Quantum Machine Learning in High Energy Physics

Sofia Vallecorsa

AI and Quantum Research - CERN IT

CERN

Outline

- Introduction
- The CERN Quantum Technlogy Initiative
- Qubits and circuits
- Quantum Machine Learning
- Applications in High Energy Physics
- Examples from CERN
- Summary

https://www.nature.com/articles/s41586-019-1666-5

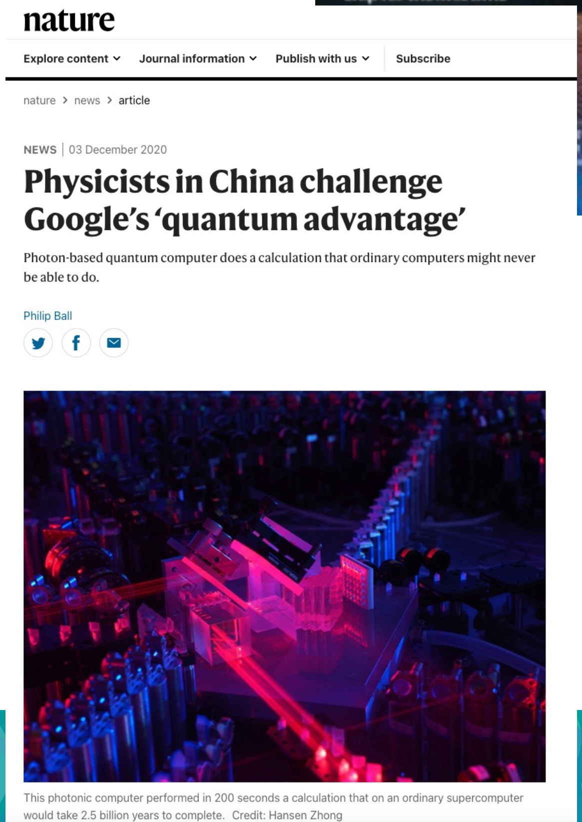
Quantum Advantage?

2019, Google: quantum advantage by solving sampling problem in 200s on Sycamore vs estimated 10k years on Summit

2020, Hefei National Lab, China: quantum advantage on boson sampling using a photonic computer

Multiple works followed, discussing and/or reducing those claims

- Exponential advantage in data representation
- Complex algorithms acceleration
  - Efficient **sampling**, **searches** and **optimization**
  - Linear algebra, matrices and machine learning
- New algorithms/methods for **cryptography** and **communication**
- **Direct simulation** of quantum systems


nature

Explore content · Journal information · Publish with us · Subscribe

nature > news > article

NEWS | 03 December 2020

Physicists in China challenge Google's 'quantum advantage'

Photon-based quantum computer does a calculation that ordinary computers might never be able to do.

Philip Ball

This photonic computer performed in 200 seconds a calculation that on an ordinary supercomputer would take 2.5 billion years to complete. Credit: Hansen Zhong

https://www.nature.com/articles/d41586-020-03434-7

CERN QTI and its Roadmap

CERN established the QTI in 2020

- Roadmap in 2021
- Publicly available on Zenodo
  - Accessed more than 6000 times

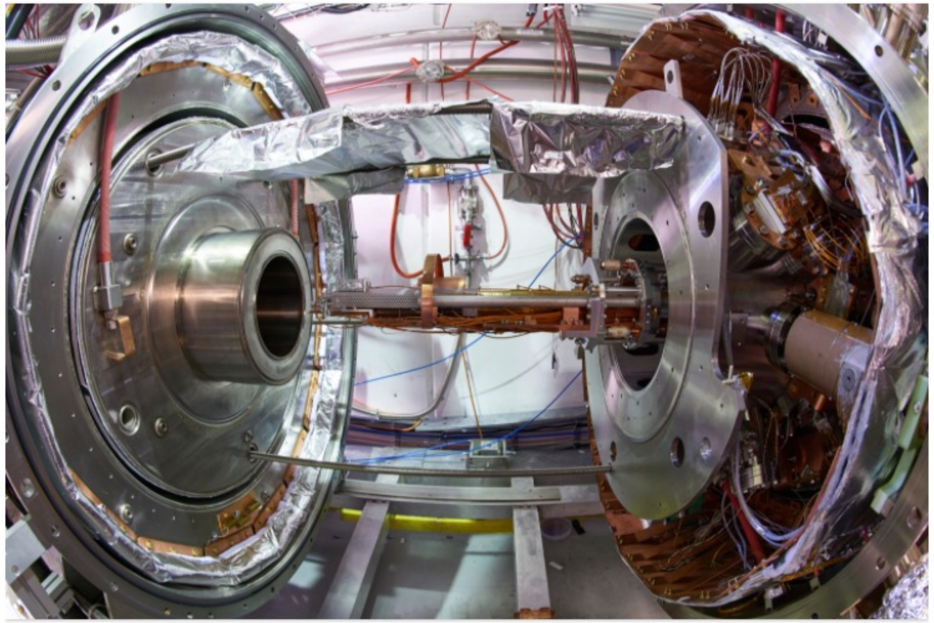
https://doi.org/10.5281/zenodo.5553774

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)

T1 - Scientific and Technical Development and Capacity Building

T2 - Co-development

APPLICATIONS | NEWS

CERN unveils roadmap for quantum technology

4 November 2021

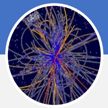
QUANTUM TECHNOLOGY INITIATIVE

Credit: CERN

T3 - Community Building

T4 - Integration with national and international initiatives and programmes

Scientific Objectives



- Assess the **areas of potential quantum advantage** in HEP (QML, classification, anomaly detection, tracking)
- Develop **common libraries of algorithms, methods, tools**; benchmark as technology evolves
- Collaborate to the development of shared, **hybrid classic-quantum infrastructures**

Computing & Algorithms

- Identify and develop techniques for **quantum simulation** in collider physics, QCD, cosmology within and beyond the SM
- Co-develop quantum computing and sensing approaches by providing **theoretical foundations** to the identifications of the areas of interest

Simulation & Theory

- Develop and promote **expertise in quantum sensing** in low- and high-energy physics applications
- Develop quantum sensing approaches with emphasis on **low-energy particle physics measurements**
- Assess **novel technologies and materials** for HEP applications

Sensing, Metrology & Materials

- **Co-develop CERN technologies relevant to quantum infrastructures** (time synch, frequency distribution, lasers)
- Contribute to the **deployment and validation of quantum infrastructures**
- Assess requirements and **impact of quantum communication on computing applications** (security, privacy)

Communications & Networks

Quantum Computing at CERN

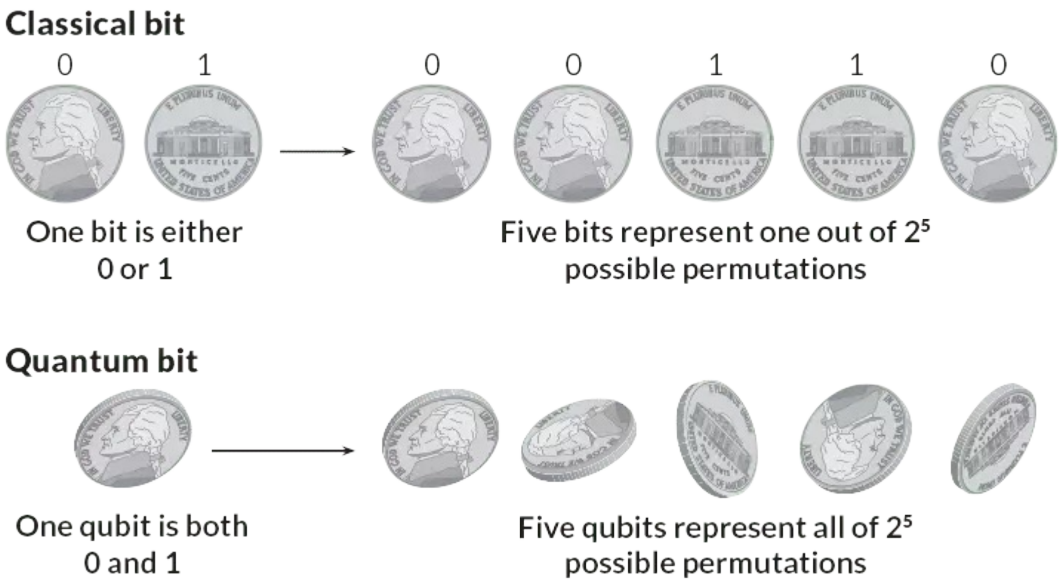
- QC is one of the four research areas in the CERN QTI
- Understand which applications can profit from quantum algorithms
  - Choose **representative use cases**
  - Understand **challenges and limitations** (on NISQ and fault tolerant hardware)
  - **Optimize** quantum algorithms
- Quantum Machine Learning algorithms are a primary candidate for investigation
  - Increasing use of ML in many computing and data analysis flows
  - Can be built as **hybrid models** where quantum computers act as accelerators
  - Efficient data handling is a challenge

Quantum Computing Intro

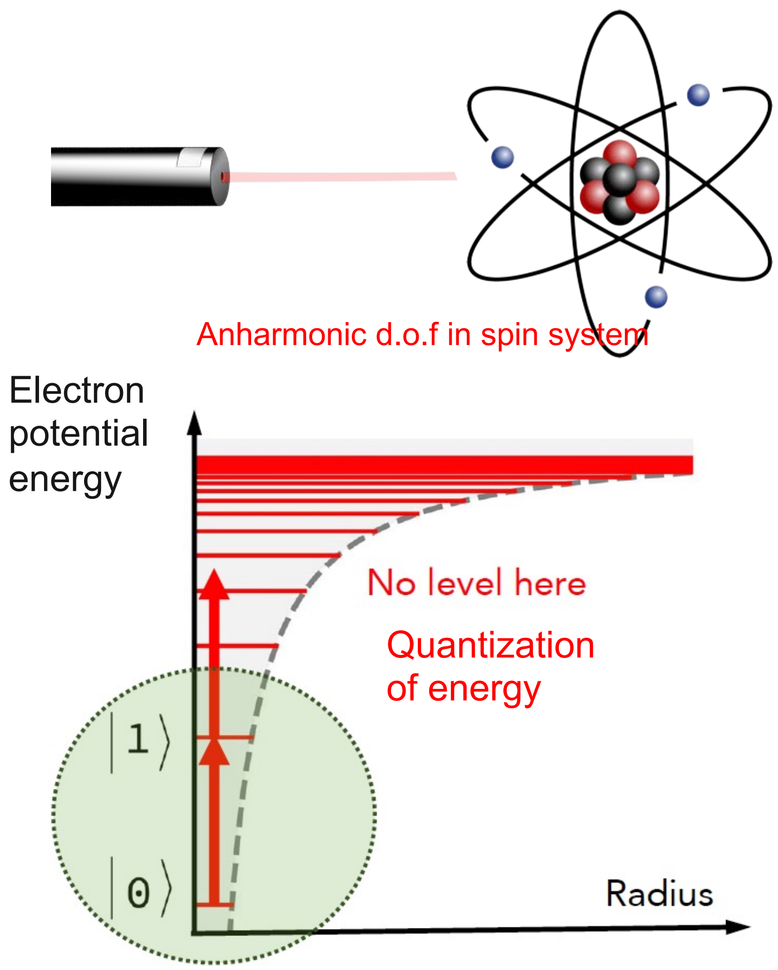
An Introduction to Quantum Computing, E. Combarro, https://indico.cern.ch/event/970905

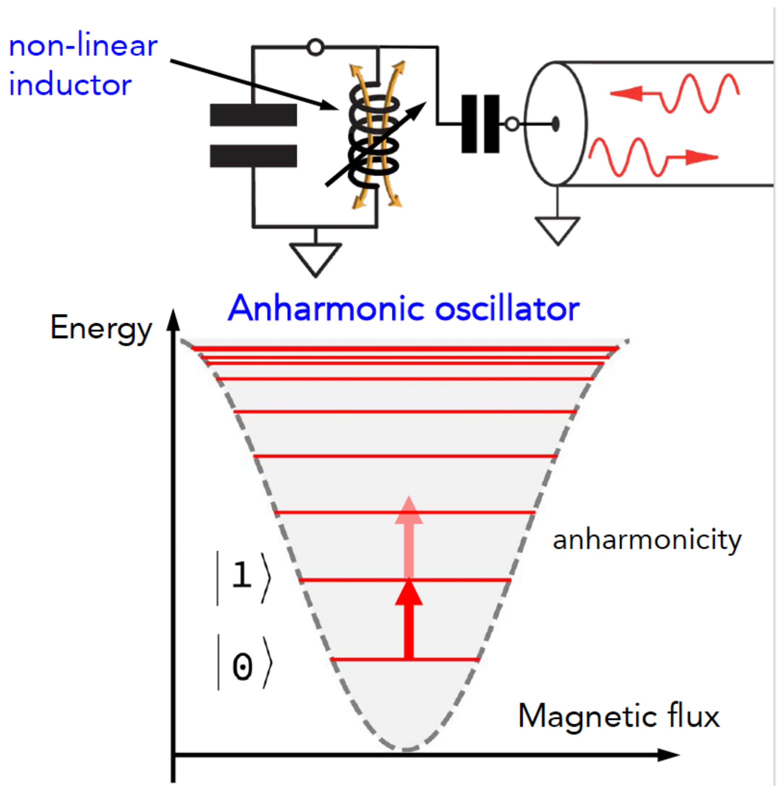
Qubit: Quantum bit

- **Classical bits are binary** “0 or 1”
- Quantum Mechanics predicts **superposition states** “simultaneously 0 and 1”
- **Superposition** can lead to highly parallel computations (**exponential speedup**)
- State of the “output qubit” has to be measured (**stochastic** nature of the result)
  - **Qubit state collapses**
- **No-cloning theorem**



Creating qubit: superconducting rings

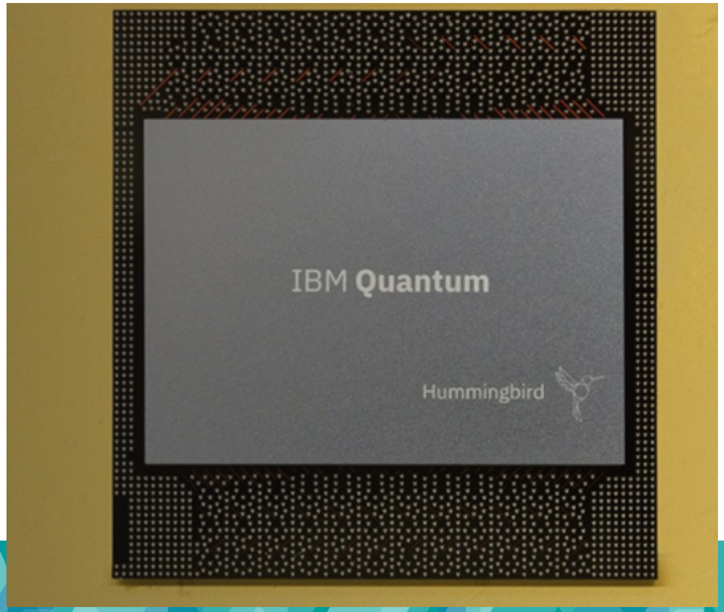




Z. Minev, Qiskit Global Summer School 2020

- Current oscillates in resistance-free circuit loop
- Injected microwave signal excites the current into super-position states

Ex. Google, IBM, …

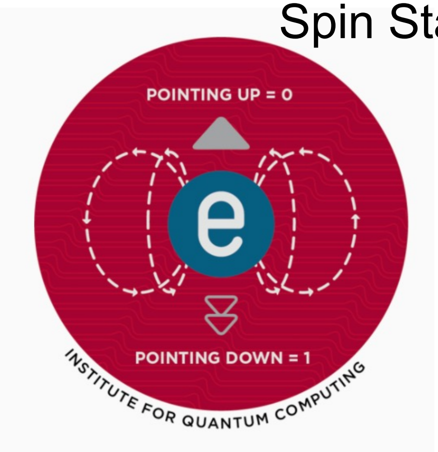


Different qubits

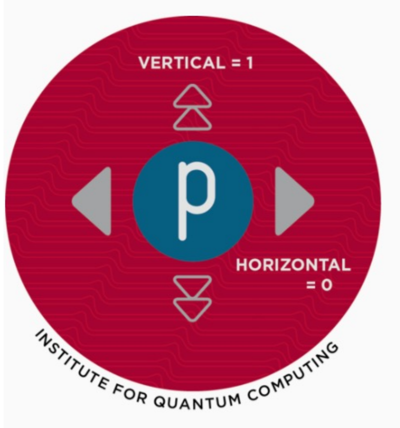
PHOTONS:

SuperConducting loops

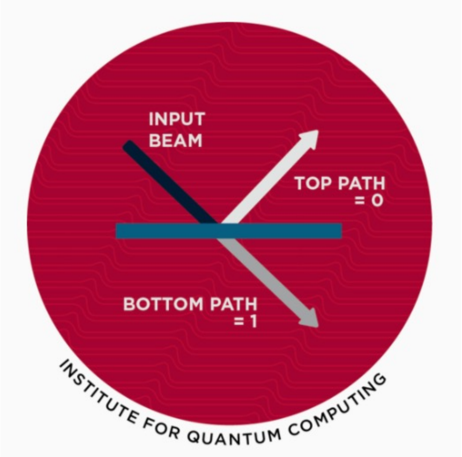
Spin States



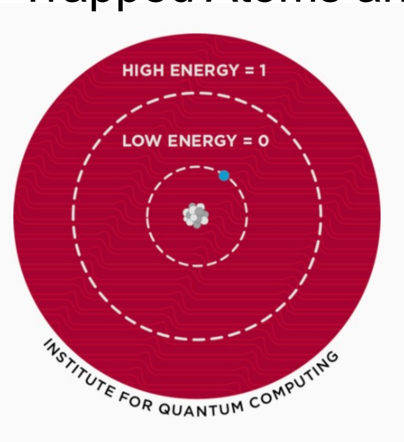
Polarization States

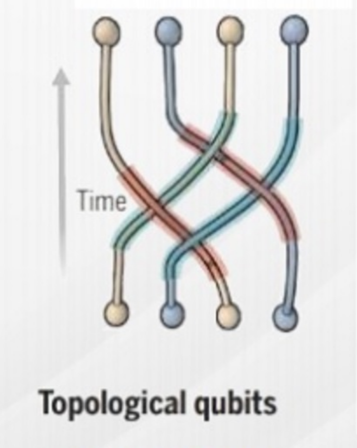


Path Qubits:

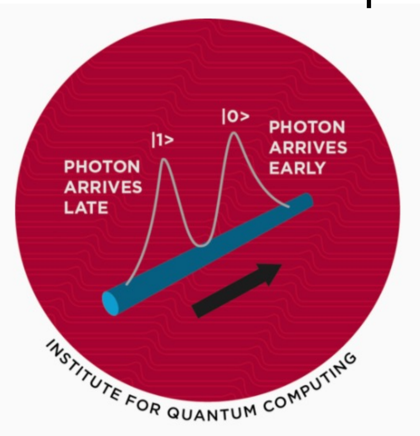


Trapped Atoms and Ions





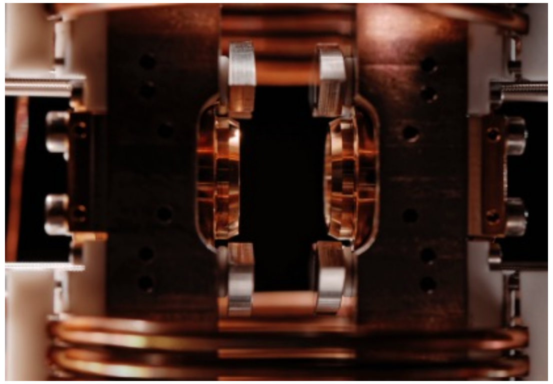
Time qubits



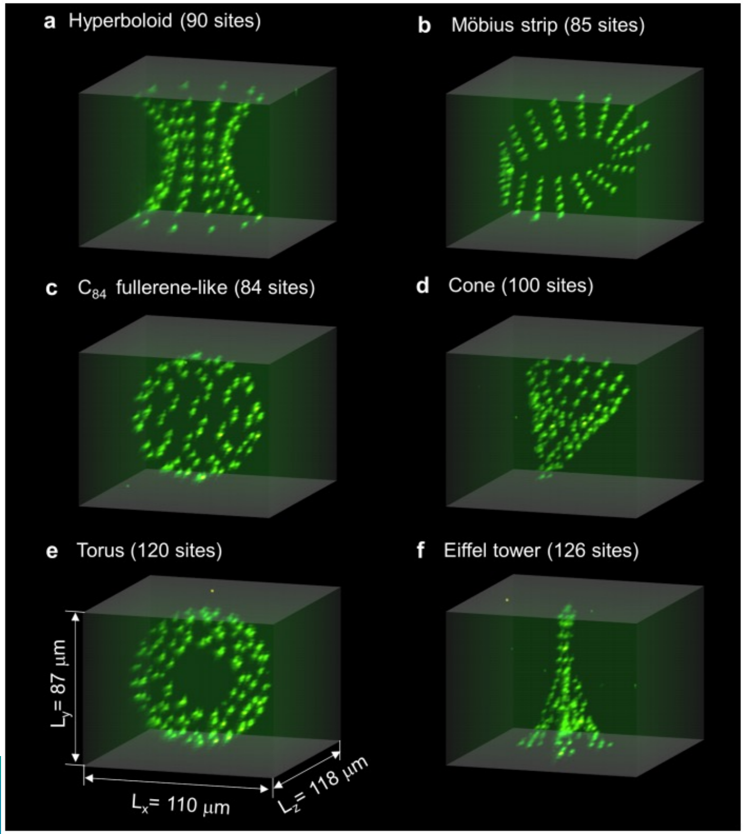
See Institute of Quantum Computing, U. of Waterloo, https://uwaterloo.ca/institute-for-quantum-computing/quantum-101/quantum-information-science-and-technology/what-qubit#Spin

Neutral atom arrays

- Configurable arrays of **single neutral atoms**
- 2 energy levels represent the qubit states
- Use **lasers** to control position and the state of the atom
  - assemble and read-out registers made of **hundreds of qubits**
  - **fully programmable quantum processing**
- **High connectivity**
- Specific computation cycle because the **register is not permanently built**
  - register preparation
  - quantum processing
  - register readout



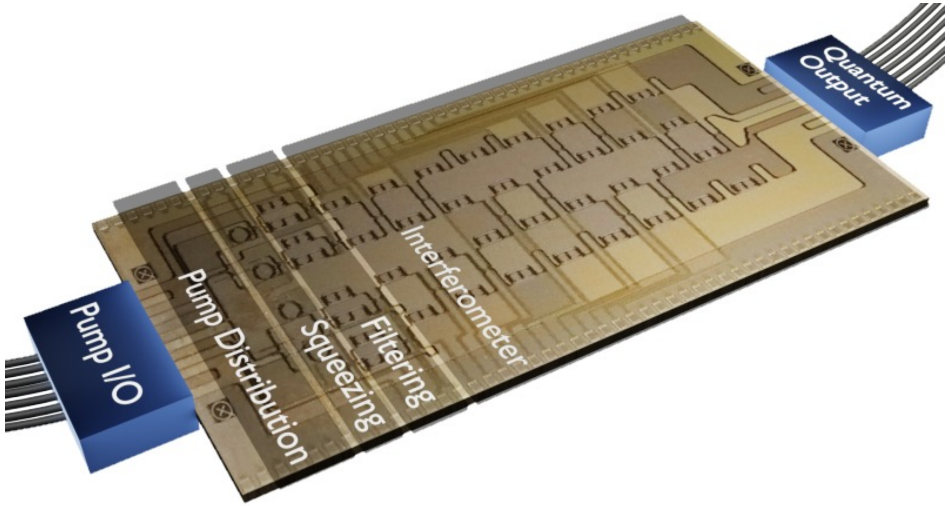
D. Barredo *et al.*, “Synthetic three-dimensional atomic structures assembled atom by atom.” arXiv:1712.02727, 2017.



