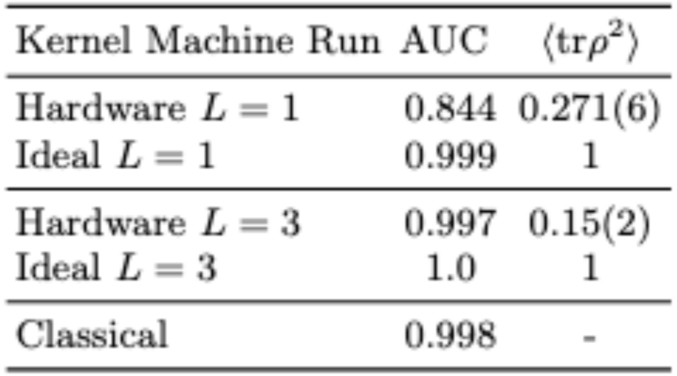


Preliminary hardware runs

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	-

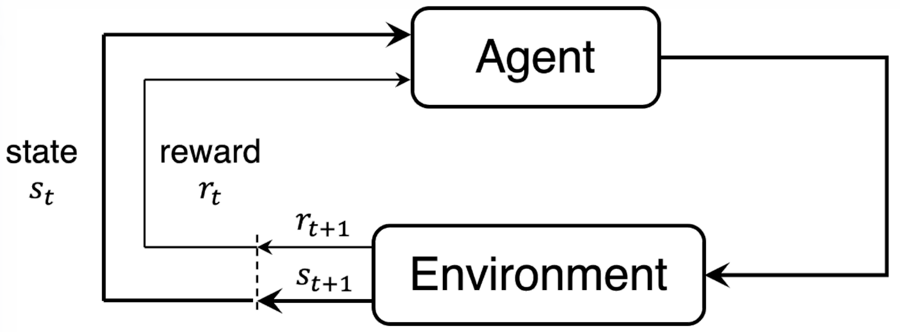
Quantum Machine Learning examples:

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

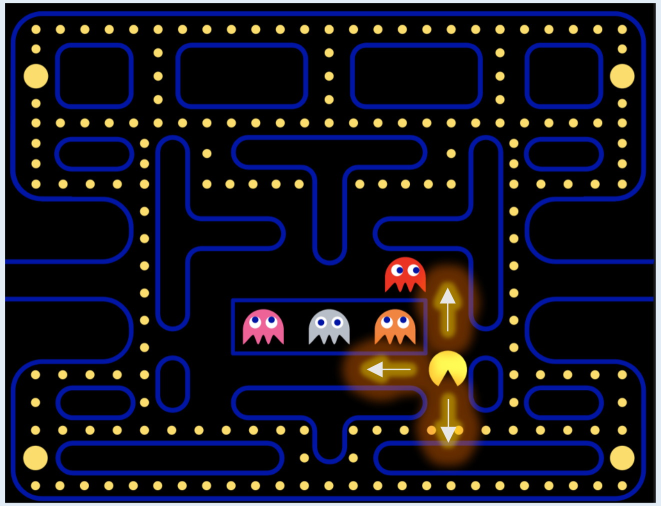
- Agent **takes actions** in environment and **collects rewards**

Q-learning

- Estimate **return using Q-function** $\mathrm{Q}(\mathbf{s},\mathbf{a})$
- **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)