PASQAL

Photonic based quantum computers

- Quantum superposition of different number of **photons in a resonator** generated by laser pulses (squeezed states)
- Set of quantum gates is implemented in a **interferometer network** (phase shifters and beam splitters)
- Photons are detected during the readout stage by **superconducting counters**
- Naturally represent **continuous variables**



https://strawberryfields.ai/photonics/hardware/details.html
https://youtu.be/v7iAqcFCTQQ

XANADU

Qubit representation

- **Dirac notation** is used to describe quantum states

Given a basis of orthogonal vectors

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

And a 2-dimensional **vector** in complex space

$$\alpha, \beta \in C^2 \qquad |\alpha|^2 + |\beta|^2 = 1$$

A quantum state is represented as

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

The **Bloch Sphere**

$$\vec{r} = \begin{bmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{bmatrix}$$

$$|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}\,|1\rangle$$

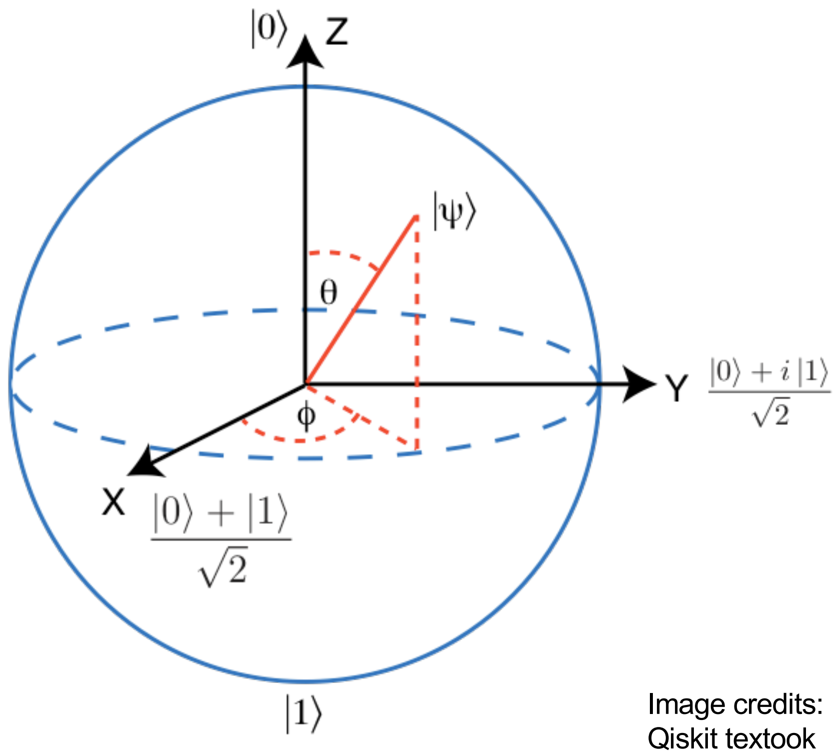


Image credits:
Qiskit textook

Quantum Gates

- Evolution of isolated quantum states follow **Schrodinger equation**
- Operations on qubits are **unitary** matrices describing state evolution
  - **Reversible operations**
  - Input and output states have the **same dimension**
  - Some classical gates (or , and, nand, xor...) **cannot be implemented directly**
  - Can **simulate** any classical computation with small overhead

$$H(t)|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$

$$UU^\dagger = U^\dagger U = I$$

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a\alpha + b\beta \\ c\alpha + d\beta \end{pmatrix}$$

$$|(a\alpha + b\beta)|^2 + |(c\alpha + d\beta)|^2 = 1$$

Example gates

The H or Hadamard gate

- The H or Hadamard gate is defined by the (unitary) matrix

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

- Its action is

$$|0\rangle - \boxed{H} - \frac{|0\rangle+|1\rangle}{\sqrt{2}}$$

$$|1\rangle - \boxed{H} - \frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

- We usually denote

$$|+\rangle := \frac{|0\rangle+|1\rangle}{\sqrt{2}}$$

$$|-\rangle := \frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

The Z gate

- The Z gate is defined by the (unitary) matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- Its action is

$$|0\rangle - \boxed{Z} - |0\rangle$$

$$|1\rangle - \boxed{Z} - -|1\rangle$$

The X or NOT gate

- The X gate is defined by the (unitary) matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

- Its action (in quantum circuit notation) is

$$|0\rangle - \boxed{X} - |1\rangle$$

$$|1\rangle - \boxed{X} - |0\rangle$$

that is, it acts like the classical NOT gate

- On a general qubit its action is

$$\alpha|0\rangle + \beta|1\rangle - \boxed{X} - \beta|0\rangle + \alpha|1\rangle$$

Other important gates

- Y gate

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

- S gate

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix}$$

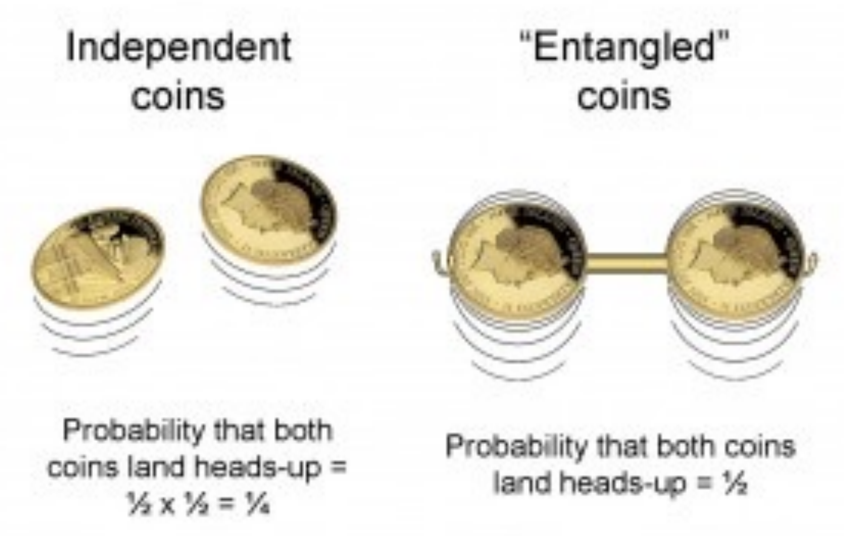
- T gate

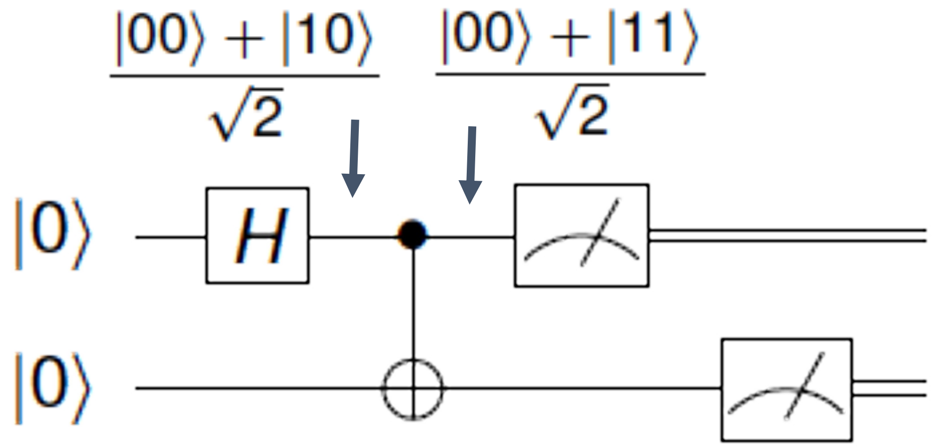
$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$$

- The gates X, Y and Z are also called, together with the identity, the Pauli gates. An alternative notation is σ_X, σ_Y, σ_Z.

Quantum entanglement

- **Quantum entanglement** creates correlation between qubit that, classically, would be independent
- Example : Bell state





Quantum circuits

Classical circuits combine **logical operations** (and, or, not, nand, and xor).

Quantum circuits use reversible gates that change the quantum states of **one, two , or more qubits**.

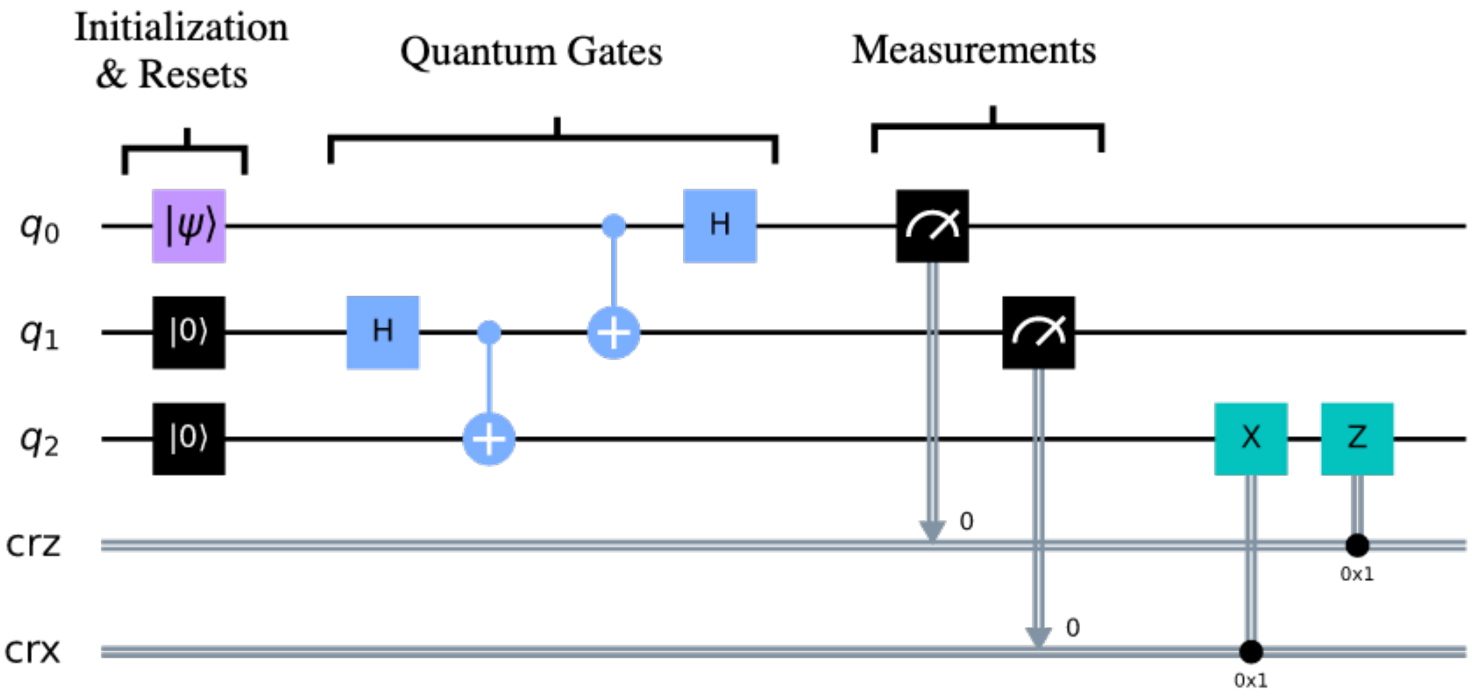
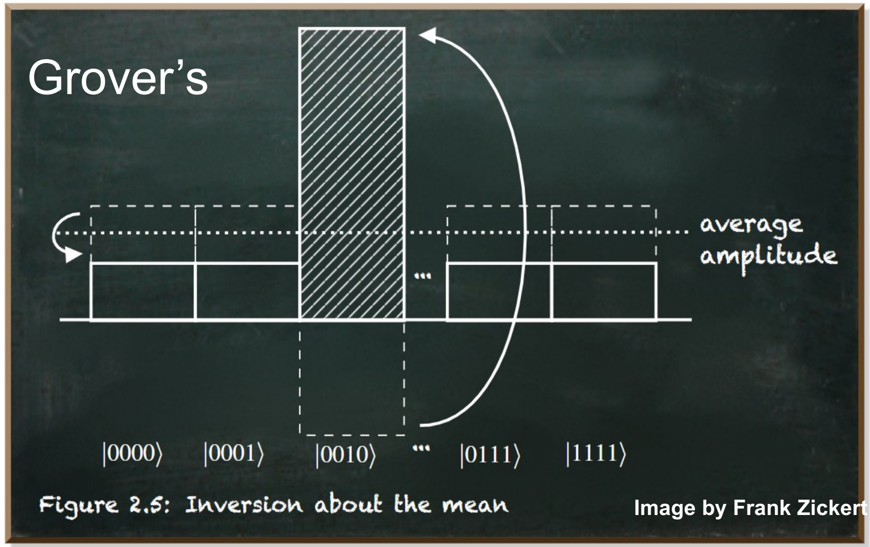


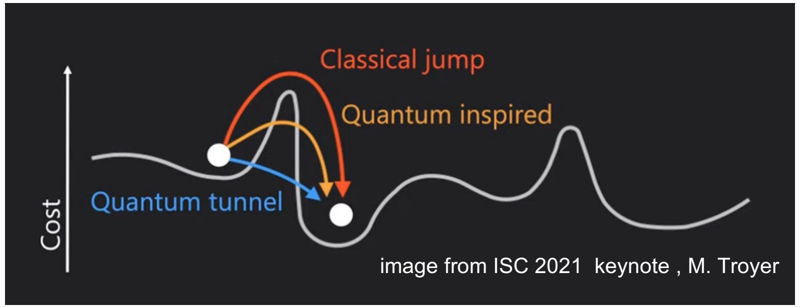
Image credits: Qiskit Textbook

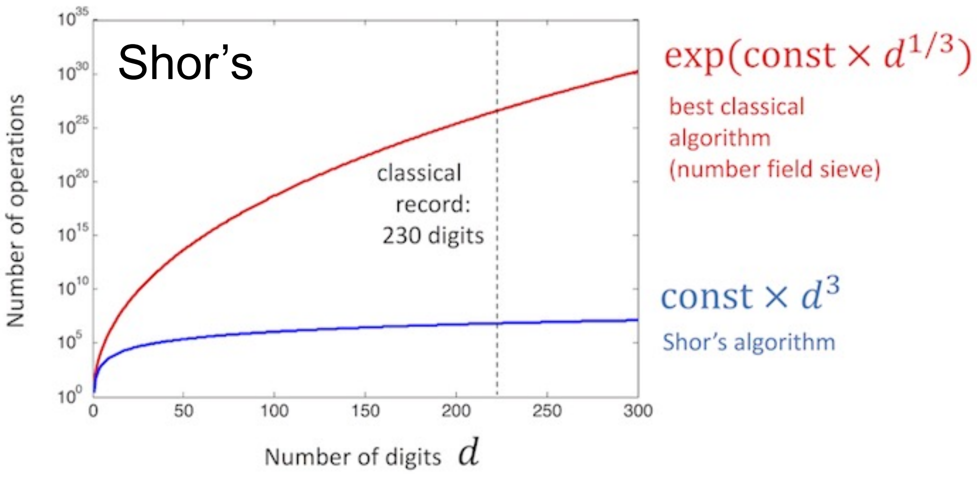
Quantum Algorithms

A collection on **http://quantumalgorithmzoo.org**

- Multiple algorithms have been studied
  - Shor algorithm for **prime factorization**
  - Grover algorithm for unsorted DB **searches**
  - Quantum **Fourier Transform**
  - …
- Quantum-inspired algorithms (emulate quantum effects on classical hardware)
- Quantum Machine Learning
- Challenge is re-thinking **algorithms design** and define fair **benchmarking** and **comparison** to classical algorithms



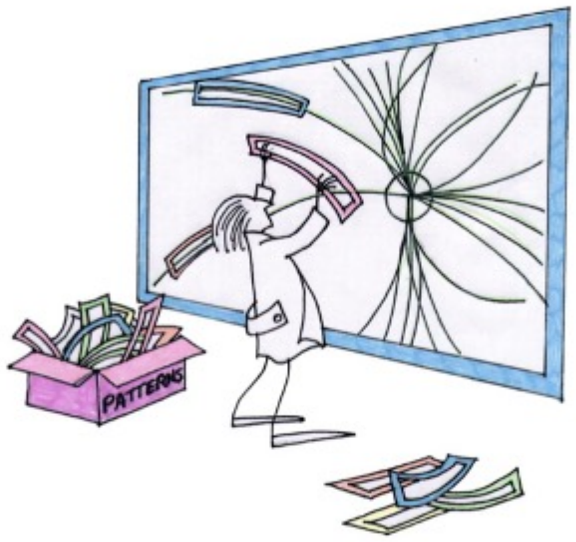


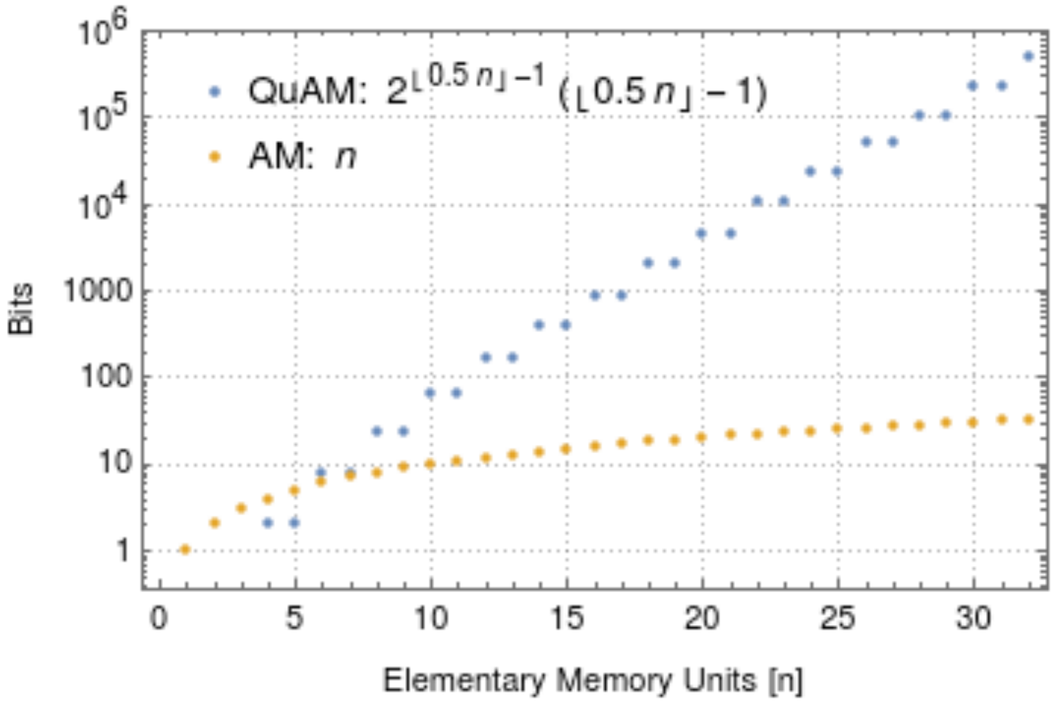


https://quantum-computing.ibm.com/composer/docs/iqx/guide/shors-algorithm

Grover algorithm for pattern recognition

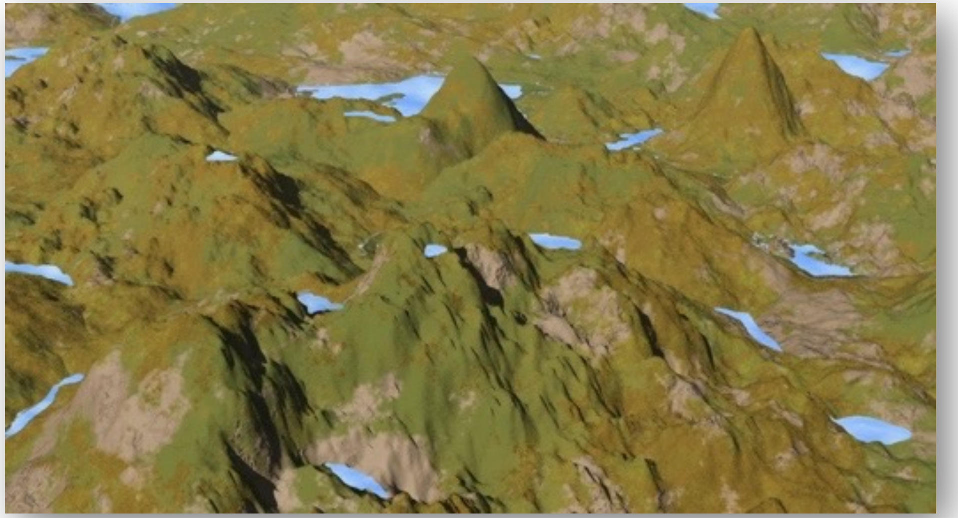
Quantum Associative Memory: Reconstruct particle trajectory by designing a DB of expected patterns and use the **generalised Grover algorithm** to match them to the detector output

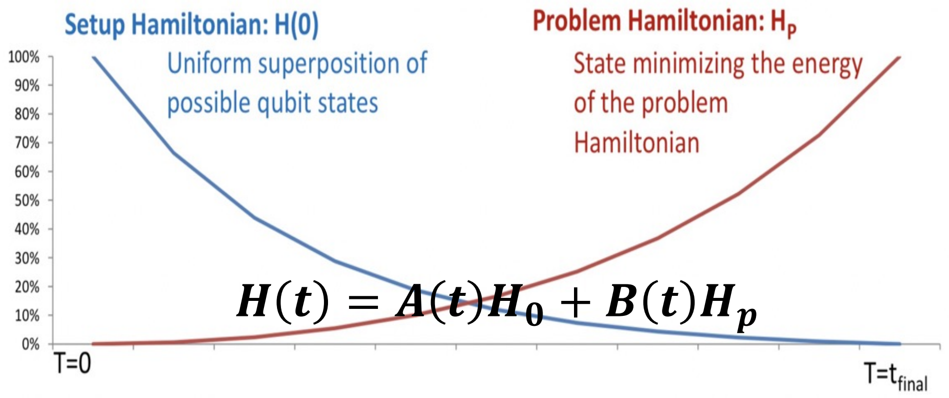




Quantum Annealing

- Annealing for optimization problems
  - PDF as a **mountain landscape**
  - Smoothly evolve probability of being at any given coordinate with time.
  - Probability increases around the coordinates of deep valleys
- Quantum systems based on **superconducting qubits**
- **D-Wave Advantage**: 5436 qubits - 15 connection (Pegasus)
  - **Quantum superposition:** scan simultaneously multiple coordinates
  - **Quantum tunneling:** reduces risk of local minima (tunnel through hills)
  - **Quantum entanglement:** discover correlations between the coordinates that lead to deep valleys.