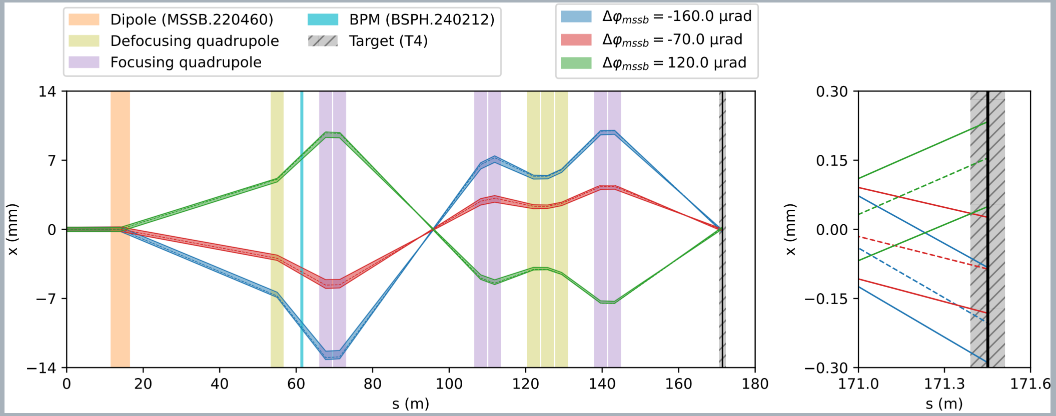
Key concept: sample-efficiency

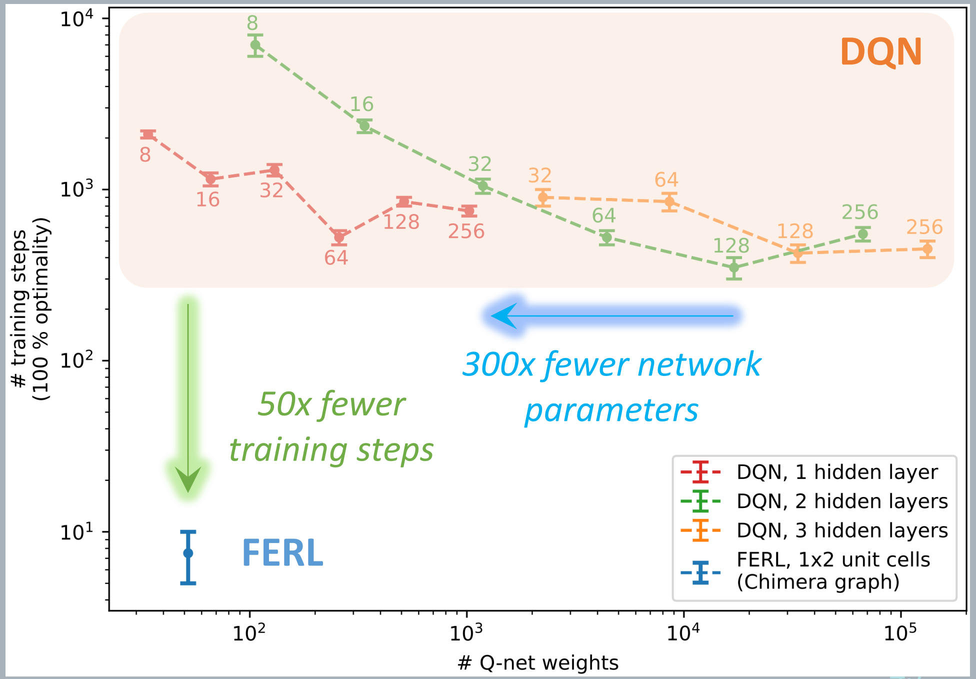
- Relevant for **particle accelerator control** given cost of beam time *(online training)*