T. Kadowaki and H. Nishimori, PRE, 58(5), pp. 5355-5363, (1998)
E. Farhi, et al., Science 292, 472 (2001)
W. Kaminsky, S. Lloyd, T. Orlando, arXiv:quant-ph/0403090

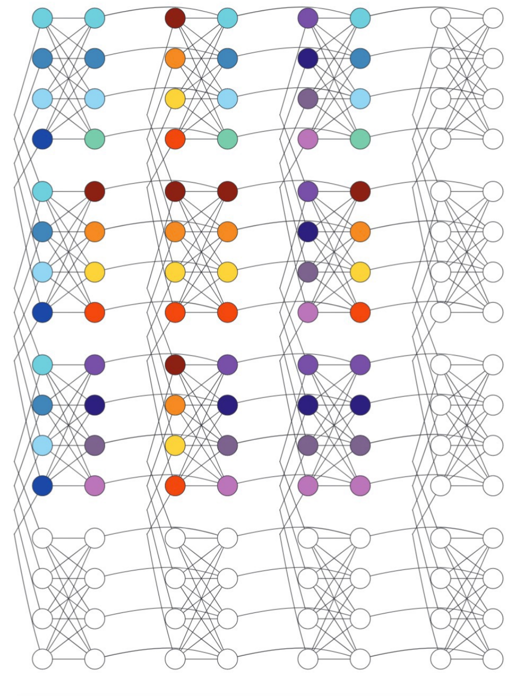
Training a classifier with QA

$$H_{\text{Ising}} = \sum_i h_i \sigma_i^z + \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z$$

Adjacent qubits

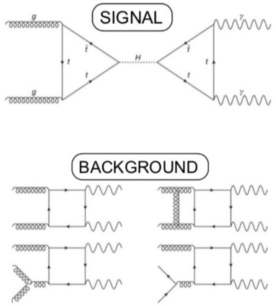
- Map the problem to a **Ising model** (spin lattice as qubit graph)
- Define Hamiltonian and **train by minimizing energy**
- First QA application to High Energy Physics

https://arxiv.org/abs/1210.8395

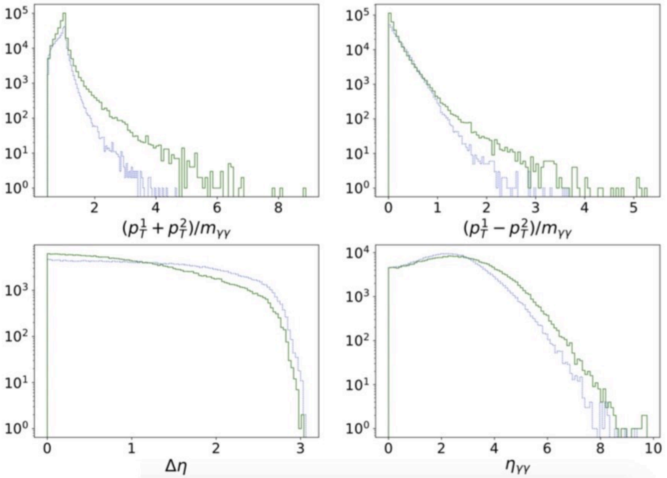


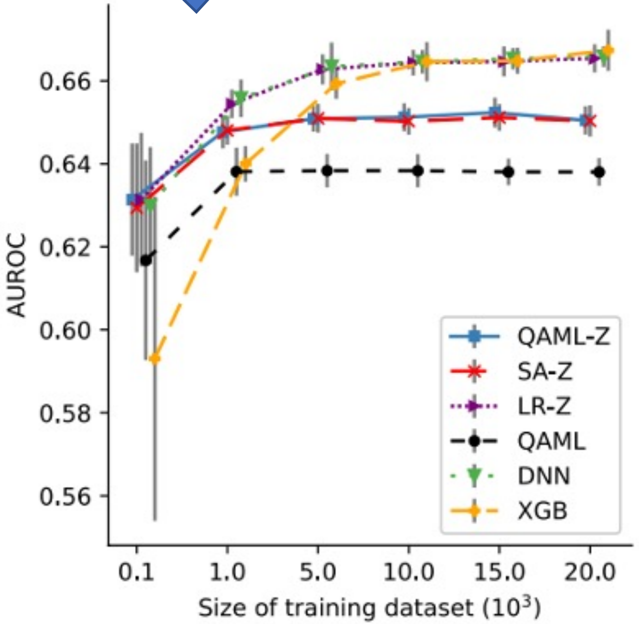
(a) K_{13} embedded in upper 3×3 sub-grid

Initial features



weak classifiers = Ising spins





Alexander Zlokapa, Alex Mott, Joshua Job, Jean-Roch Vlimant, Daniel Lidar, and Maria Spiropulu. **Quantum adiabatic machine learning by zooming into a region of the energy surface.** Physical Review A, 102:062405, 2020. DOI:10.1103/PhysRevA.102.062405.

Today’s challenges

- **Noisy Intermediate-Scale Quantum** devices
    - Limitations in terms of **stability** and **connectivity**
    - **De-coherence**, measurement errors or gate level errors (**noise**)
        - Specific **error mitigation techniques**
        - **Circuit optimisation**
        - Prefer algorithms **robust against noise**
- Quantum computers initially integrated in **hybrid quantum-classical infrastructure**
    - Engineering, cooling, I/O
    - Hybrid algorithms, QPU as accelerators

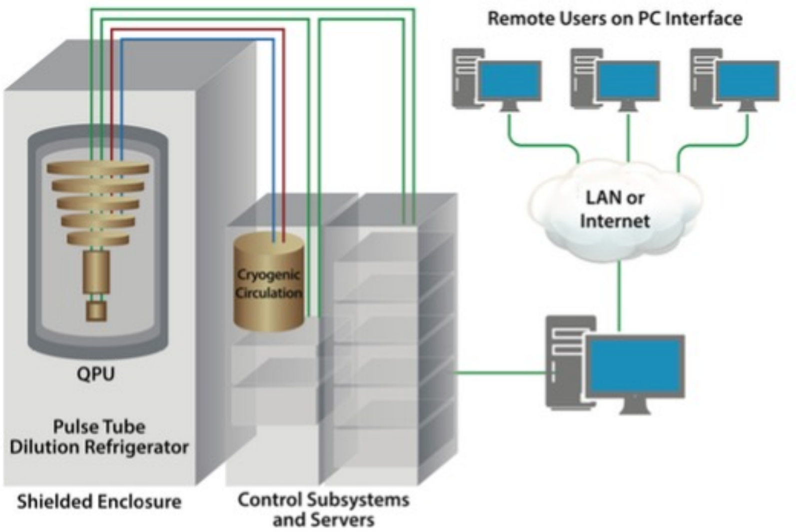
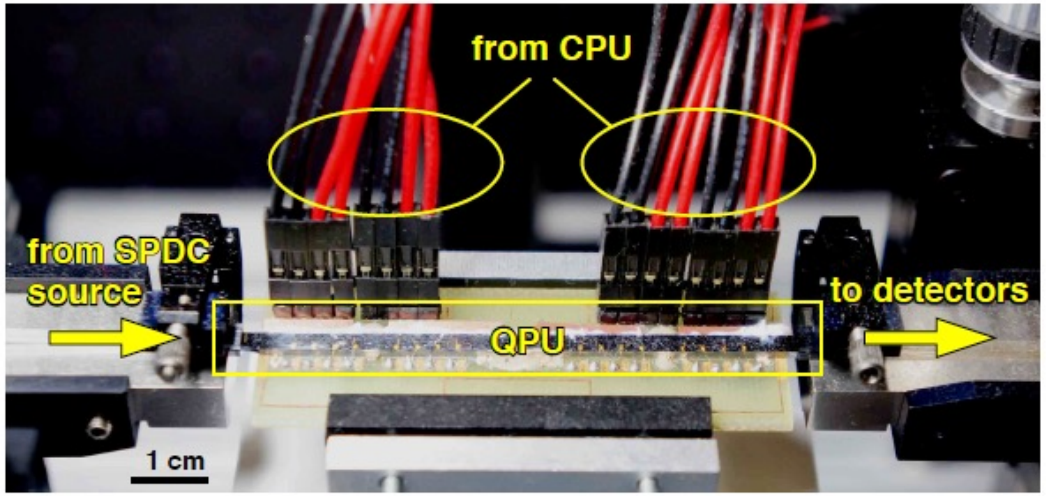


Image: D-Wave tutorial

Peruzzo, A. "A variational eigenvalue solver on a quantum processor. eprint." *arXiv preprint arXiv:1304.3061* (2013).



Development Roadmap | Executed by IBM On target

IBM **Quantum**

2019	2020	2021	2022	2023	2024	2025	Beyond 2026
Run quantum circuits on the IBM cloud	Demonstrate and prototype quantum algorithms and applications	Run quantum programs 100x faster with Qiskit Runtime	Bring dynamic circuits to Qiskit Runtime to unlock more computations	Enhancing applications with elastic computing and parallelization of Qiskit Runtime	Improve accuracy of Qiskit Runtime with scalable error mitigation	Scale quantum applications with circuit knitting toolbox controlling Qiskit Runtime	Increase accuracy and speed of quantum workflows with integration of error correction into Qiskit Runtime

Model Developers

- Prototype quantum software applications → Quantum software applications
  - Machine learning | Natural science | Optimization

Algorithm Developers

- Quantum algorithm and application modules
  - Machine learning | Natural science | Optimization
- Quantum Serverless
  - Intelligent orchestration
  - Circuit Knitting Toolbox
  - Circuit libraries

Kernel Developers

- Circuits
- Qiskit Runtime
  - Dynamic circuits
  - Threaded primitives
  - Error suppression and mitigation
  - Error correction

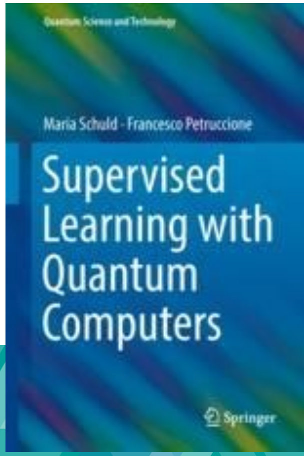
System Modularity

Falcon	27 qubits
Hummingbird	65 qubits
Eagle	127 qubits
Osprey	433 qubits
Condor	1,121 qubits
Flamingo	1,386+ qubits
Kookaburra	4,158+ qubits
Scaling to	10K-100K qubits with classical and quantum communication
Heron	133 qubits x p
Crossbill	408 qubits

Quantum Machine Learning

QML tutorials and resources https://pennylane.ai

Supervised Learning with Quantum Computers
Maria Schuld
Francesco Petruccione



Quantum Machine Learning

Use **Quantum Computing** to accelerate **ML/DL.**

Quantum circuits are **differentiable** and can be trained **minimizing a cost function** dependent on training data:

1. **Feature extraction** and **data encoding**
   - How to represent classical data in quantum states?
2. **Model definition** (kernel based or variational)
   - Design wrt data
3. **Optimisation** and **convergence in Hilbert space**
   - **Convergence vs expressivity**
   - Barren plateau and vanishing gradients
   - Gradient-free or gradient-based optimisers
   - …

Different tools can enable hybrid computations

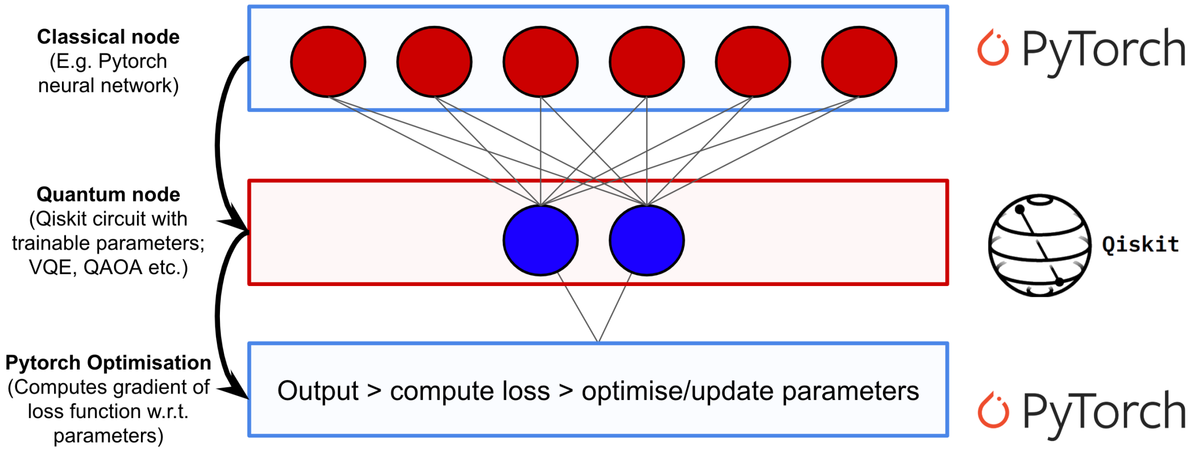
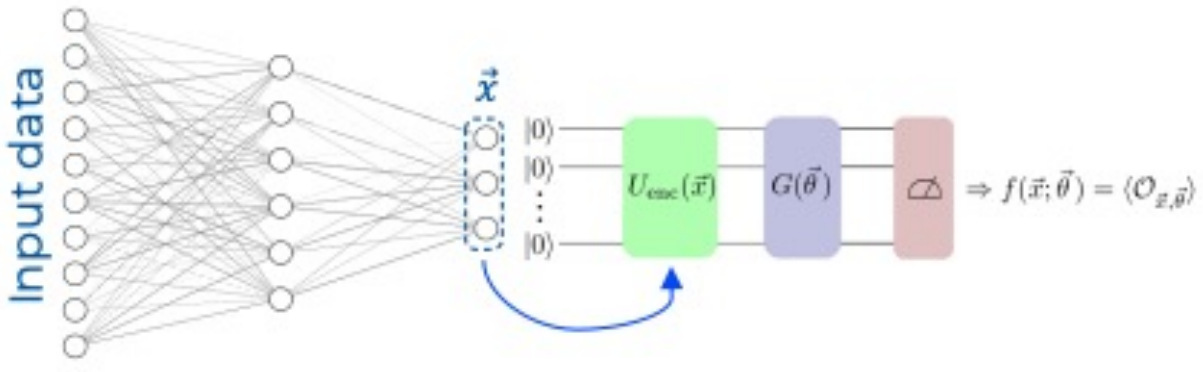


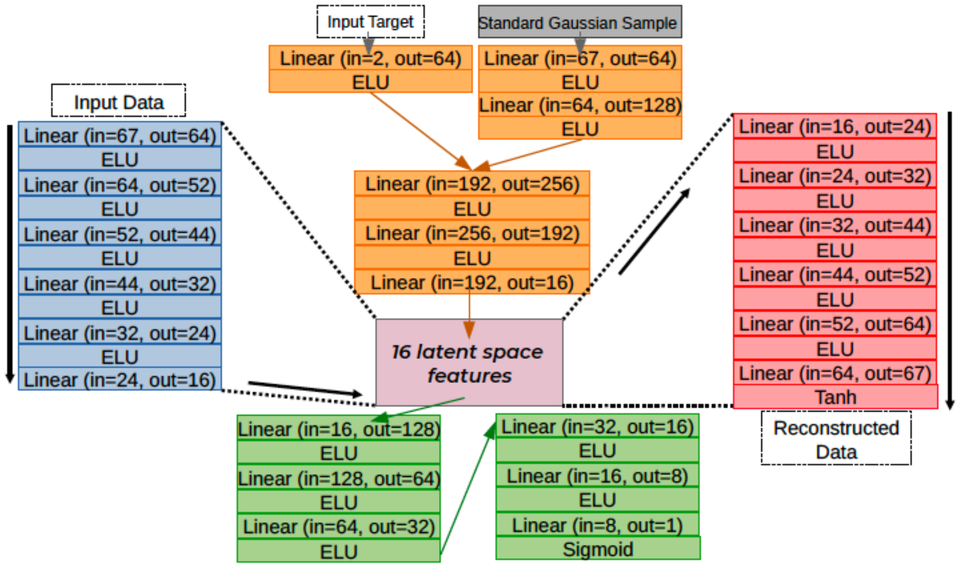
Image credit Qiskit.org/textbook

Dimensionality reduction and feature extraction

Dimensionality reduction/feature extraction

- Reduce size of classical data
- Optimize input (PCA, Auto-Encoders..)
- **Pre-trained or co-trained** in hybrid setup





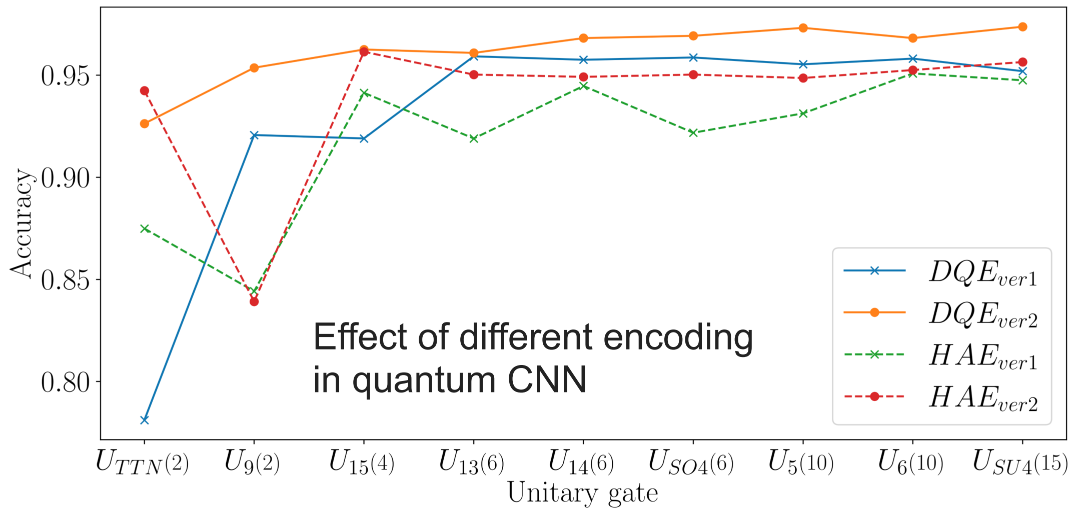
Belis, Vasilis, et al. "**Higgs analysis with quantum classifiers**." *EPJ Web of Conferences*. Vol. 251. EDP Sciences, 2021.

Feature selection + Model	AUC
AUC + QSVM	0.66 ± 0.01
PyTorch AE + QSVM	0.62 ± 0.03
AUC + SVM rbf	0.65 ± 0.01
PyTorch AE + SVM rbf	0.62 ± 0.02
KMeans + SVM rbf	0.61 ± 0.02

Patrick Odagiu, 2021 : End-to-end Sinkhorn autoencoder with a classifier NN (green).
Sinkhorn part cosists of an encoder (blue), decoder (red) and noise generator (orange).

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

Quantum embedding



Data embedding in quantum states :
compromise between exponential compression and circuit depth

In some cases: **data re-uploading**

1) Amplitude Encoding

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

\+ Exponential compression

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

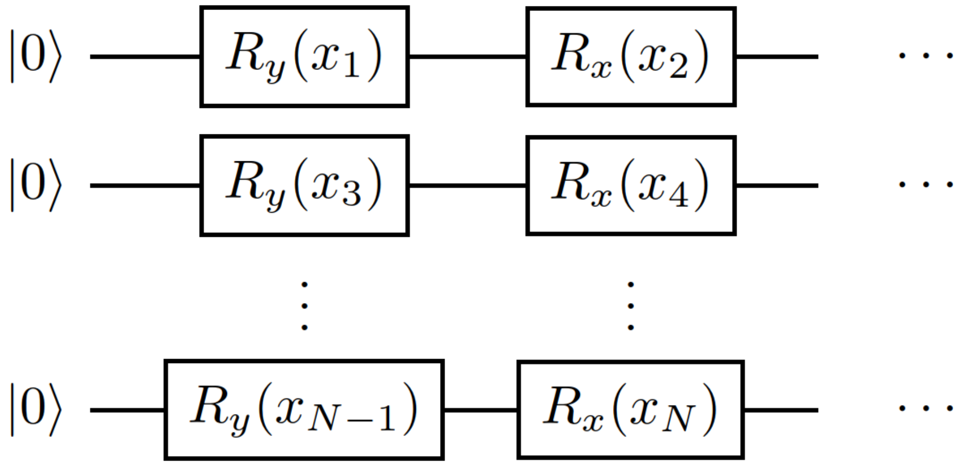
− Polynomial number of gates

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

Quantum Embedding

2) Dense Qubit Encoding

$$|\phi(x)\rangle = \bigotimes_{j=1}^{\frac{N}{2}} \left(e^{-i\frac{x_{\frac{N}{2}+j}}{2}\sigma_x} e^{-i\frac{x_j}{2}\sigma_y}\right)$$



+ Easy to implement

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

− No compression in resources

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

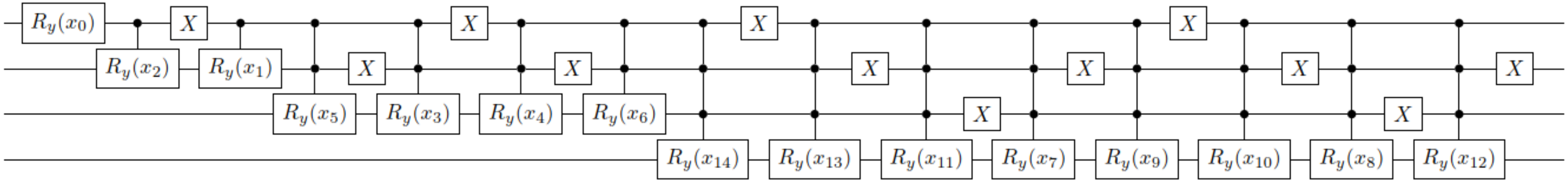
Quantum Embedding

3) Hybrid Angle Encoding

$$|\phi(x)\rangle = \bigotimes_{k=1}^{b} \left(\sum_{i=1}^{2^m} \prod_{j=0}^{m-1} \cos^{1-i_j}(x_{g(j),k}) \sin^{i_j}(x_{g(j),k}) |i\rangle_k \right)$$

Encode $b \times 2^m$ values into $b \times m$ qubits

- Compromise between Amplitude and Qubit Encoding
- Still requires too many two qubit gates



Model definition

Kernel methods

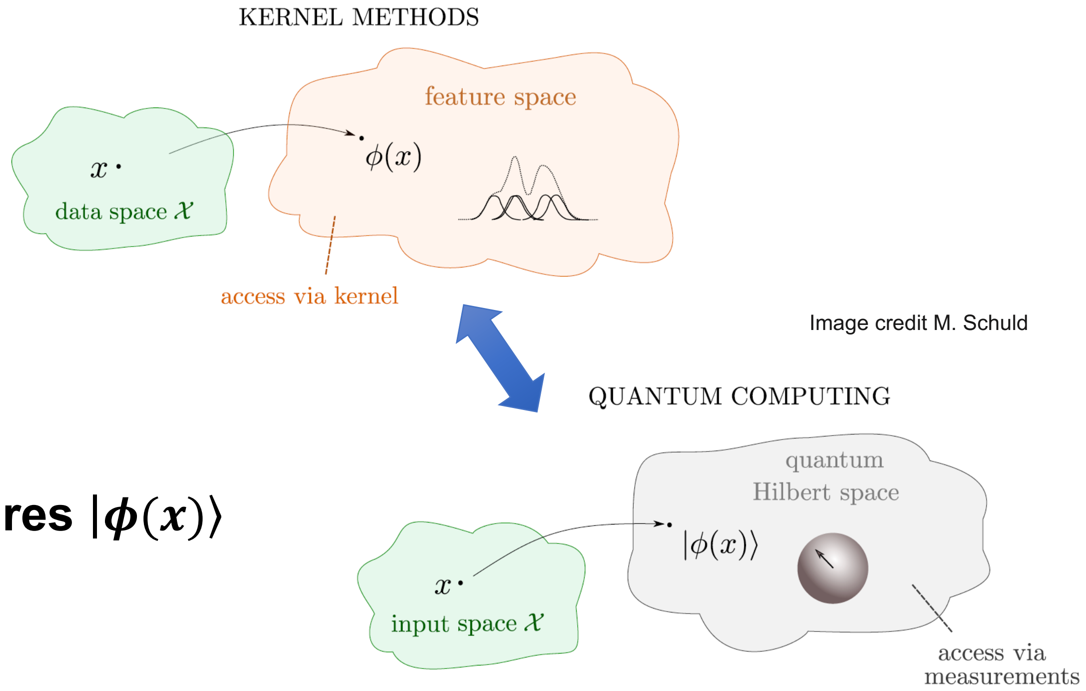


Image credit M. Schuld

Feature maps as **quantum kernels**

Use quantum computers to create **classically intractable features** $|\boldsymbol{\phi}(\boldsymbol{x})\rangle$

- Build inner product of feature vectors → $O(N_{data}^2)$
- Use classical **kernel-based training**
  - **Convex** losses**, global** minimum
- Identify classes of kernels that relate to specific data **structures**[1]
- Given a variational circuit of the form $U(x,\vartheta) = \mathcal{V}_\vartheta U_\phi(x)$, can define a quantum kernel method with better accuracy: $|\boldsymbol{\phi}(\boldsymbol{x})\rangle = U_\phi(x)|0\rangle$
- **Classically: not all machine learning models can be described by kernel methods.**

Schuld, Maria. "Supervised quantum machine learning models are kernel methods." *arXiv preprint arXiv:2101.11020* (2021).