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

1^{st} study: 1D beam steering

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





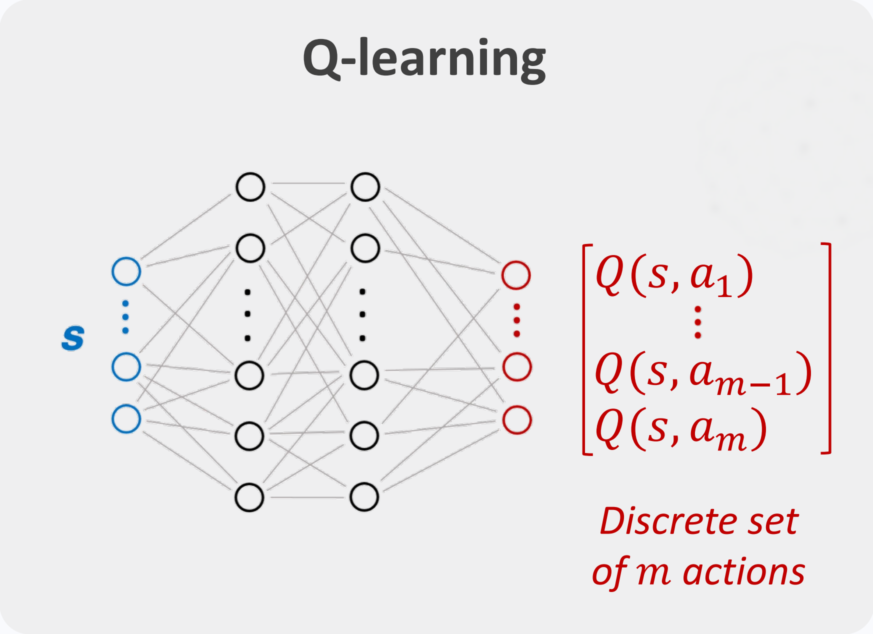
50x fewer training steps

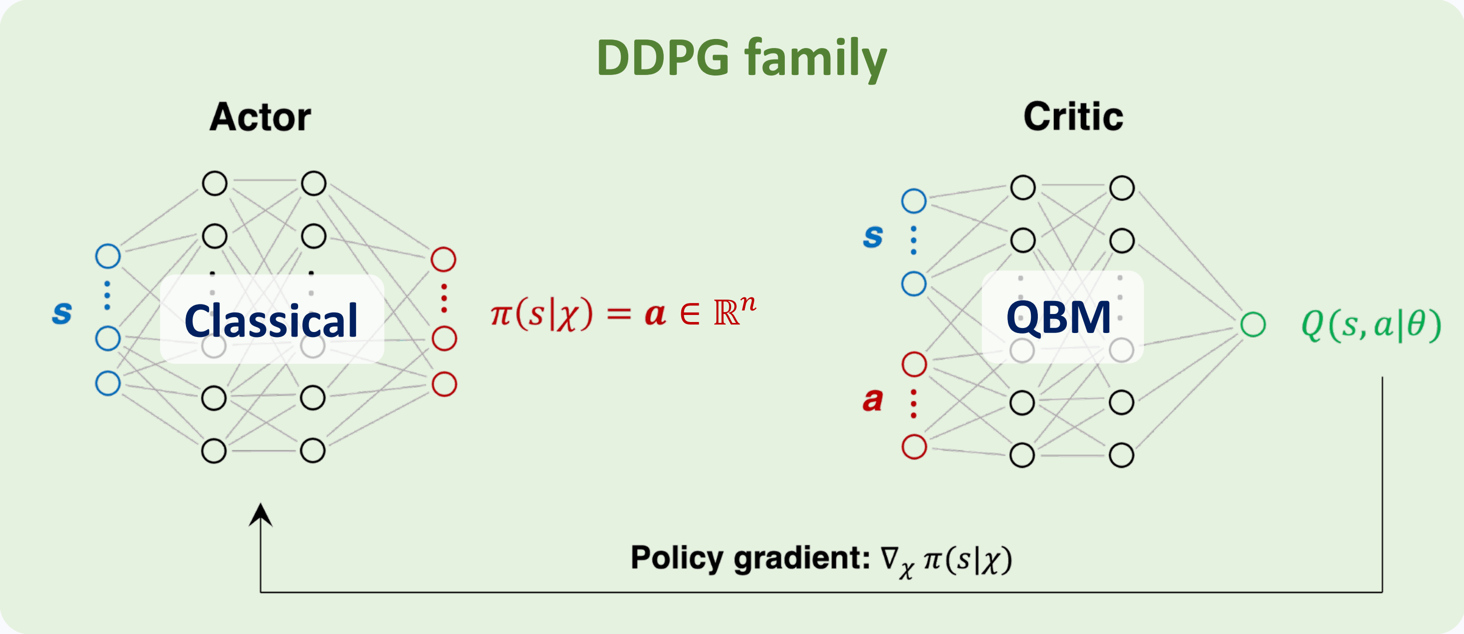
300x fewer network parameters

Developing a hybrid actor-critic scheme

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

- **QBM replaces classical critic net**



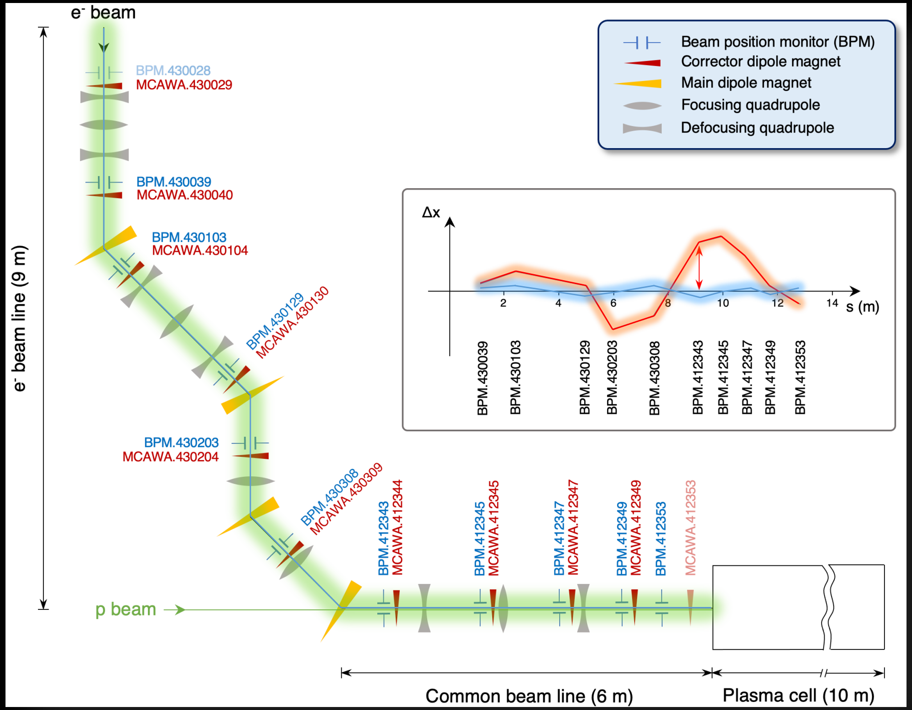


2nd study: 10D continuous beam steering

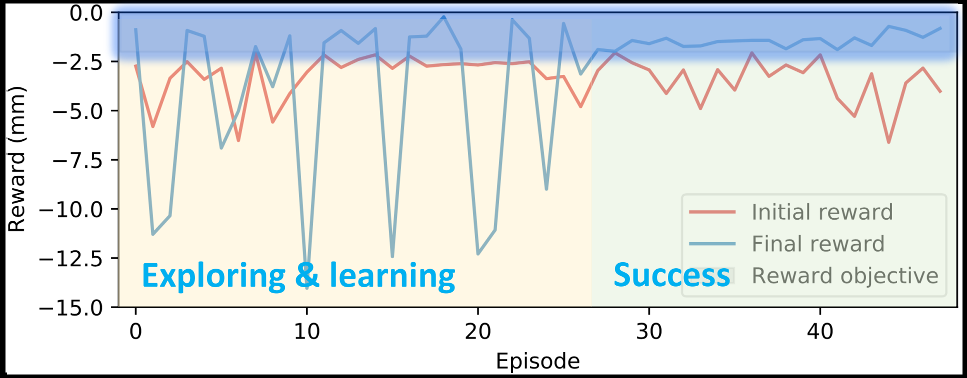
Environment: e⁻ beam line of AWAKE

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

AWAKE: Advanced Proton Driven Plasma Wakefield Acceleration Experiment



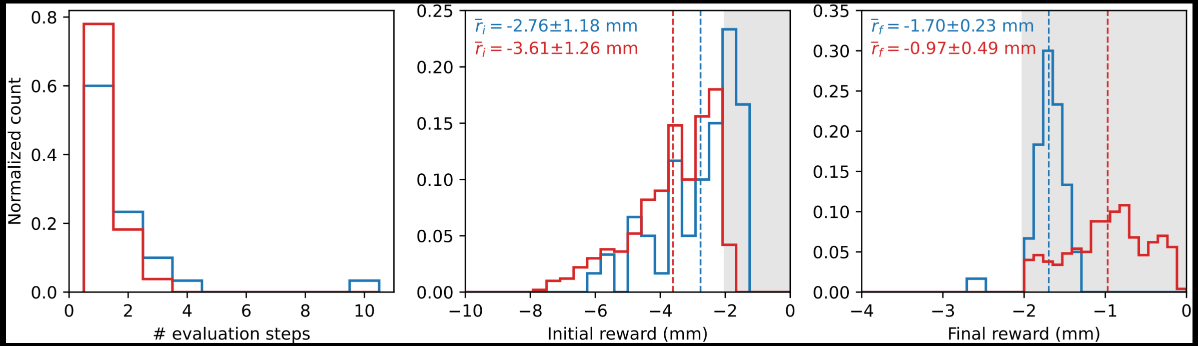
Training: on D-Wave Advantage quantum annealer (QA)