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

Quantum Support Vector Machine

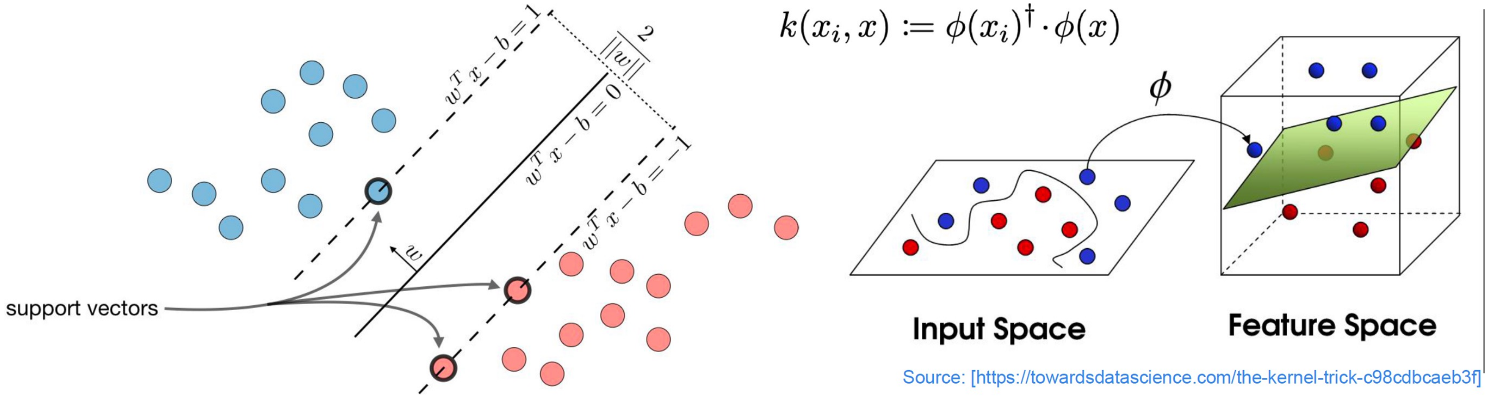
SVM are **kernel methods**:

Trained to find the optimal separating plane

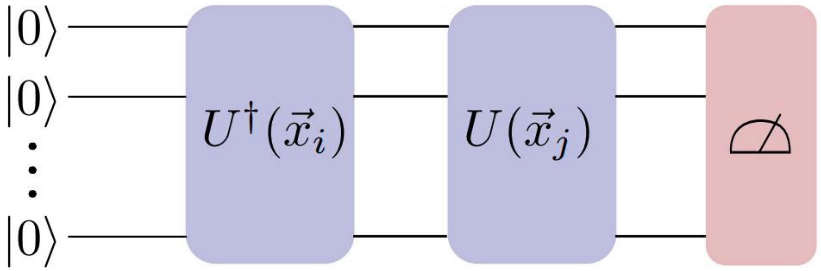
Quantum SVM use feature maps as kernels

Feature maps enable SVM to design non-linear decision boundaries

Feature maps in high dimensionality space improve separation power



Source: [https://towardsdatascience.com/the-kernel-trick-c98cdbcaeb3f]



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

NB:

- Quantum kernels sampled on quantum device
- Minimisation step is classical

Model definition

Variational algorithms

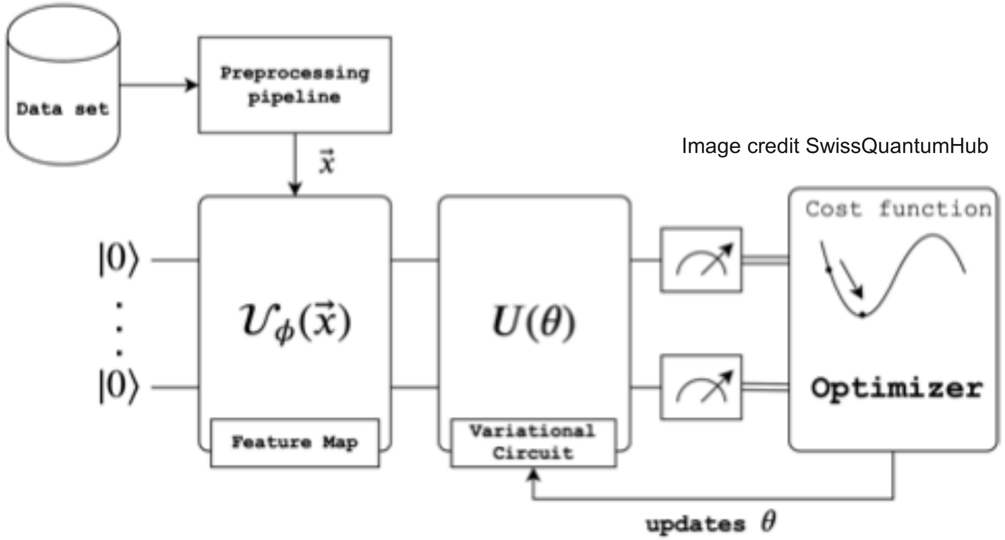


Image credit SwissQuantumHub

Define a **parametric quantum circuit** with trainable parameters ϑ

$$U(x,\vartheta)$$

Given an observable O, build a model

$$y(x,\vartheta) = \langle 0 | U^\dagger(x,\vartheta) O U(x,\vartheta) | 0 \rangle$$

- **Trained using gradient-free** or **gradient-based** optimization in a classical loop
    - Backpropagation and auto-differentiation
- **Data Embedding** $\mathcal{V}_\phi(x)$ can be **learned**
- Improve performance by designing architectures to **leverage data symmetries**[1]
- There are quantum circuits that **hard to simulate classically**

[1] Bogatskiy, Alexander, et al. "**Lorentz group equivariant neural network for particle physics.**" *International Conference on Machine Learning*. PMLR, 2020.

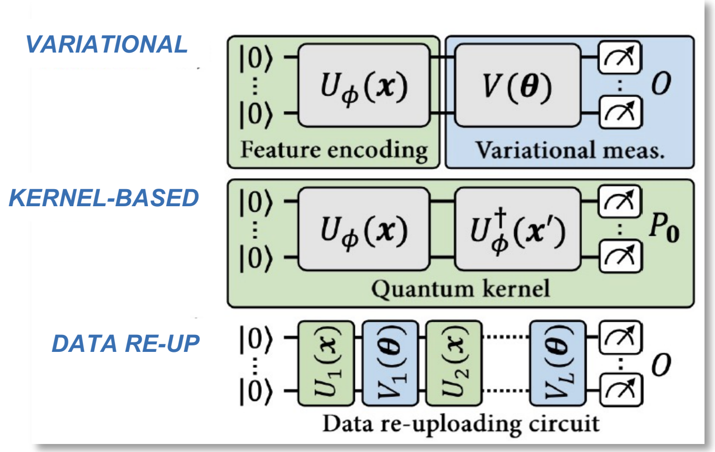
Equivalent interpretations?

Characterize the behaviour of different models**, similarity and links among them** and **link to data properties.**

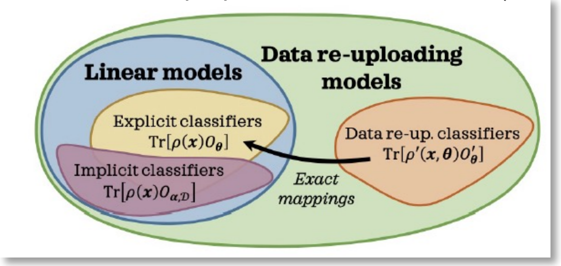
Ex:

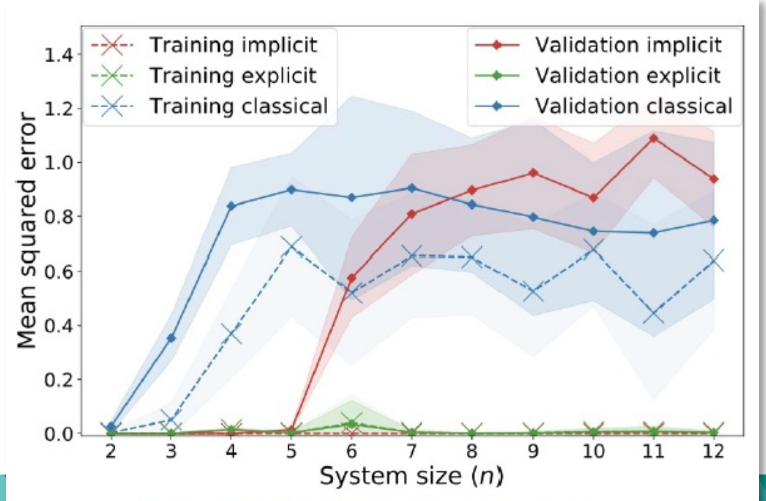
- Data Re-Uploading circuits: alternating data encoding and variational layers.
    - Represented as **explicit linear models** (variational) in larger feature space

    → can be reformulated as **implicit models** (kernel)
- Representer theorem: implicit models achieve **better accuracy**
    - Explicit models exhibit **better generalization** performance



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





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

See M. Grossi summary at the 2022 CERN Openlab Technical Workshop : https://indico.cern.ch/event/1100904/contributions/4775169/

Model Convergence and Barren Plateau

Given the size of the Hilbert space a compromise between **expressivity**, **convergence** and **generalization** performance is needed.

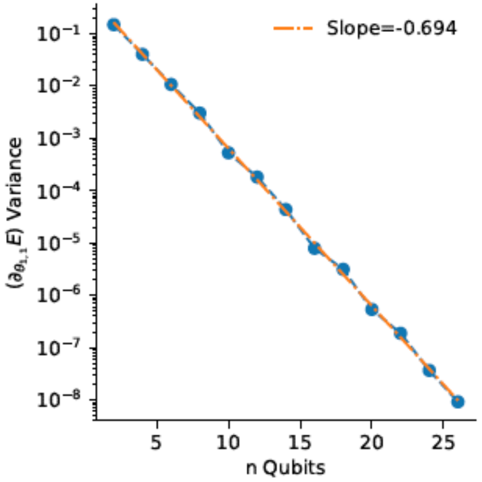
Classical gradients **vanish exponentially** with the number of layers (J. McClean *et al.*, arXiv:1803.11173)

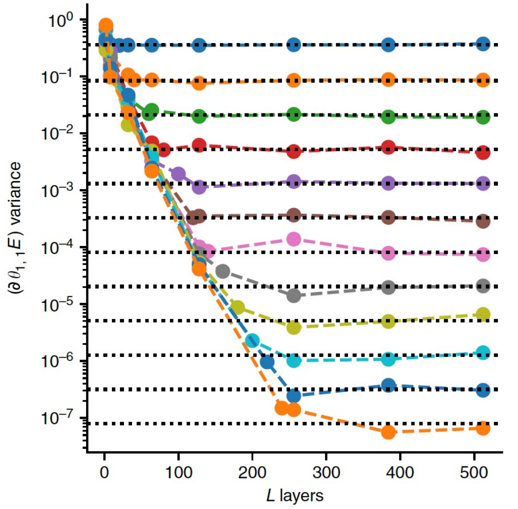
- Convergence still possible if gradients consistent between batches.

Quantum gradient decay exponentially in the number of qubits

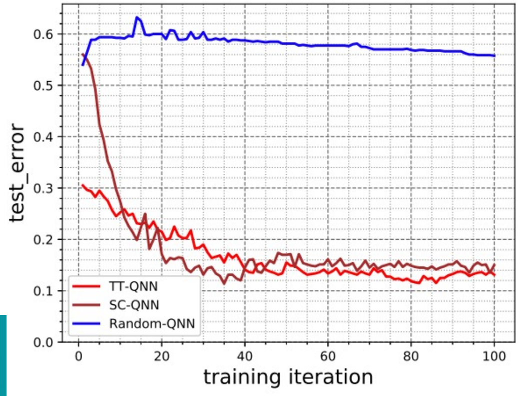
- Random circuit initialization
- Loss function locality in shallow circuits (M. Cerezo *et al.*, arXiv:2001.00550)
- Ansatz choice: TTN, CNN (Zhang *et al.*, arXiv:2011.06258, A Pesah, *et al.*, *Physical Review X* 11.4 (2021): 041011.)
- Noise induced barren plateau (Wang, S *et al.*, Nat Commun 12,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

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

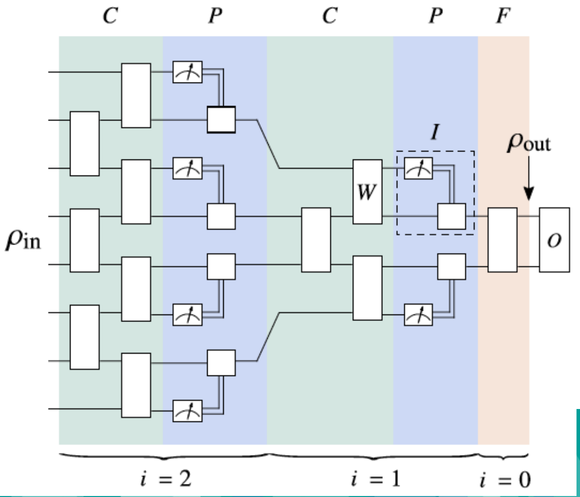




TTN for MNIST classification (8 qubits), Zhang *et al.*, arXiv:2011.06258

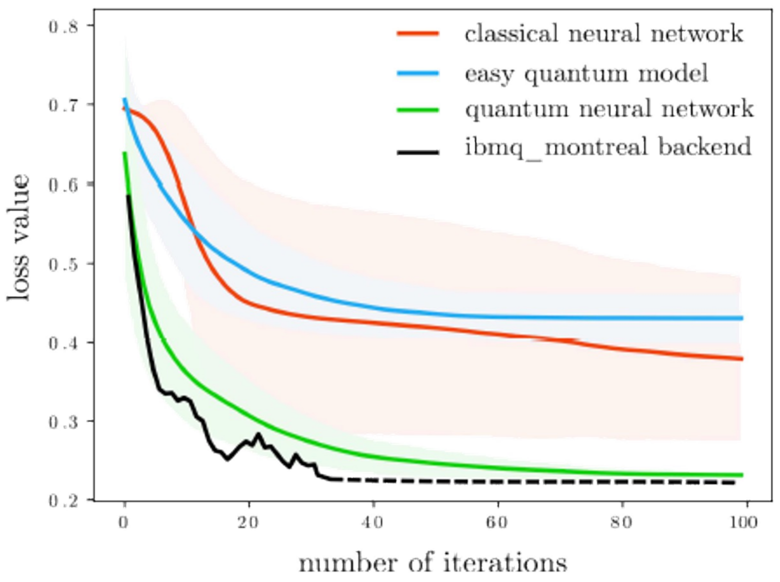


QCNN: A Pesah, *et al.*, *Physical Review X* 11.4 (2021): 041011



Defining quantum Advantage for QML

Abbas, Amira, et al. "**The power of quantum neural networks.**" *Nature Computational Science* 1.6 (2021): 403-409.



Different possible definitions

Runtime speedup

Sample complexity

Representational power

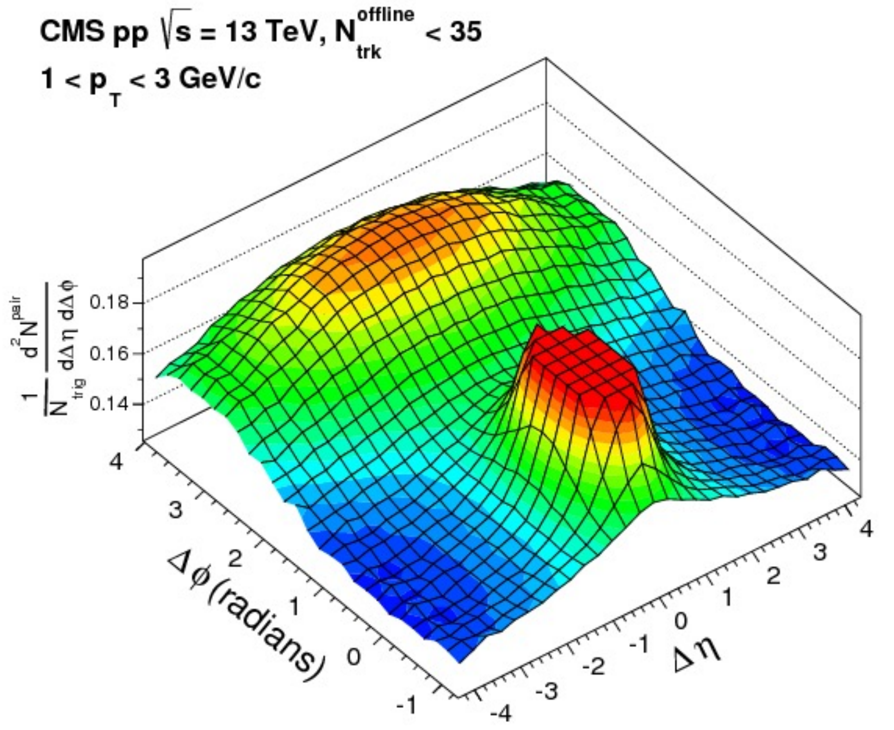
Classical Intractability: a quantum algorithm that cannot be efficiently simulated classically

- No established recipe for classical data
- Need to use the whole exponential advantage in Hilbert space, but will it converge ?

(Algorithm expressivity vs convergence and generalization)

Kübler, Jonas, Simon Buchholz, and Bernhard Schölkopf. "**The inductive bias of quantum kernels.**" Advances in Neural Information Processing Systems 34 (2021).
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

Practical advantage



Khachatryan, Vardan, et al. "**Measurement of Long-Range Near-Side Two-Particle Angular Correlations in p p Collisions at s= 13 TeV**." *Physical review letters* 116.17 (2016): 172302.

Practical implementation vs asymptotic complexity

Data embedding
NISQ vs ideal quantum devices
Realistic applications

Performance metrics and fair comparison to classical models

HEP data is classical, but originally produced by quantum processes. It is these **intrinsically quantum correlations** we are trying to identify

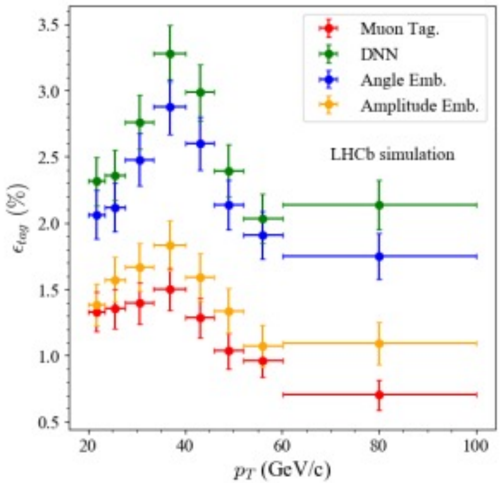
A change of paradigm could reflect in interesting insights

- What are natural building blocks for QML algorithms?
- How can we construct useful bridges between QC and learning theory?
- How can we make quantum software ready for ML applications?

Schuld, Maria, and Nathan Killoran. "**Is quantum advantage the right goal for quantum machine learning?**." *arXiv preprint arXiv:2203.01340* (2022).

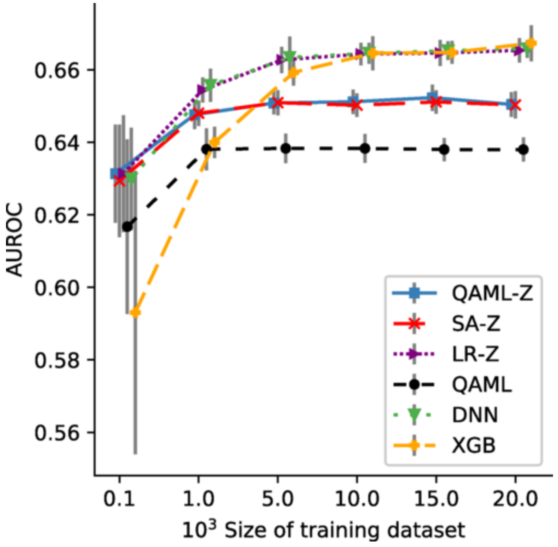
Quantum Machine Learning examples

QML in High Energy Physics

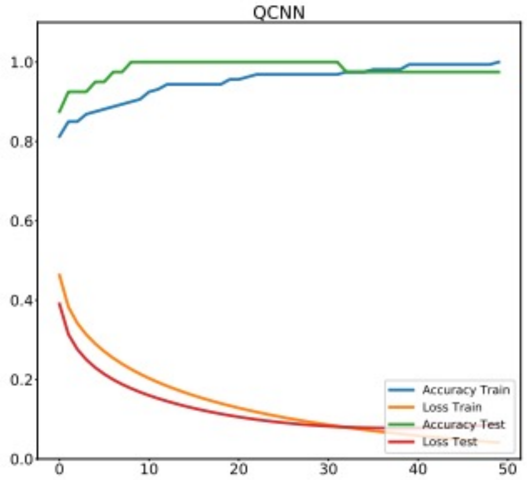


Alessio Gianelle, Patrick Koppenburg, Donatella Lucchesi, Davide Nicotra, Eduardo Rodrigues, Lorenzo Sestini, Jacco de Vries, and Davide Zuliani. **Quantum Machine Learning for b-jet identification.** arXiv:2202.13943, 2022.

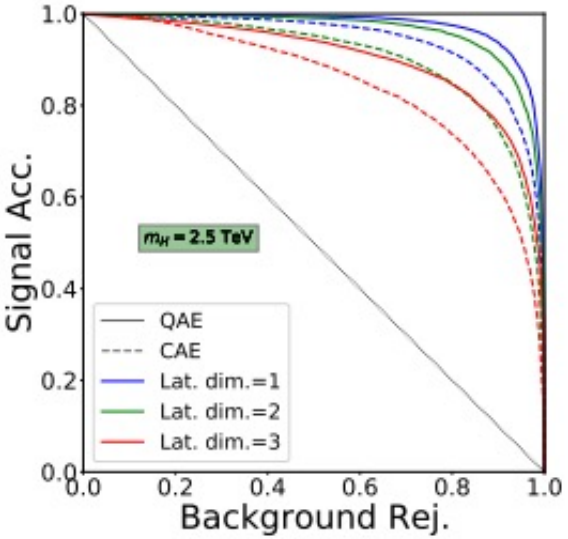
Alexander Zlokapa, Alex Mott, Joshua Job, Jean-Roch Vlimant, Daniel Lidar, and Maria Spiropulu. **Quantum adiabatic machine learning by zooming into a region of the energy surface.** Physical Review A, 102:062405, 2020. DOI:10.1103/PhysRevA.102.062405.



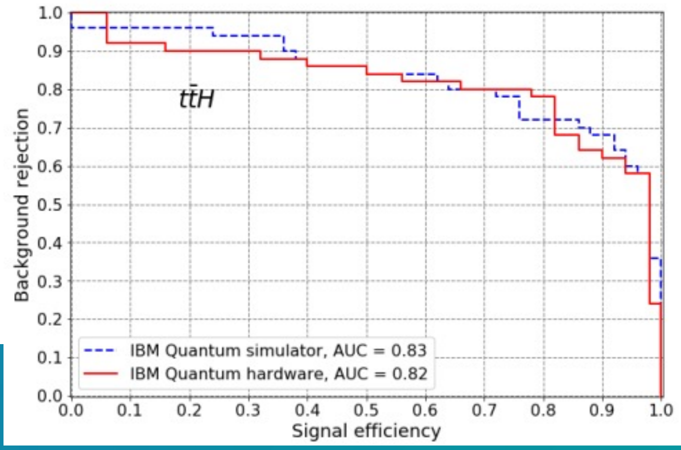
Samuel Yen-Chi Chen, Tzu-Chieh Wei, Chao Zhang, Haiwang Yu, and Shinjae Yoo. **Quantum convolutional neural networks for high energy physics data analysis.** arXiv preprint: 2012.12177, 2020.



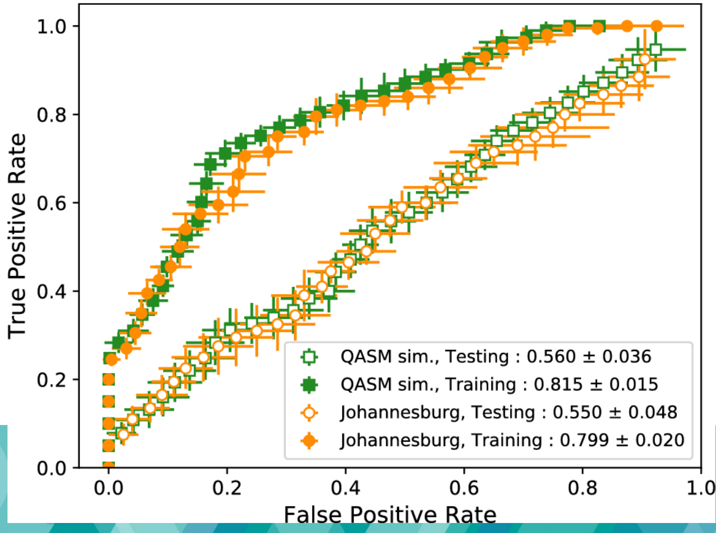
Vishal S Ngairangbam, Michael Spannowsky, and Michihisa Takeuchi. **Anomaly detection in high-energy physics using a quantum autoencoder**. arXiv preprint arXiv:2112.04958, 2021.



Sau Lan Wu, Jay Chan, Wen Guan, Shaojun Sun, Alex Wang, Chen Zhou, Miron Livny, Federico Carminati, Alberto Di Meglio, Andy C Y Li, and et al. **Application of quantum machine learning using the quantum variational classifier method to high energy physics analysis at the lhc on ibm quantum computer simulator and hardware with 10 qubits**. Journal of Physics G: Nuclear and Particle Physics, 48(12):125003, Oct 2021

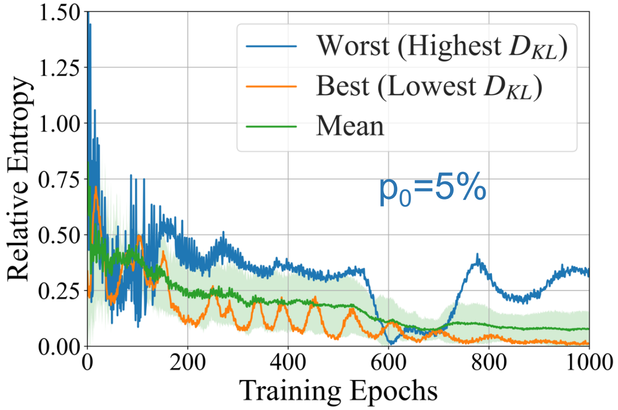


Koji Terashi, Michiru Kaneda, Tomoe Kishimoto, Masahiko Saito, Ryu Sawada, and Junichi Tanaka. **Event classification with quantum machine learning in 20 high-energy physics**. Computing and Software for Big Science, 5(1), January 2021.

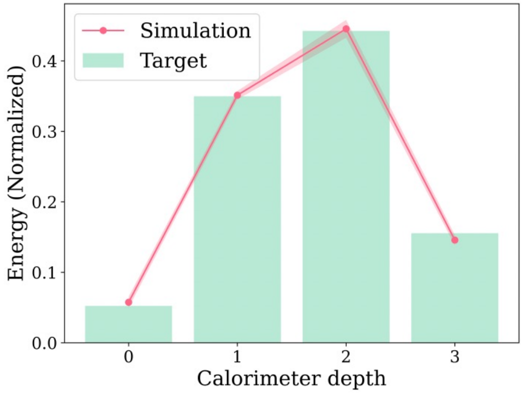


QML at CERN

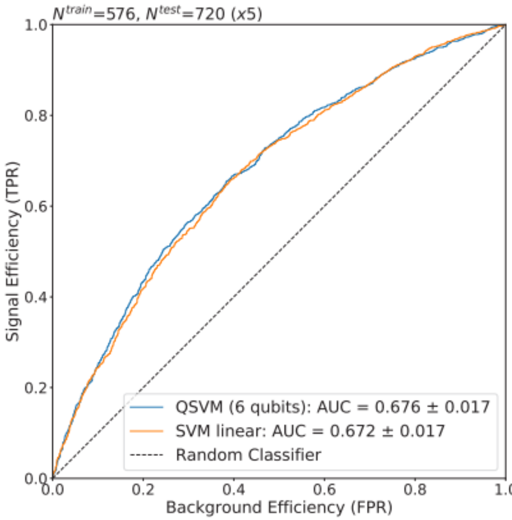
Borras, Kerstin, et al. "**Impact of quantum noise on the training of quantum Generative Adversarial Networks**." *arXiv preprint arXiv:2203.01007* (2022).



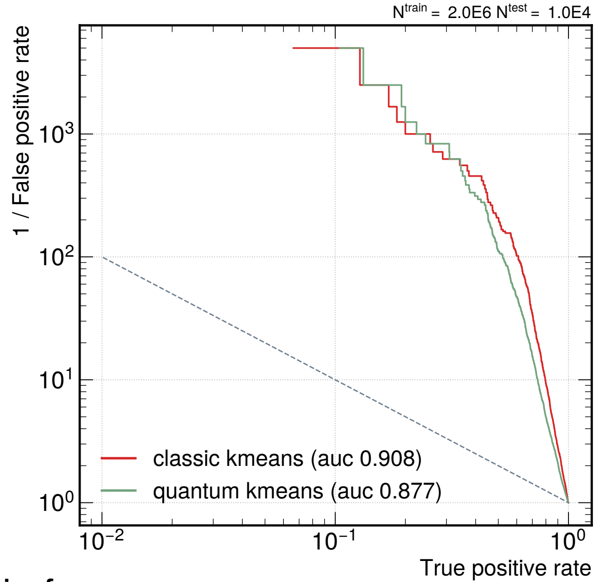
Chang S.Y. *et al.*, **Running the Dual-PQC GAN on Noisy Simulators and Real Quantum Hardware**, QTML2021, ACAT21



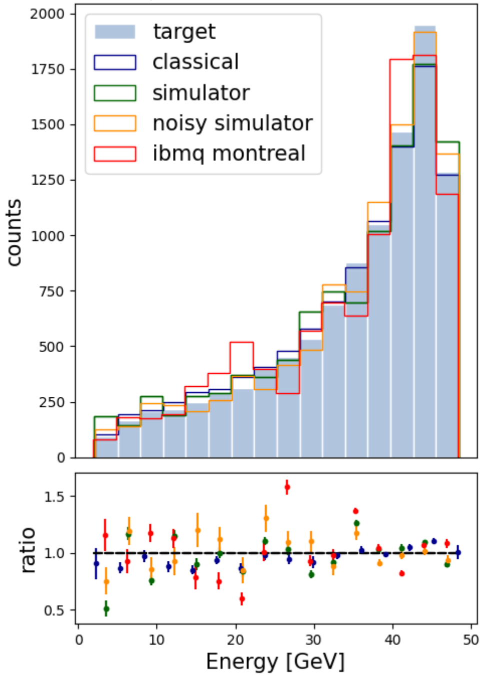
Vasilis Belis, Samuel González-Castillo, Christina Reissel, Sofia Vallecorsa, Elías F. Combarro, Günther Dissertori, and Florentin Reiter. **Higgs analysis with quantum classifi**ers. EPJ Web of Conferences, 251:03070, 2021



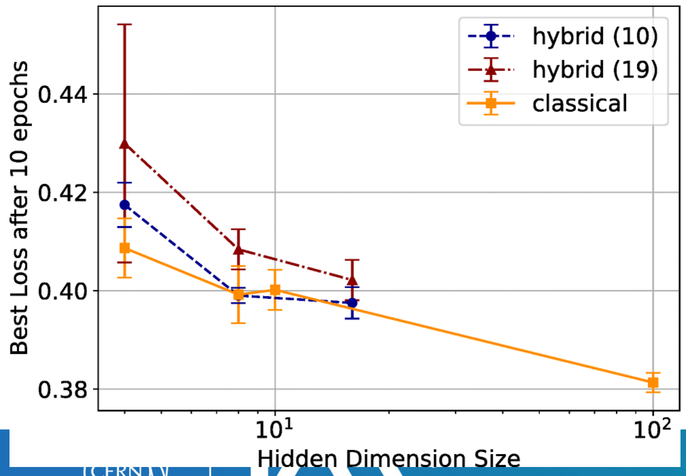
Kinga Wozniak, **Unsupervised clsutering for a Randall–Sundrum Graviton at 3.5TeV narrow resonance, 5th IML workshop, May 2022**



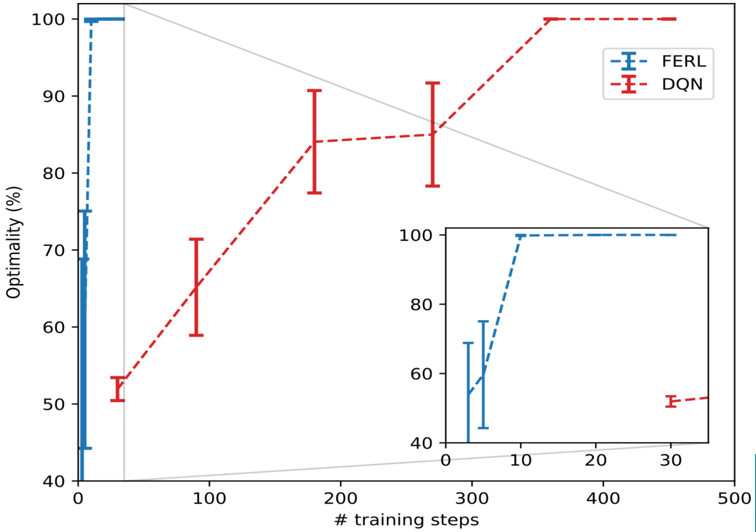
O. Kiss, **Quantum Born Machine for event generation**, ACAT2021



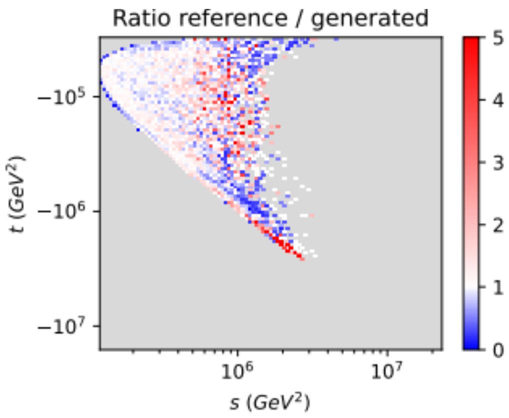
Tüysüz, Cenk, et al. "**Hybrid quantum classical graph neural networks for particle track reconstruction**." *Quantum Machine Intelligence* 3.2 (2021): 1-20.



M. Shenk, V. Kain, **Quantum Reinformcement Learning**, BQiT 2021, 2022 CERN openlab Tech Workshop



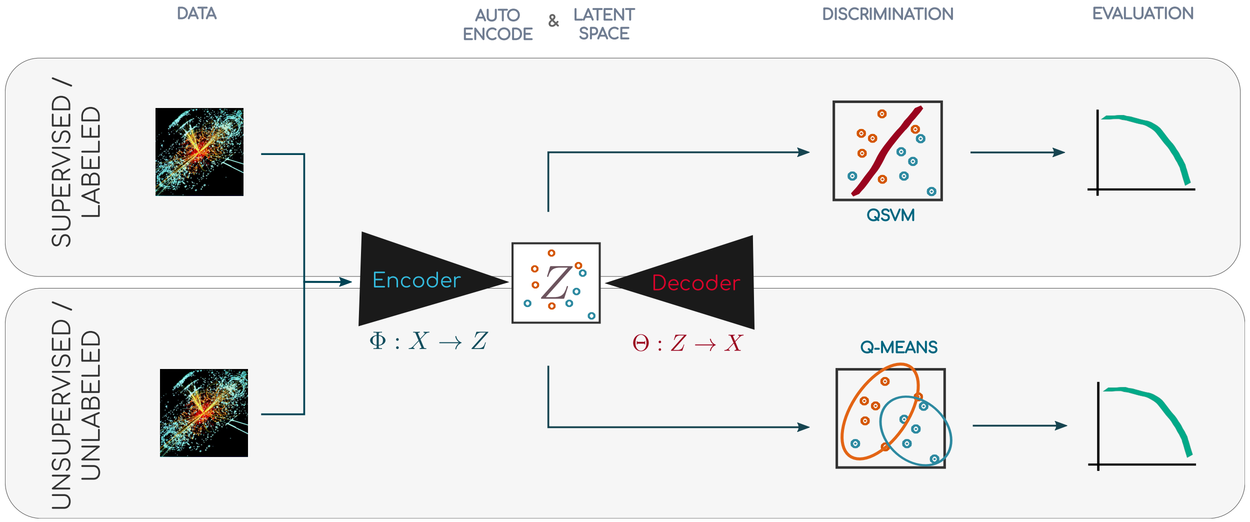
Bravo-Prieto, Carlos, et al. "**Style-based quantum generative adversarial networks for Monte Carlo events**." *arXiv preprint arXiv:2110.06933* (2021).

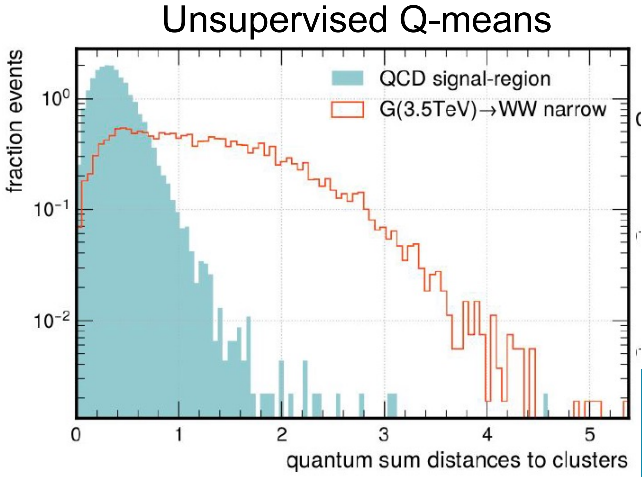


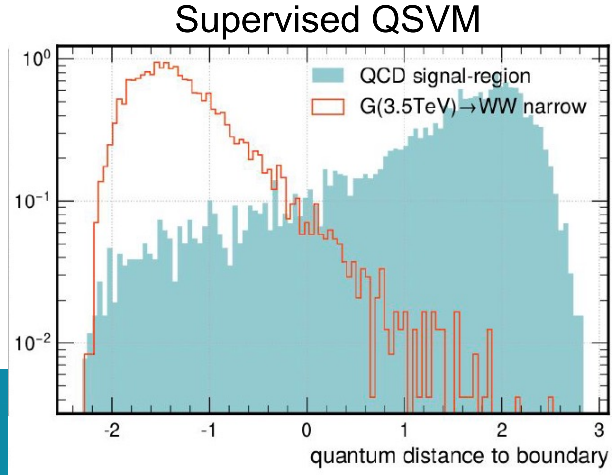
Kinga Wozniak, **Unsupervised clsutering for a Randall–Sundrum Graviton at 3.5TeV narrow resonance, 5th IML workshop, May 2022**

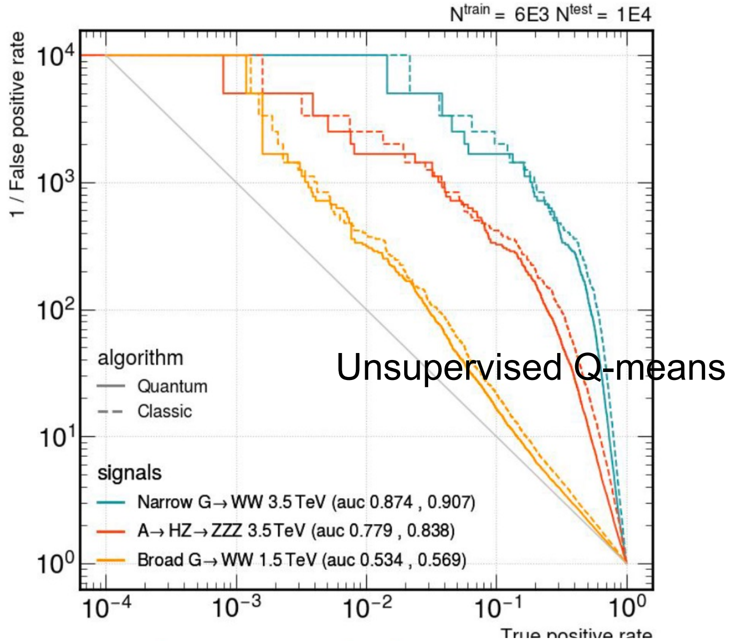
Hybrid setup for anomaly detection

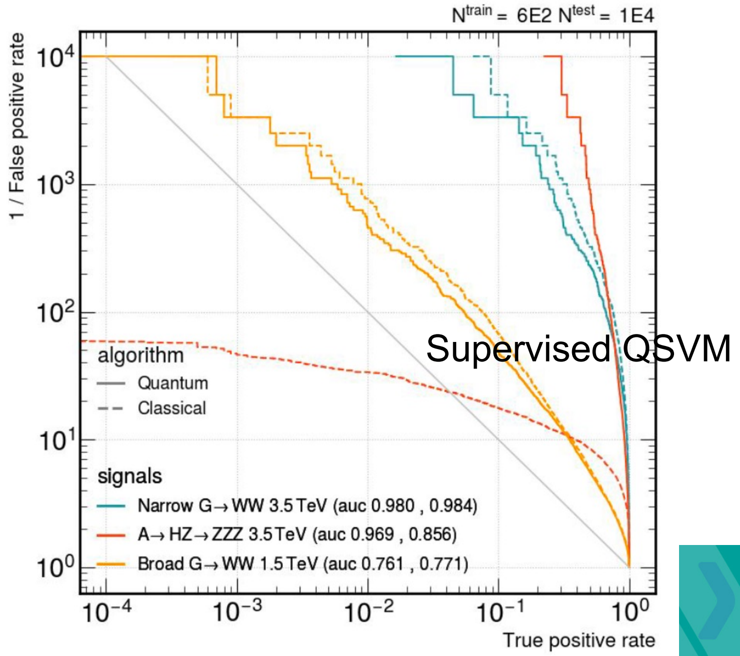
Di-jet events ($\Delta\varphi$, $\Delta\eta$, p_T). Train AE on **QCD sidebands**.
Train classifiers on **signal region**.











Boltzman Machines

Amin, Mohammad H., et al. "**Quantum boltzmann machine.**" *Physical Review X* 8.2 (2018): 021050.

Ex. Compute expected value of physical observable

- In statistical mechanics define a probability function $\pi(x) = \frac{e^{-E(x)}}{\sum_x e^{-E(x)}}$

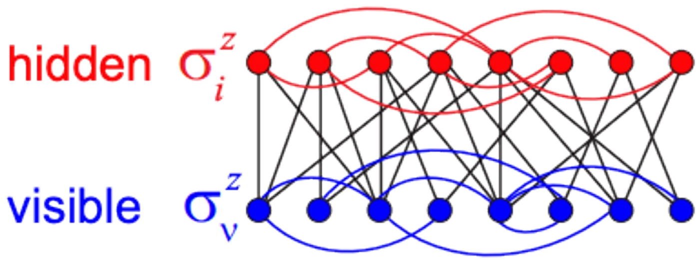
- Minimize the free energy -ln $\mathcal{Z}$ (intractable in general) $\mathcal{Z} = \sum_x e^{-E(x)}$
- Define a variational free energy, for a normalized variational probability q(x)

$$L = \sum_x \mathrm{q(x)ln} \frac{q(x)}{e^{-E(x)}} = \langle \mathrm{E(x)} + \ln \mathrm{q(x)} \rangle_{x \sim q(x)} \qquad L + \ln Z = KL(q||\pi) \geq 0$$

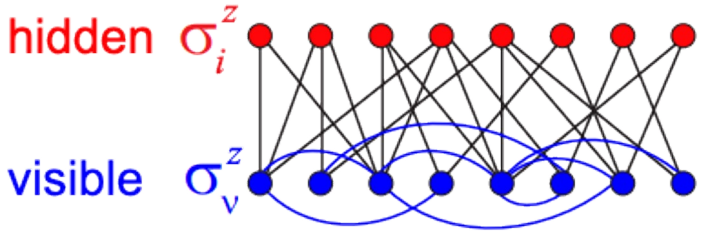
- L is upper bound of physical free energy – ln $\mathcal{Z}$

Quantum Boltzman Machines: replace the energy function with Hamiltonian of a qubit graph (transverse field Ising model)

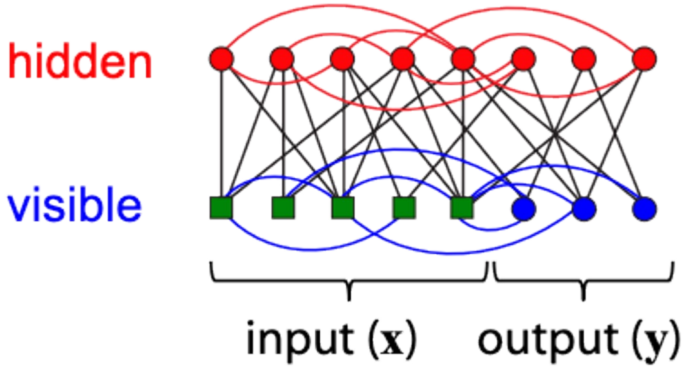
$$H = -\sum_a b_a \sigma_a^z - \sum_{a,b} w_{ab} \sigma_a^z \sigma_b^z$$



Restricted BM:



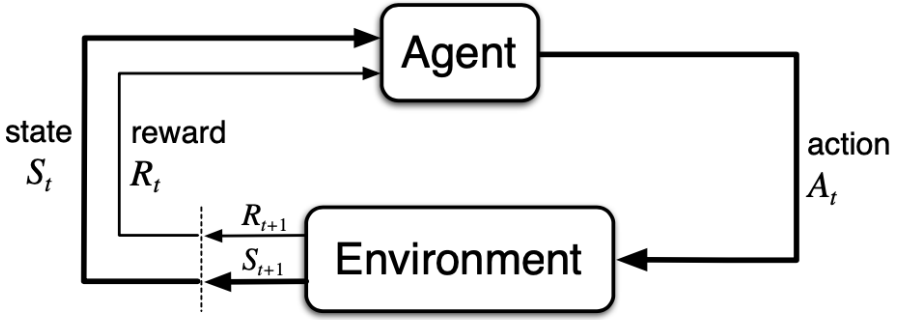
Discriminative learning:



Reinforcement Learning

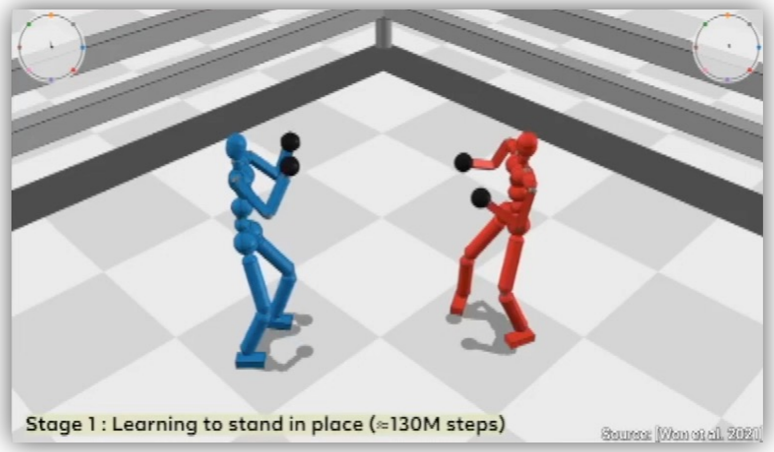
Agent interacts with environment
- **Receives reward after every action**
- Learns through **trial-and-error**
- **Training sample:** $(s_t, a_t, r_t, s_{t+1}, d_t)$

Decision making
- Agent follows **policy** $\boldsymbol{\pi}$: $S \to A$
- **Goal:** find optimal policy π^*
- **Optimal ⇔ maximizing return:** $G_t = \sum_k \gamma^k R_{t+k}$



RL book: Sutton & Barto

Expected return can be estimated through ***value function*** $\boldsymbol{Q(s,a)}$
- Helps answering: **"Best action to take in given state?"**
- Not a priori known, but **can be learned iteratively**



https://www.youtube.com/watch?v=SsJ_AusntiU
https://www.youtube.com/watch?v=Lu56xVlZ40M
https://www.youtube.com/watch?v=imOt8ST4Ej

M. Shenk, V. Kain
BQiT 2021
2022 CERN openlab Tech Workshop

Quantum Reinforcement Learning

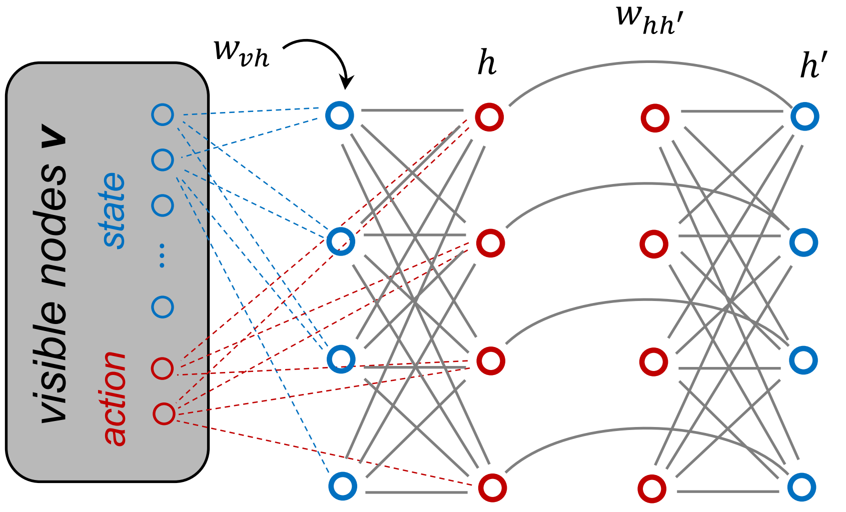
Q-learning: learn $\boldsymbol{Q(s,a)}$ using **function approximator**