Objective

Evaluation: on actual beam line

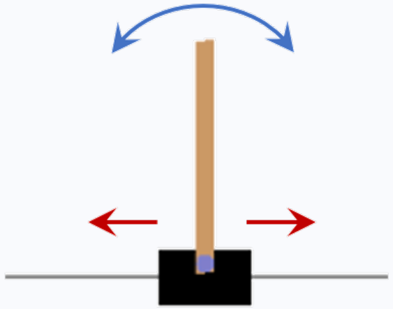
Real vs. simulated QA



- Agent minimizes rms in **1 step in 60 % cases**
- **Hyperparameter** tuning with **simulated** QA

3rd study: Cartpole-v1

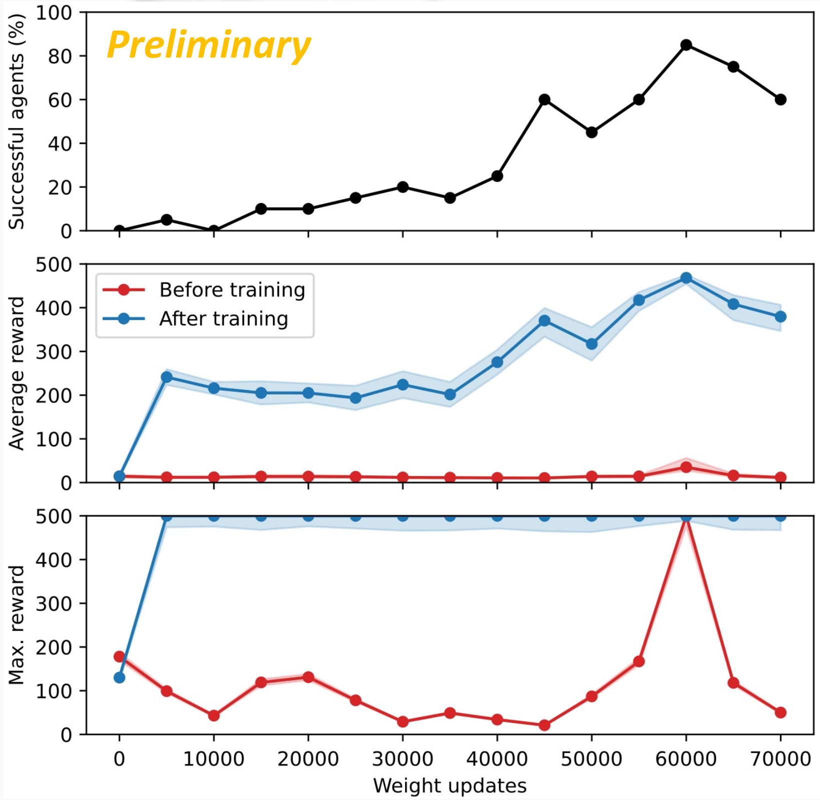
Discrete action problem, non-linear dynamics



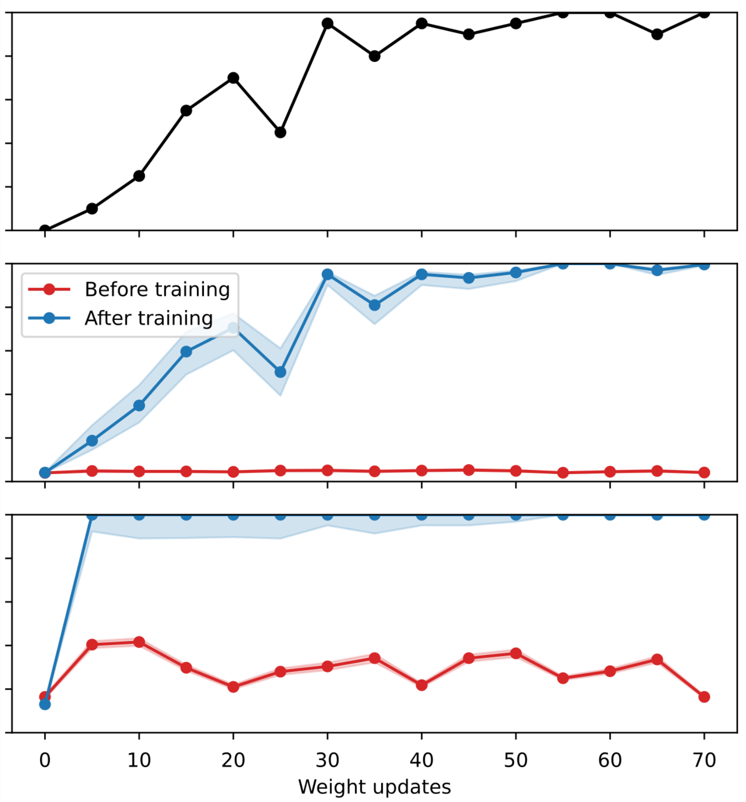
- **Cartpole-v1:** official OpenAI gym env from classic control problems domain
- Continuous state (4D), **discrete action** *(right, left)* problem with **non-linear dynamics**
- Terminate episodes after **max. 500 steps**