- **DQN:** Deep Q-learning *(feed-forward neural network)*
- **FERL:** Free energy based RL *(quantum Boltzmann machine)*

Free Energy RL: clamped Quantum Boltzman Machine

- **Network of coupled, stochastic, binary units** (spin up / down)
- $\hat{\boldsymbol{Q}}(\boldsymbol{s},\boldsymbol{a})$ **≈ negative free energy** of classical spin configurations c
- **Sampling** c using **(simulated) quantum annealing**
- **Clamped:** visible nodes not part of QBM; accounted for as biases
- **Using 16 qubits of D-Wave Chimera graph**
- **Discrete, binary-encoded** state and action spaces

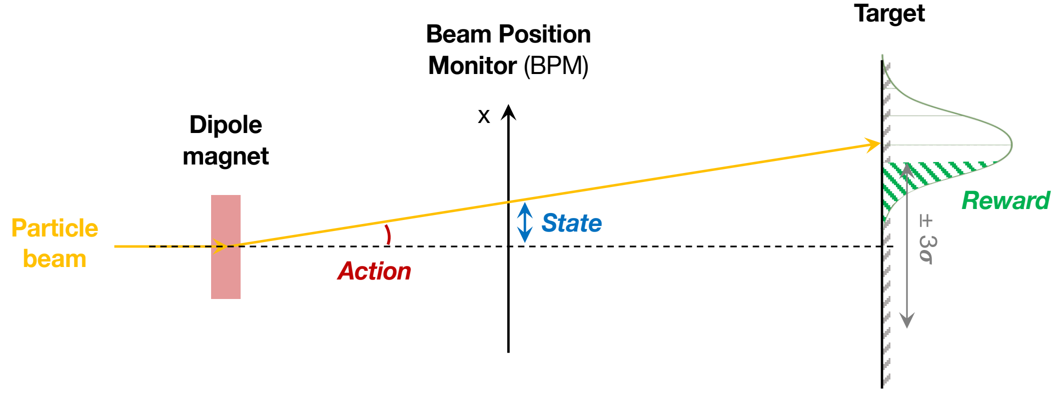
Clamped QBM



$$\hat{Q}(s,a) \approx -F(\boldsymbol{v}) = -\langle H_{\boldsymbol{v}}^{\mathrm{eff}} \rangle - \frac{1}{\beta}\sum_{c} \mathbb{P}(c|\boldsymbol{v}) \log \mathbb{P}(c|\boldsymbol{v})$$

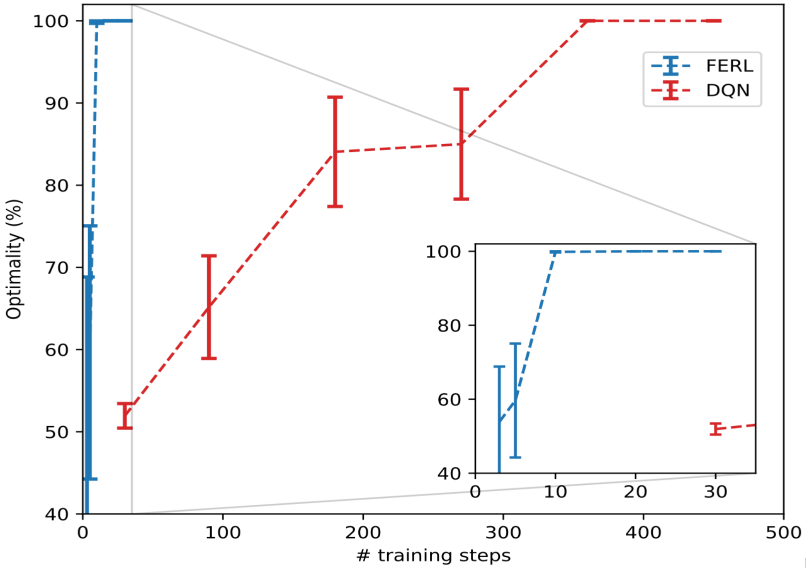
Beam optimisation in linear accelerator

- **Action: deflection** angle
- **State: BPM** position
- **Reward: integrated beam intensity on target**
- **Optimality**: what fraction of possible states does agent take the right decision

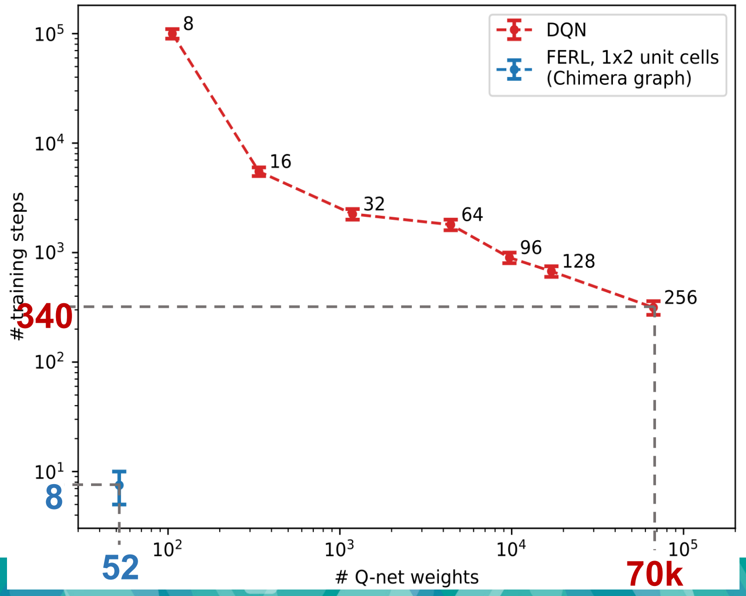


- **Training efficiency:** FERL **massively** outperforms classical Q-learning *(8±2 vs. 320±40 steps)*
- **Descriptive power:** QBM can reach high performance with **much fewer weights** than DQN *(52 vs. ~70k)*

Training efficiency



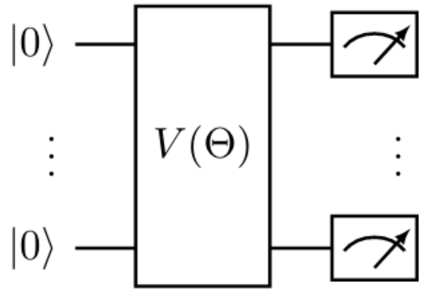
Training efficiency vs. # Q-net / QBM weights



Quantum Circuit Born Machine

Sample from a variational wavefunction $|\psi(\theta)\rangle$ with probability given by the **Born rule**:

$$p_\theta(x) = |\langle x|\psi(\theta)\rangle|^2$$



- Only able to generate **discrete PDFs** (continuous in the limit #qubits $\rightarrow \infty$)
- Train using **Maximum Mean Discrepancy**:

$$\mathrm{MMD}(P,Q) = \mathbb{E}_{\substack{X\sim P\\ Y\sim P}}[K(X,Y)] + \mathbb{E}_{\substack{X\sim Q\\ Y\sim Q}}[K(X,Y)] - 2\mathbb{E}_{\substack{X\sim P\\ Y\sim Q}}[K(X,Y)]$$

with K a gaussian kernel

- **Pros**: relativly easy to optimize, **Cons**: empircally less efficient than an adversarial approach

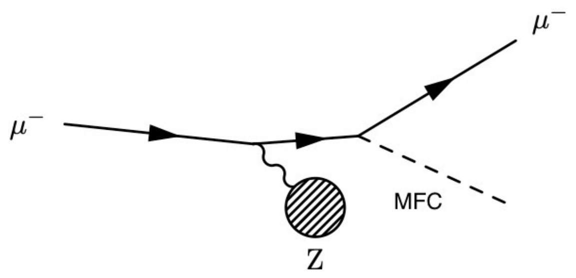
Coyle, B., Mills, D. et al, **The Born supremacy.** In: *npj Quantum Inf 6, 60* (2020)

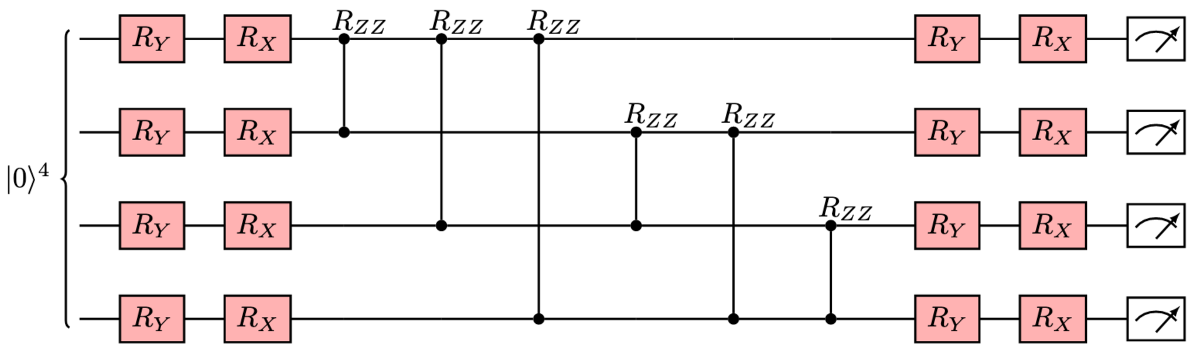
QCBM for event generation

Muon Force Carriers predicted by several theoretical models:

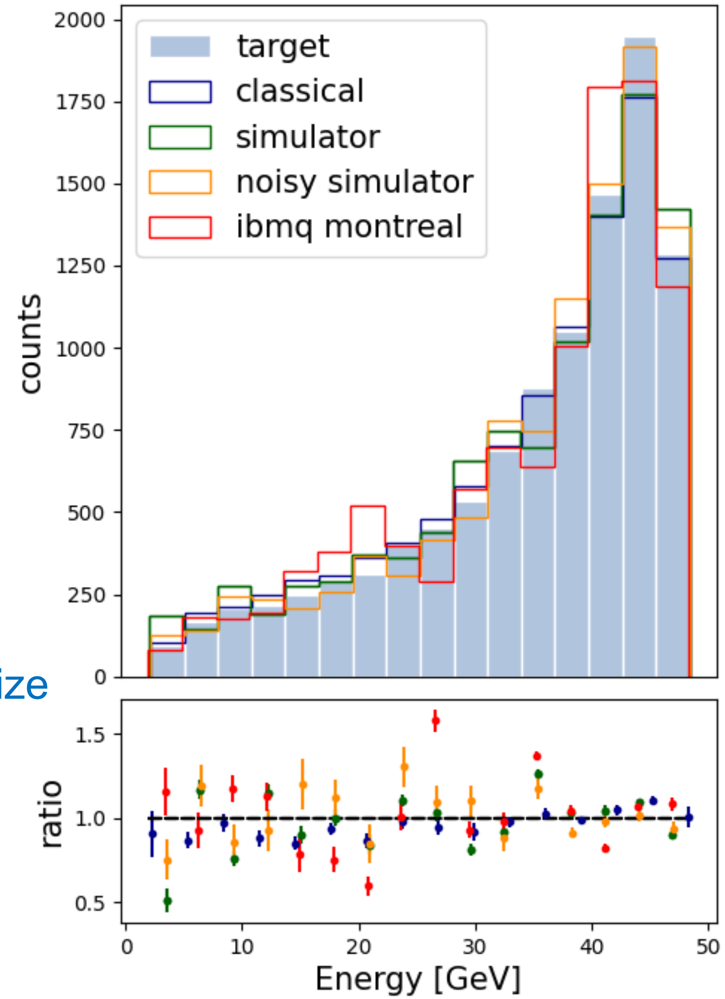
- Could be detected by muon fixed-target experiments (FASER) or muon interactions in calorimeters (ATLAS)[1].

Generate E, p_t, η of outgoing muon and MFC





Perfect simulator
Noisy simulator (IBMQ casablanca) (no error mitigation)
IBMQ Montreal
Classical GMMD of size (15,128, 256,128,16,1)
Easy GMMD ~ QCBM in size

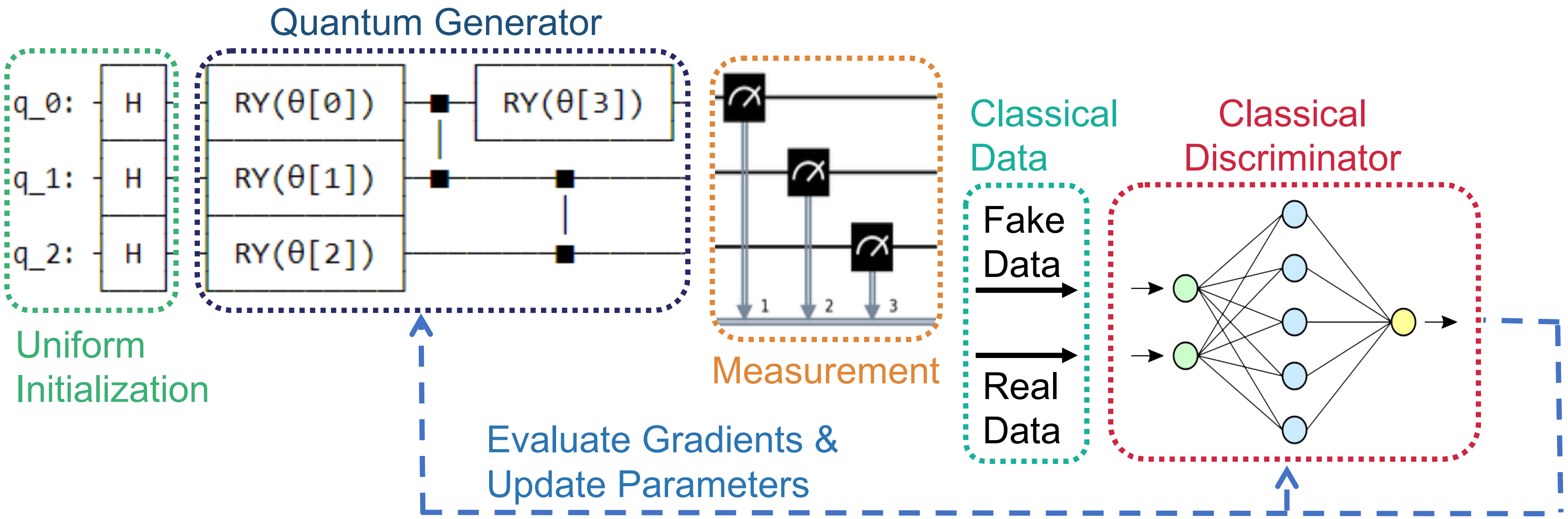


1 Galon, I, Kajamovitz, E et al. "Searching for muonic forces with the ATLAS detector". In: Phys. Rev. D 101, 011701 (2020)

Quantum Generative Adversarial Networks

Density estimation by comparison

- Sample-based comparison between **estimated** q(x) and **true distribution** p(x)
- Multiple implementations, mostly classical-quantum hybrid
- Used for
  - Data generation
  - PDF loading on quantum systems
  - Anomaly detection

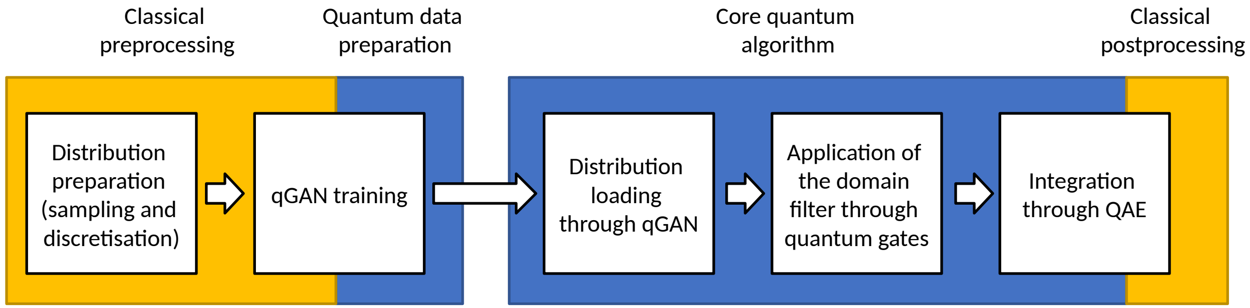


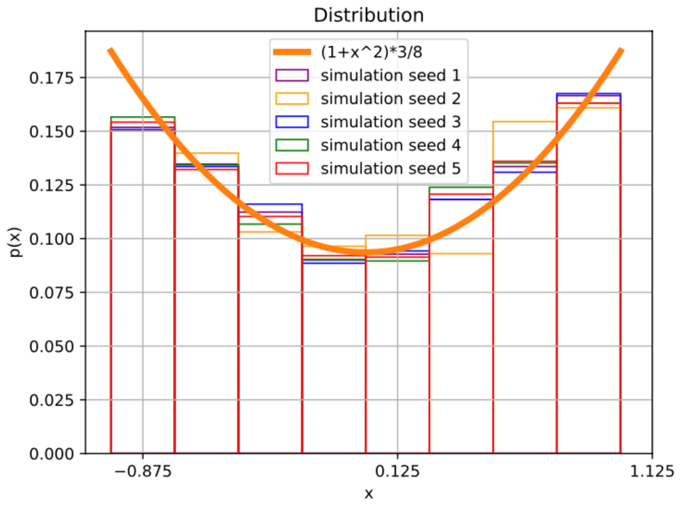
qGAN as a data loader

Use **Quantum Amplitude Estimation** to accelerate **Monte Carlo Integration**

Data encoding into quantum states affects the quality of the integration

Test different approaches including **QGAN**





Loading of $1 + x^2$ distribution:

- 10k events
- 3 qubits
- circular entanglement

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

FR-PHENO-2022-01

arXiv:2201.01547v1 [hep-ph] 5 Jan 2022

Quantum integration of elementary particle processes

Gabriele Agliardi[1,2]*, Michele Grossi[3]†, Mathieu Pellen[4]‡, Enrico Prati[5,6]§

[1] Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy
[2] IBM Italia S.p.A., Via Circonvallazione Idroscalo, I-20090 Segrate (MI), Italy
[3] CERN, 1 Esplanade des Particules, Geneva CH-1211, Switzerland
[4] Albert-Ludwigs-Universität Freiburg, Physikalisches Institut, Hermann-Herder-Straße 3, D-79104 Freiburg, Germany
[5] Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy
[6] National Inter-university Consortium for Telecommunications (CNIT), Viale G.P. Usberti, 181/A Pal.3, I-43124 Parma, Italy

Abstract

We apply quantum integration to elementary particle-physics processes. In particular, we look at scattering processes such as $e^+e^- \to q\bar{q}$ and $e^+e^- \to q\bar{q}'W$. The corresponding probability distributions can be first appropriately loaded on a quantum computer using either quantum Generative Adversarial Networks or an exact method. The distributions are then integrated using the method of Quantum Amplitude Estimation which shows a quadratic speed-up with respect to classical techniques. In simulations of noiseless quantum computers, we obtain per-cent accurate results for one- and two-dimensional integration with up to six qubits. This work paves the way towards taking advantage of quantum algorithms for the integration of high-energy processes.

*E-mail: gabrielefrancesco.agliardi@polimi.it
†E-mail: michele.grossi@cern.ch
‡E-mail: mathieu.pellen@physik.uni-freiburg.de
§E-mail: enrico.prati@ifn.cnr.it

qGAN for event generation

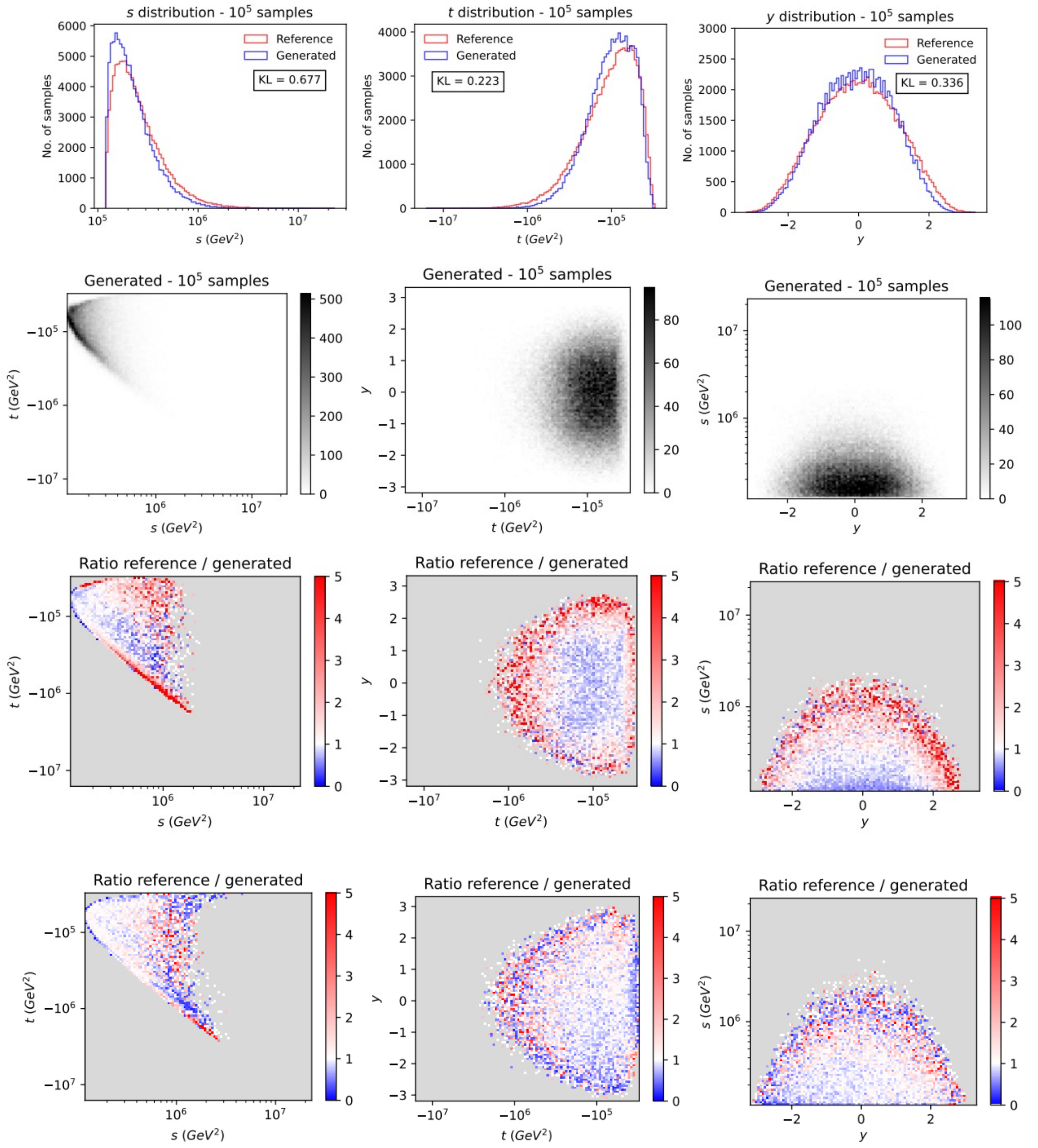
Generate Mandelstam (s,t) + y variables for t-tbar production

Introduce a style-based approach

	$pp \to t\bar{t}$ **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

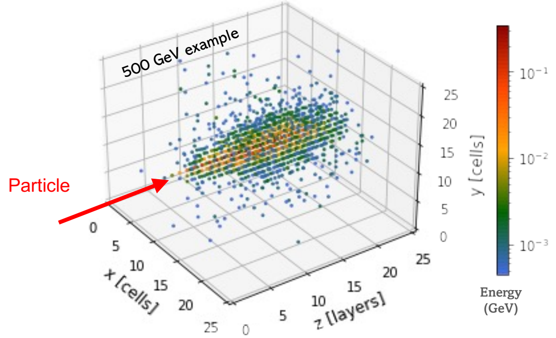
Quantum simulator



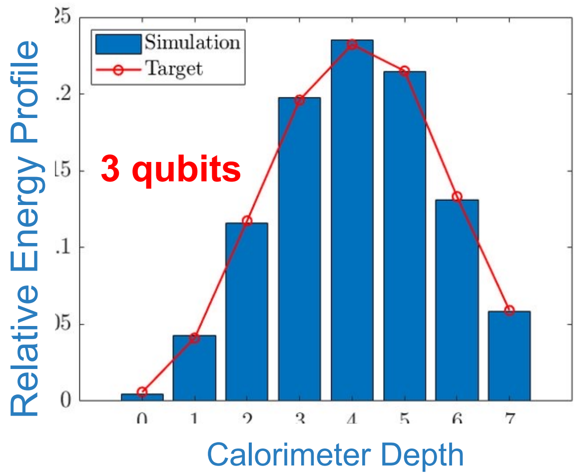
Bravo-Prieto, Carlos, et al. "**Style-based quantum generative adversarial networks for Monte Carlo events**." *arXiv preprint arXiv:2110.06933* (2021).

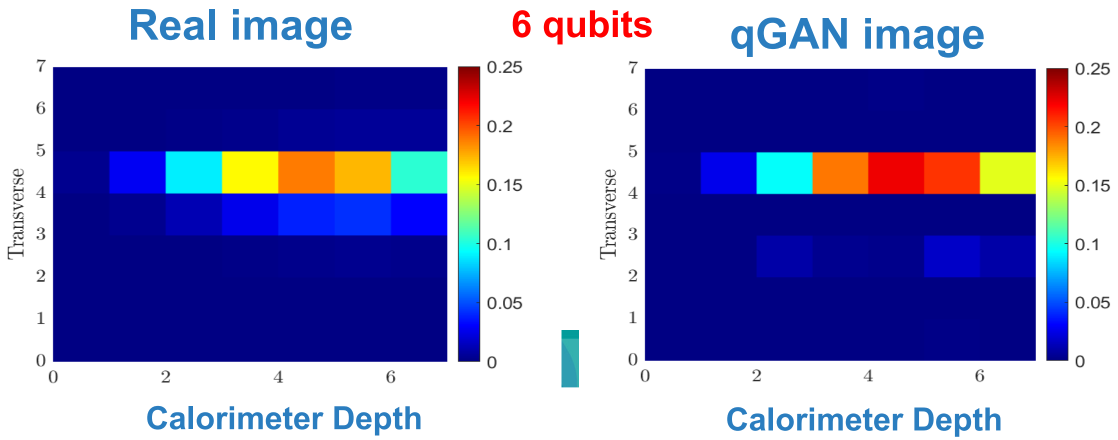
Increasing generated dimensionality

Energy Profiles in Calorimeters



- Calorimeter simulation is one of the main use cases for classical GAN in HEP
- Represented as a 3D regular grid
- Reduce to:
  - 1D distribution along the calorimeter depth (8 pixel)
  - **2D distribution on the y-z plane (64 pixel)**





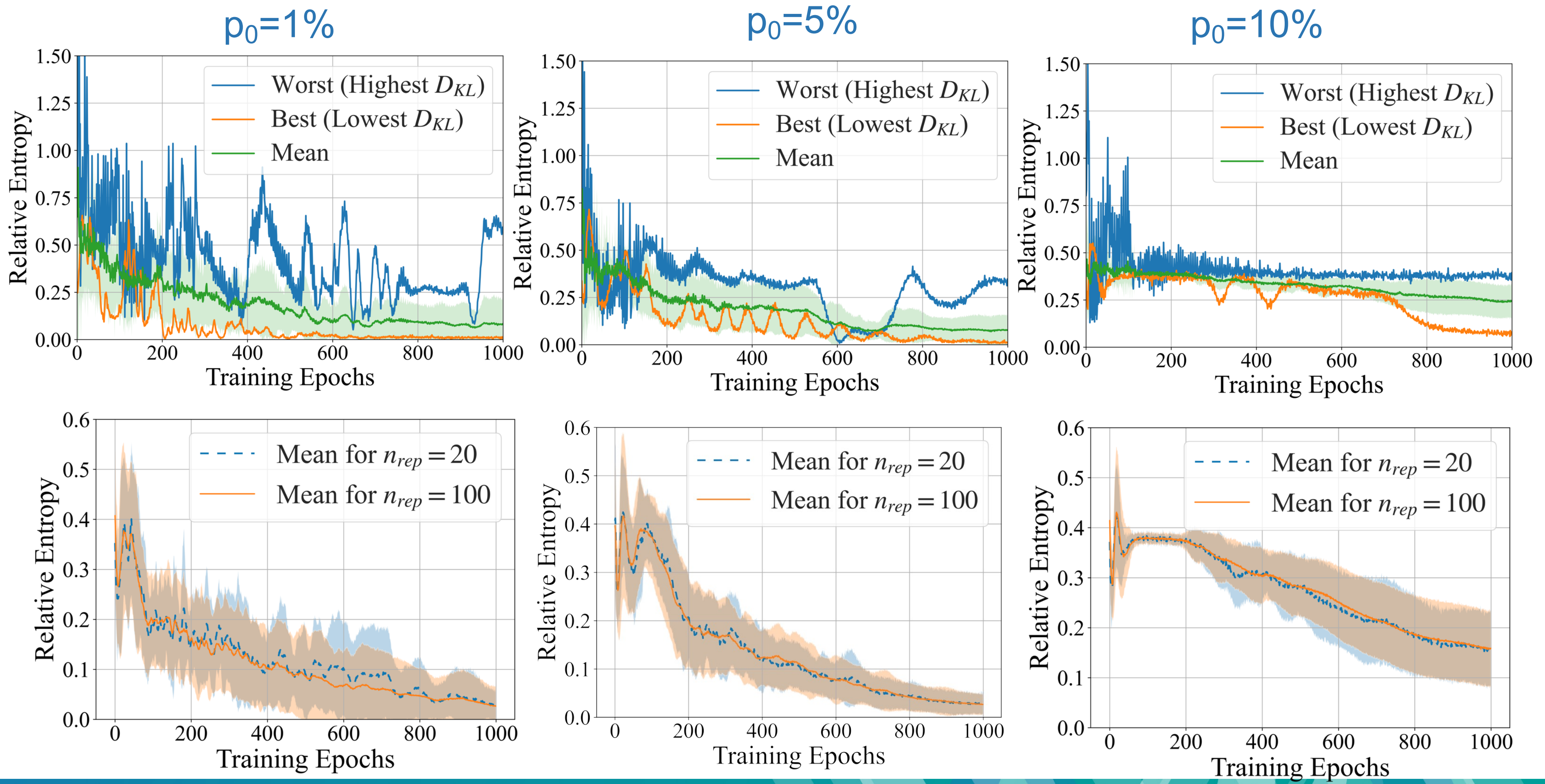
Rehm, Florian, et al. "Quantum Machine Learning for HEP Detector Simulations." (2021).

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

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

Readout noise effect on GAN training

- Training is up to ~5% readout **noise tolerant**
- **Higher readout noise reduces** accuracy
- **Intrinsic instability** in the training process

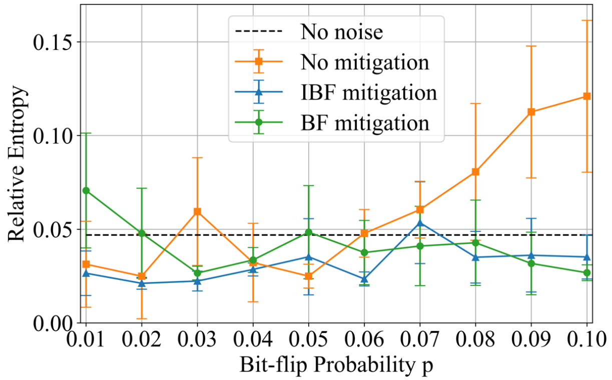


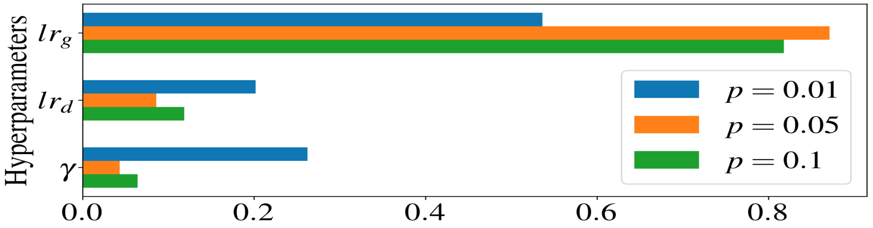
Running the model on noisy devices

Florian Rehm, S. Y. Chang:
https://arxiv.org/abs/2203.01007

Train on noisy simulator

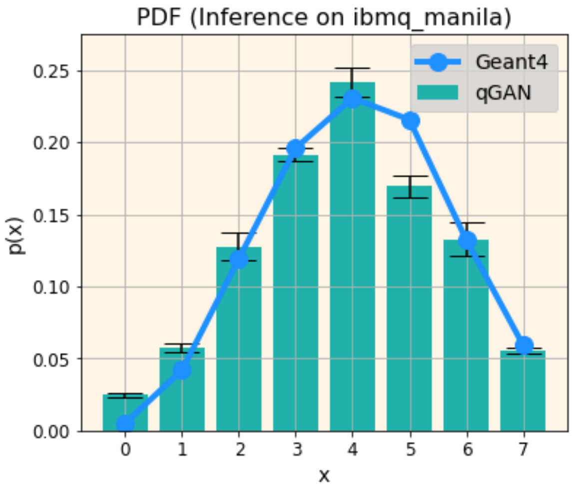
- Evaluate importance of training hyperparameters
- **Error mitigation needed only** for **higher** noise level

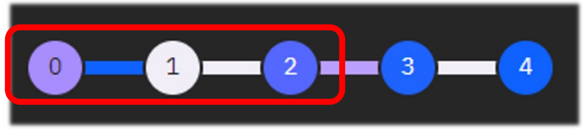




Inference on IBM Q Manila hardware

- Maintain good physics perfomance





Qubit Number	0	1	2
Readout Error	2.34%	2.66%	2.05%
CX-gate Error	1.11%		1.75%

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 lowes

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

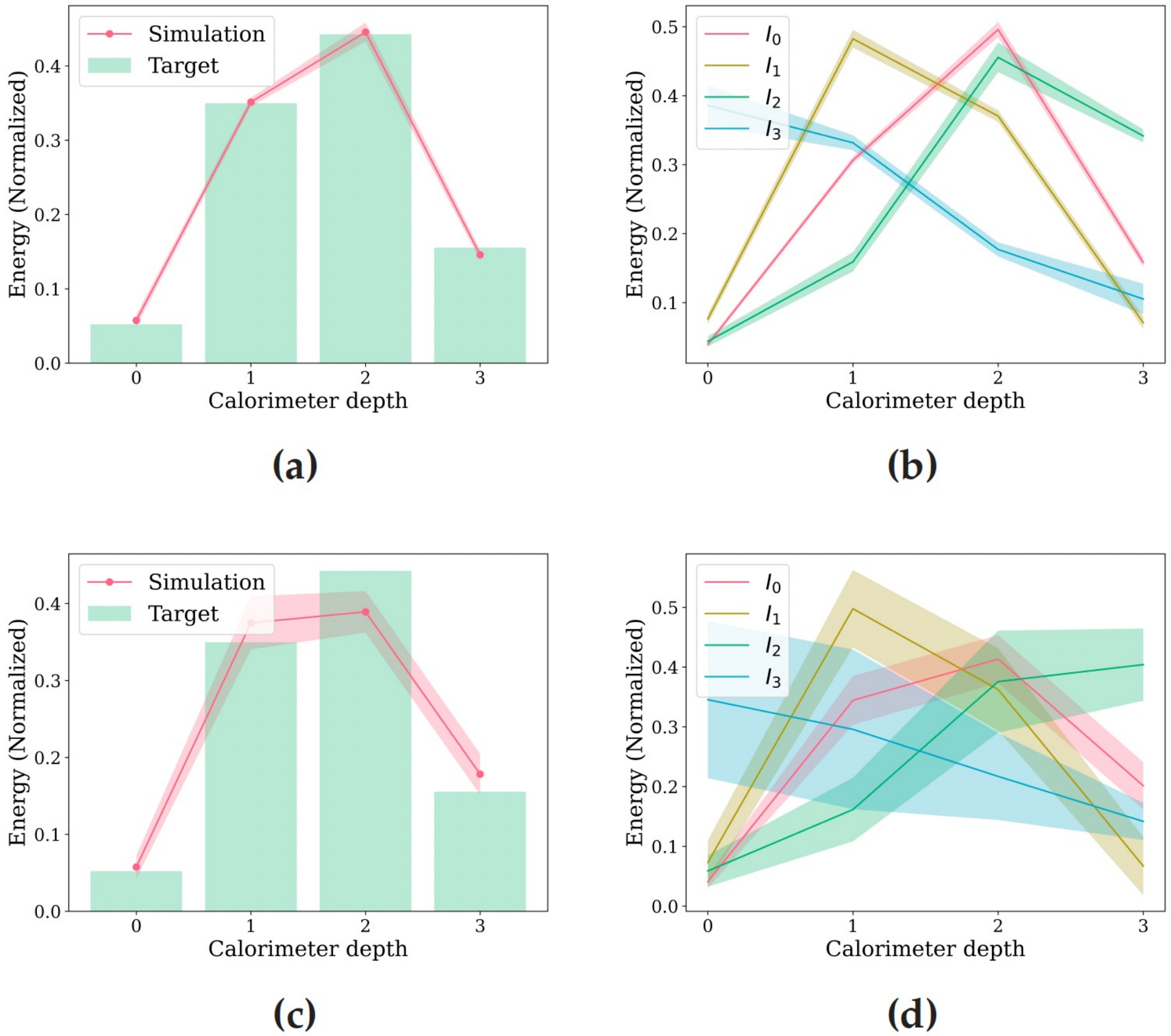


Figure 4: Mean (a,c) and individual images (b,d) obtained by inference test on ibmq_jakarta (a,b) and IONQ (c,d).

Summary

Research on QML applications in High Energy Physics is producing a large number of prototypes

- So far focus on different steps of data processing in «controlled environment»
- Some **preliminary hints** of advantage in terms of input feature size and representational power
- Mostly we do «**as good as classical methods**»
- Need **more robust studies** to relate quantum model architecture and performance to data sets
- Identify use cases where **quantum approach** could be **more effective** than classical machine/deep learning
- Studying QML algorithms today can build links between **QC and learning theory**

GET IN TOUCH IF YOU'D LIKE TO CONTRIBUTE!

Thanks!

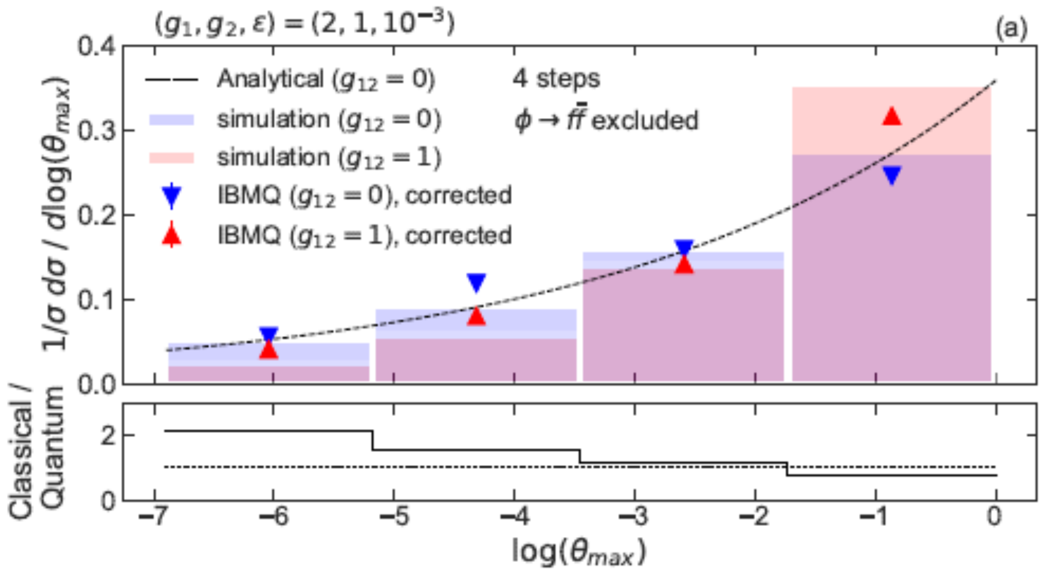
Sofia.Vallecorsa@cern.ch

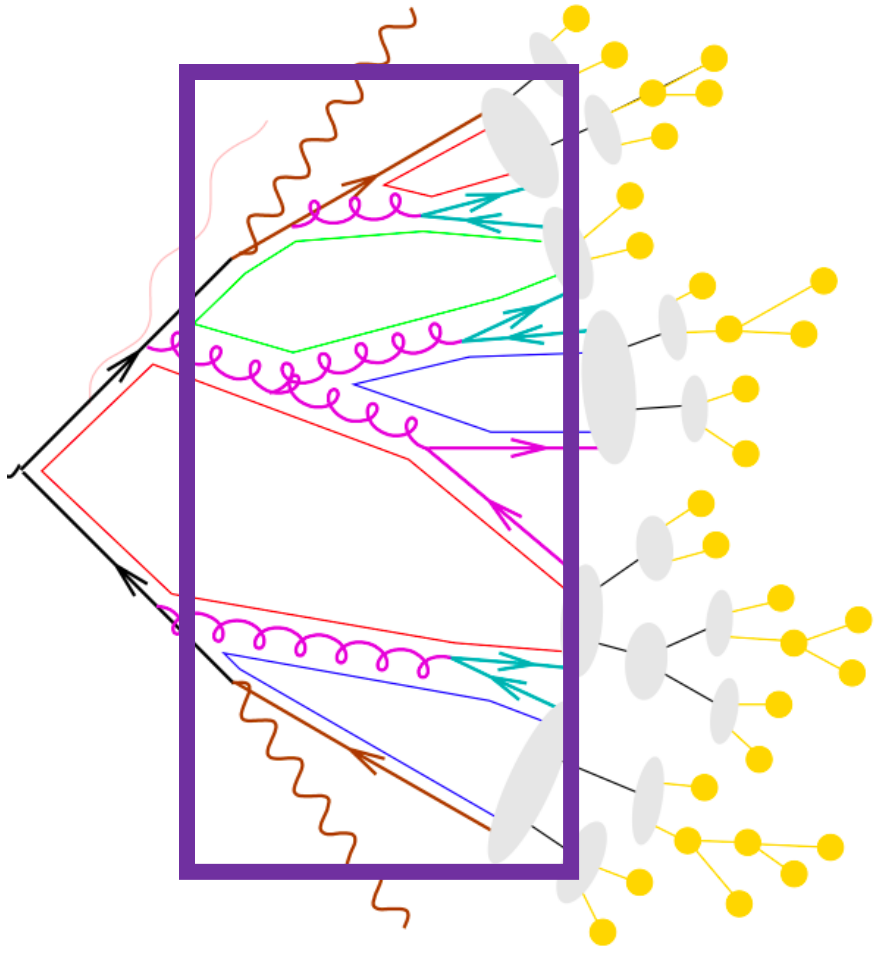
CERN and the Quantum Technology Intiative

Direct simulation

- Quantum computer can **naturally simulate quantum systems** (reproduce the evolution of an Hamiltonian via the Schrodinger equation)
- Quantum **chemistry** and **physics**

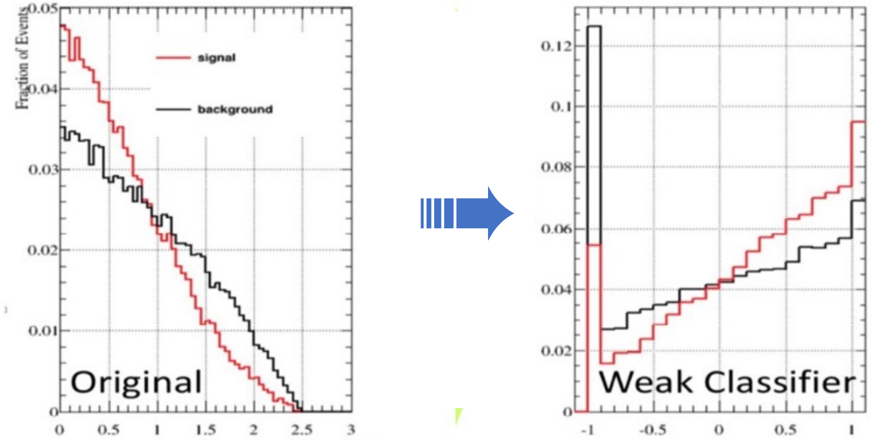
- Exploit entanglement between qubits on a quantum computer to **simulate correlations in the parton shower**





arxiv:1904.03196

Weak → Strong classifier



$h_i(x) \in [-1,1]$ are functions of the variables such that

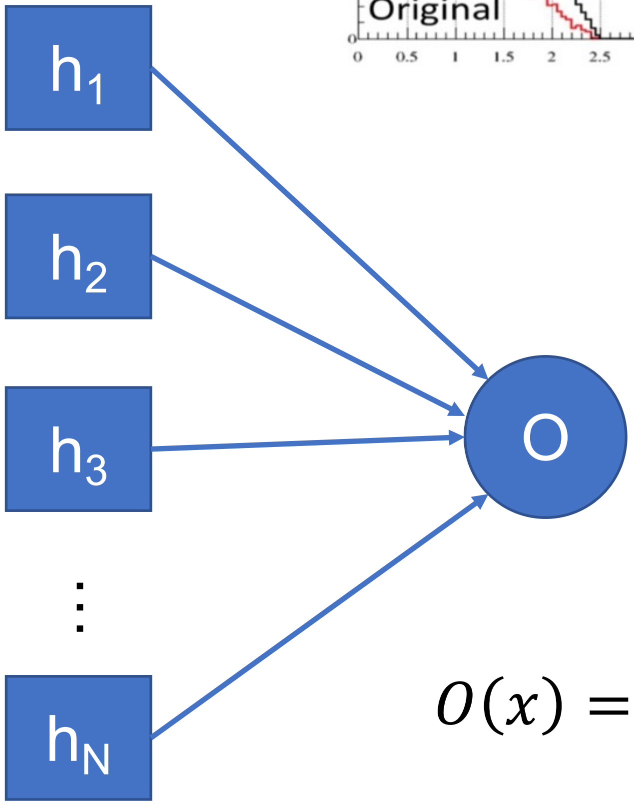
P(S|h_i>0) > P(B|h_i>0)

P(B|h_i<0) > P(S|h_i<0)

i.e.

h_i>0 probably Signal

h_i<0 probably Background

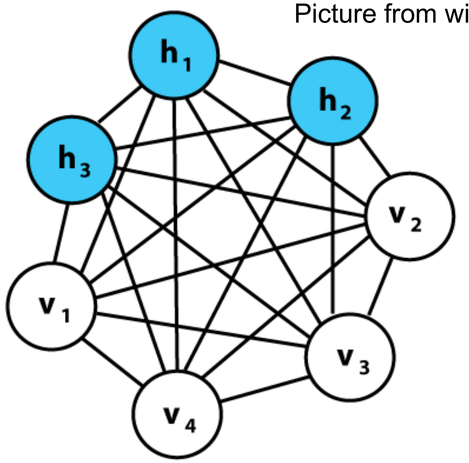


$$O(x) = \sum_i w_i h_i(x)$$

https://arxiv.org/abs/1109.0325

Quantum Boltzmann Machines

Picture from wikipedia



Train GBM parameters to learn the underlying data distribution

→ Generative models

→ Feature mapping

→ Can act as classifiers by clamping part of the visible units

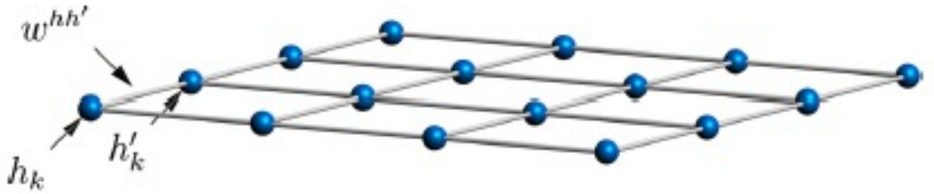
Classical Energy:

$$\mathcal{H}_{\mathbf{v}}(\mathbf{h}) = - \sum_{v \in V, h \in H} w^{vh} v h - \sum_{\{h,h'\} \subseteq H} w^{hh'} h h'$$

Quantum BM: qubits graph as a **transverse field Ising model**

Hamiltonian:

$$\mathcal{H}_{\mathbf{v}} = - \sum_{v \in V, h \in H} w^{vh} v \sigma_h^z - \sum_{\{h,h'\} \subseteq H} w^{hh'} \sigma_h^z \sigma_{h'}^z - \Gamma \sum_{h \in H} \sigma_h^x ,$$



Solving QBMs on the annealer

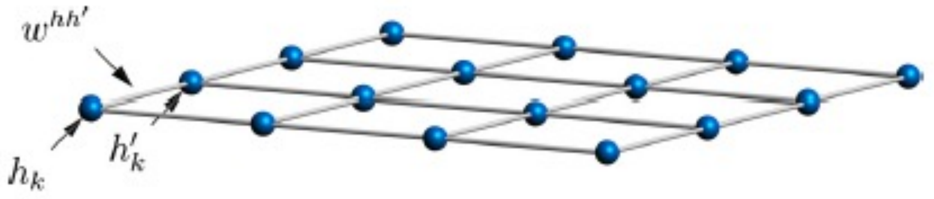
$$\mathcal{H}_{\mathbf{v}} = -\sum_{v\in V,\, h\in H} w^{vh} v \sigma_h^z - \sum_{\{h,h'\}\subseteq H} w^{hh'} \sigma_h^z \sigma_{h'}^z - \Gamma \sum_{h\in H} \sigma_h^x ,$$

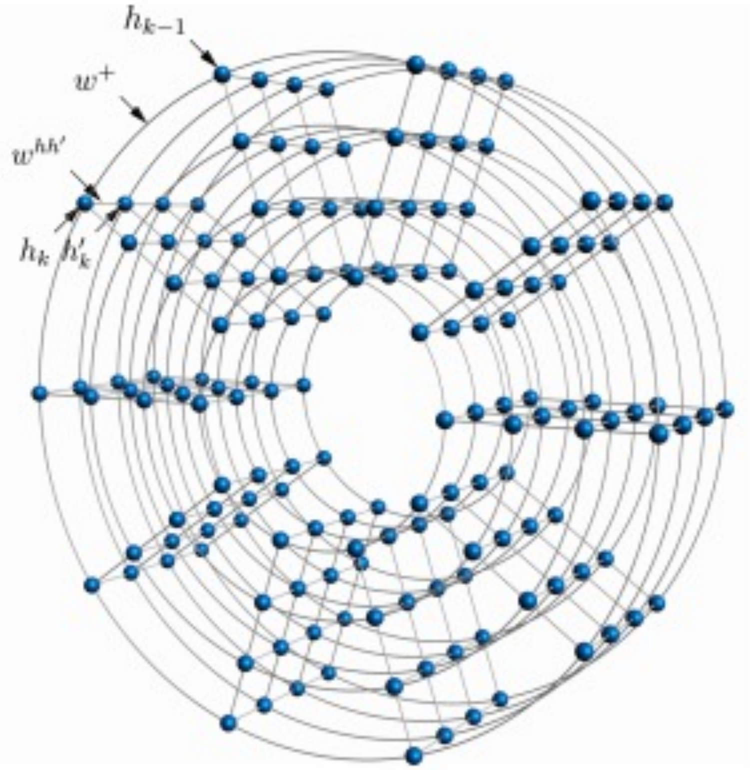
Spins configurations can only be measured along one axis.