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

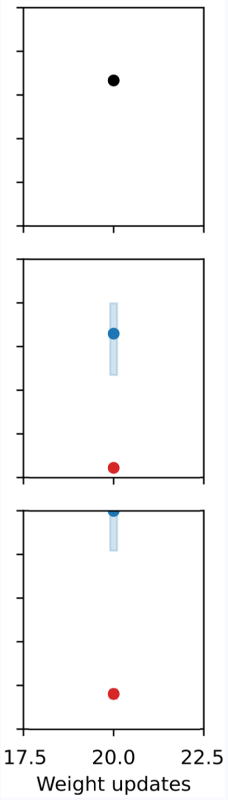
DQN

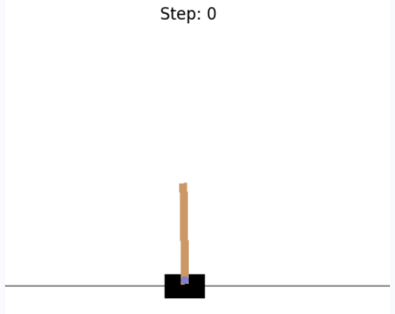


FERL *(simulated QA)*

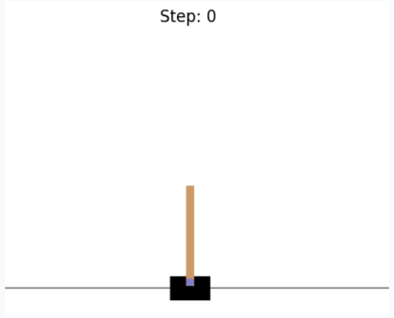


FERL *(trained on D-Wave)*





before



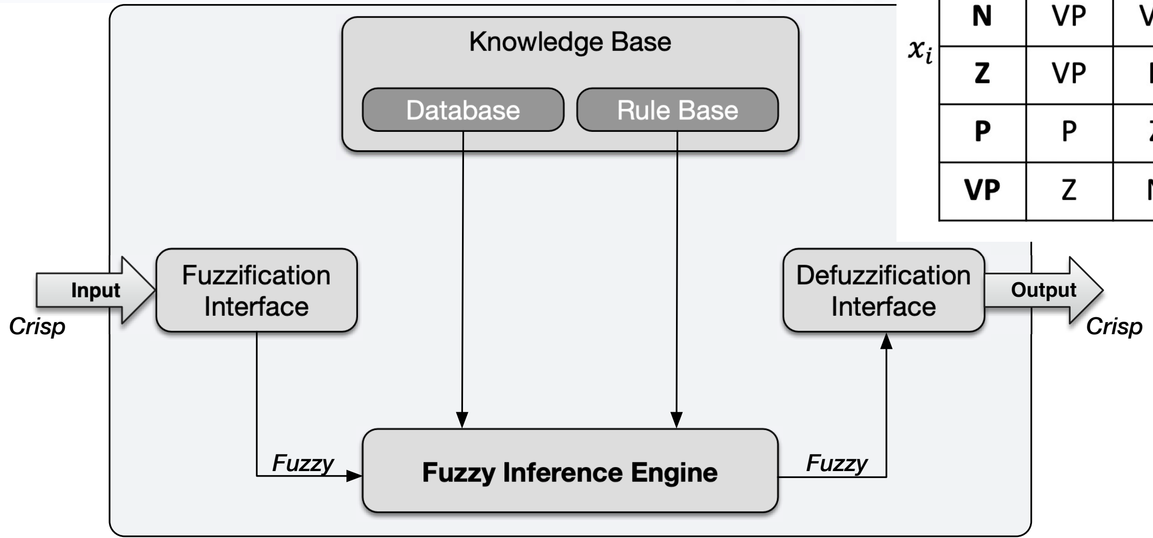
after

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