Measuring along σ^z → σ^x collapses

$\mathcal{H}_v$ can't be measured directly → use an approximation (Suzuki-Trotter representation)

Stack a set of classical Ising models (one dimension higher)





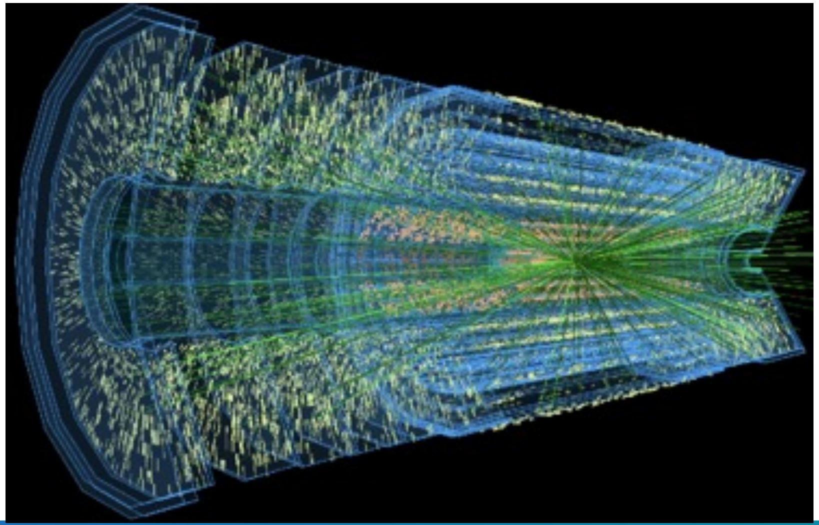
Cenk Tuysuz
arXiv:2012.01379

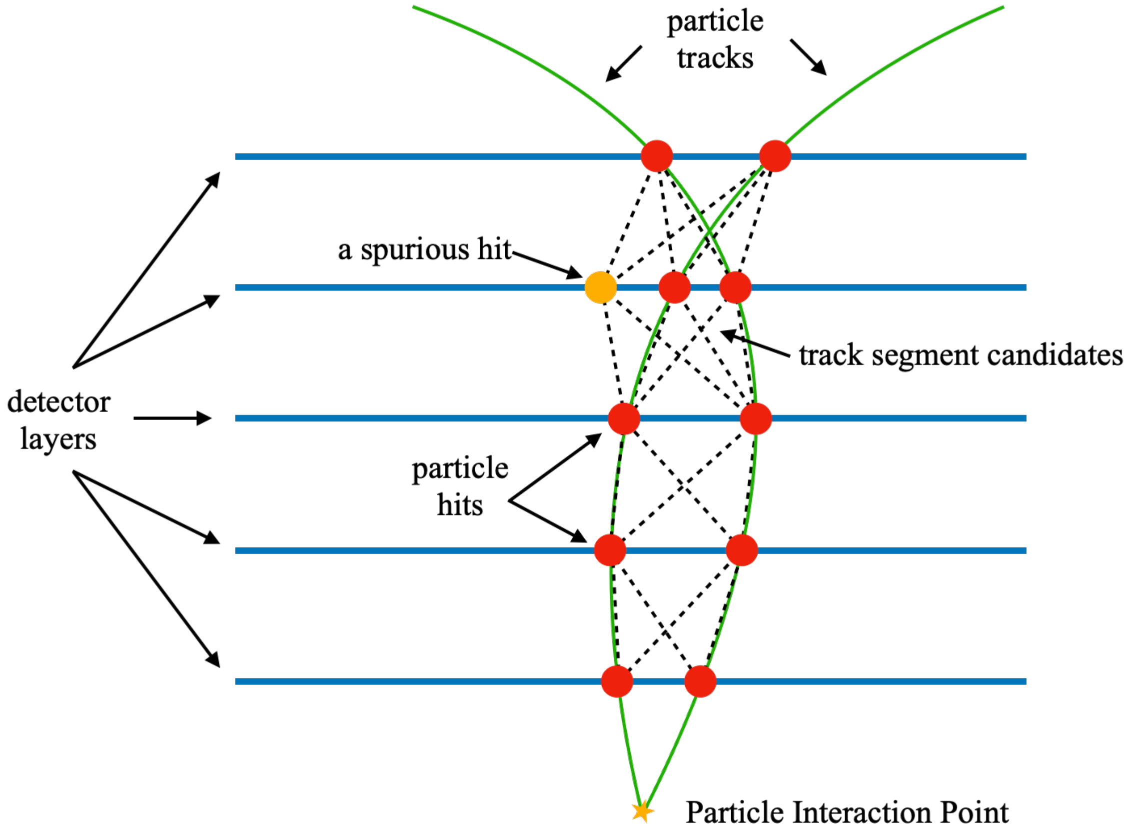
arXiv:2109.12636

Charged particle tracking

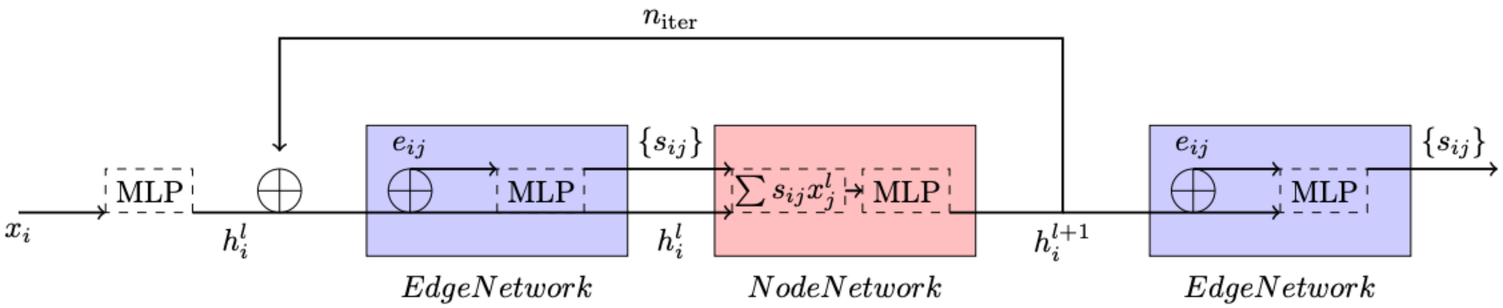
Graph Neural Networks for particle trajectory reconstruction

- Data as a **graph of connected hits**
- Connect hits using **geometric** constraints
- Embedding requires **large graphs** (~10^5 nodes)





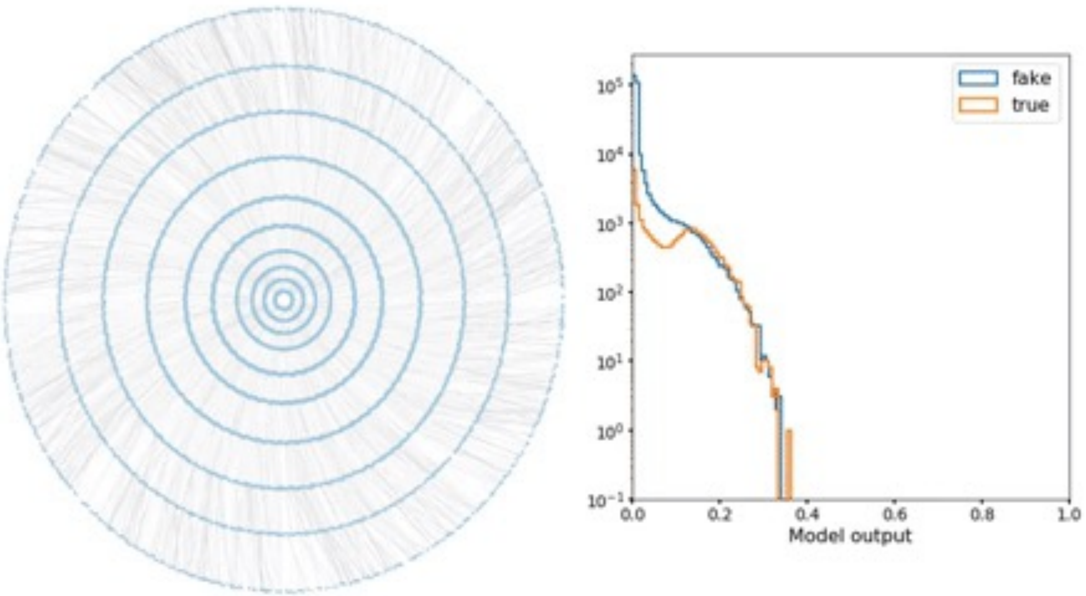
GNN for particle tracking



arxv:2007.00149

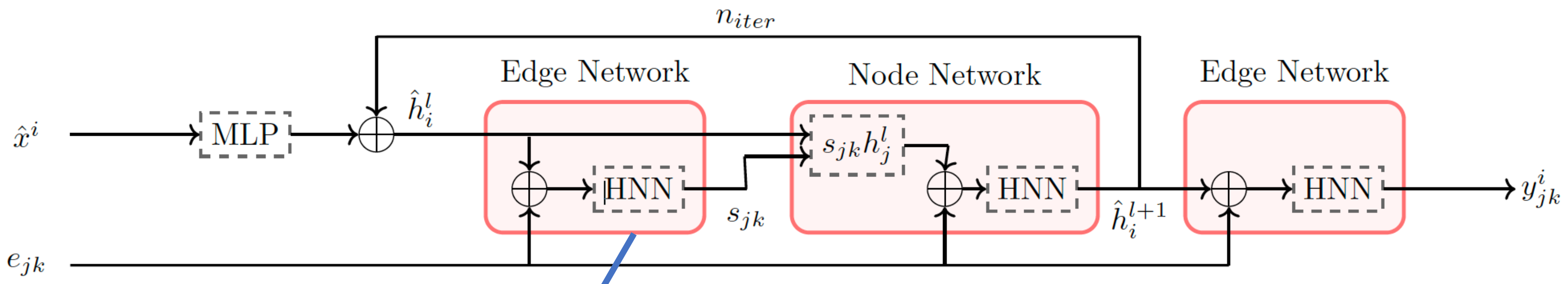
Exa.TrkX

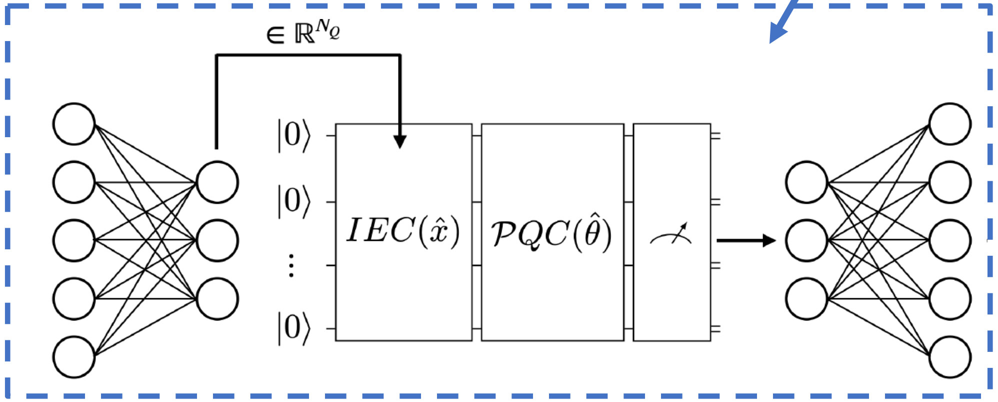
https://exatrkx.github.io/

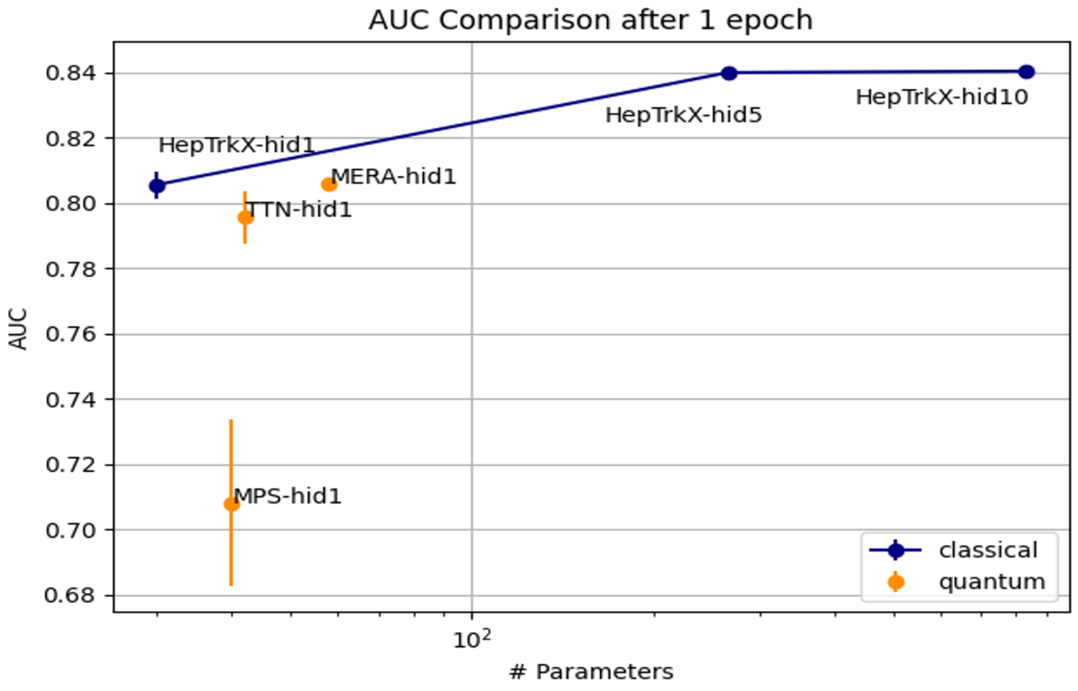


Quantum models

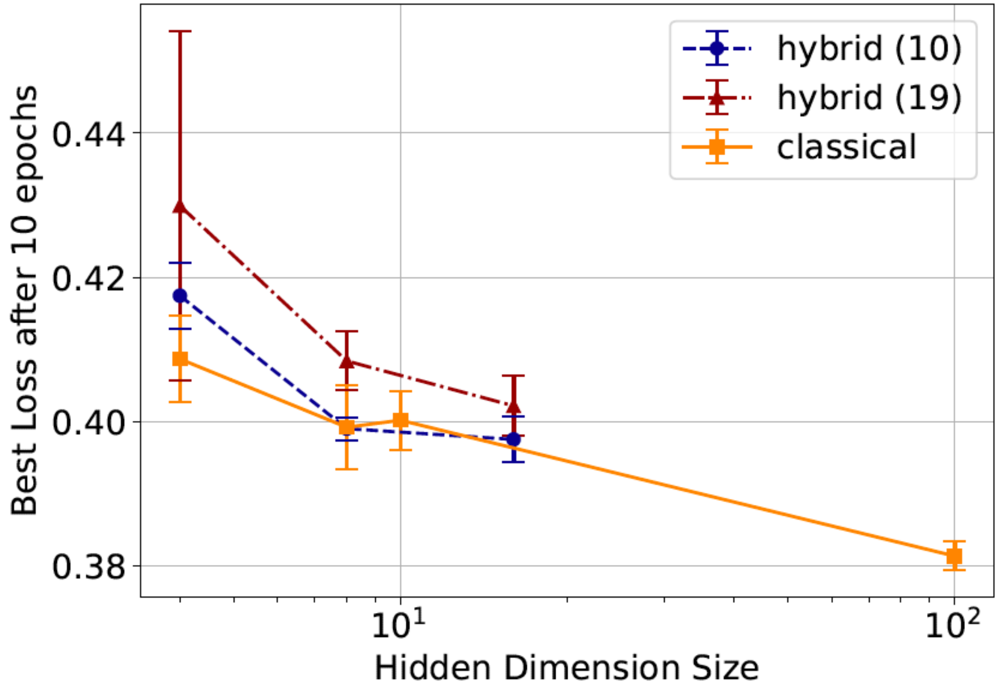
Replace Edge and Node networks with **hybrid classifiers**

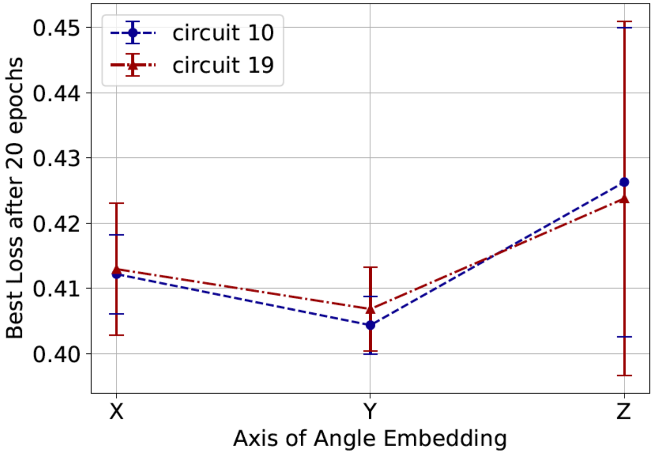




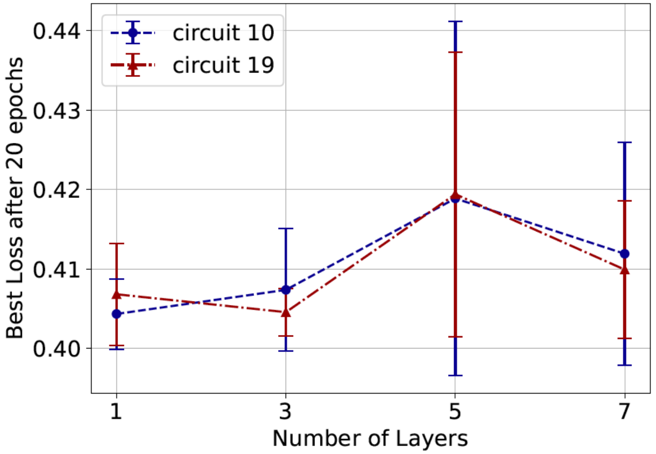


Quantum circuit systematics

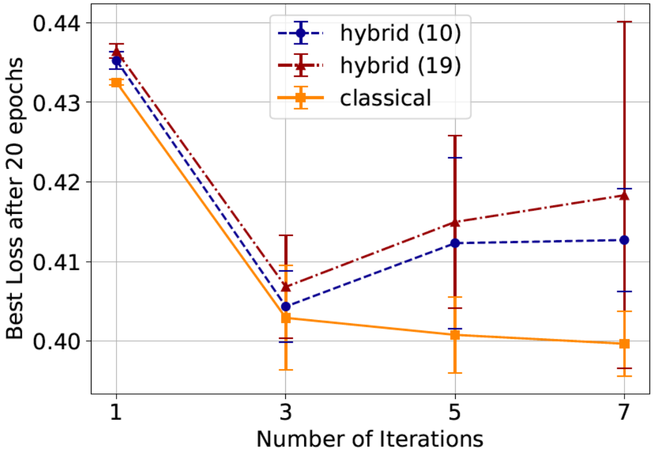




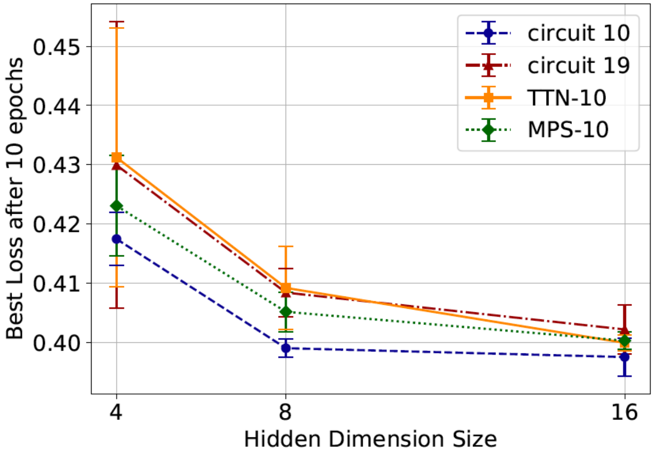
(a) Axis of angle embedding comparison.



(b) Number of layers (N_L) comparison.



(c) Number of iterations (N_I) comparison.



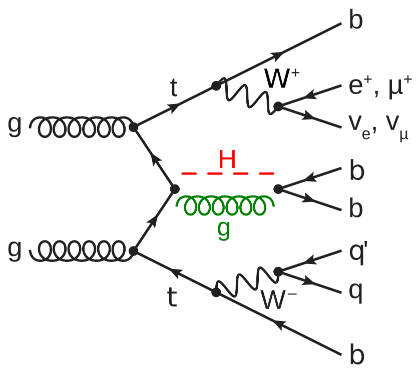
(d) Hidden dimension size ($N_D = N_Q$) comparison.

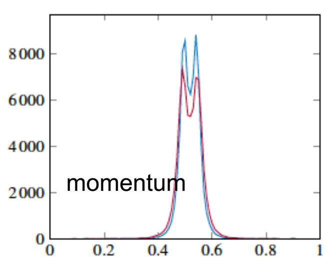
V. Belis, S. Gonzalez-Castillo
BQiT 2021
vCHEP2021
arXiv:2104.07692

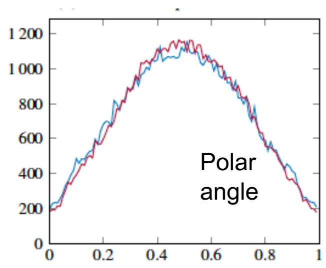
Quantum SVM for Higgs classification

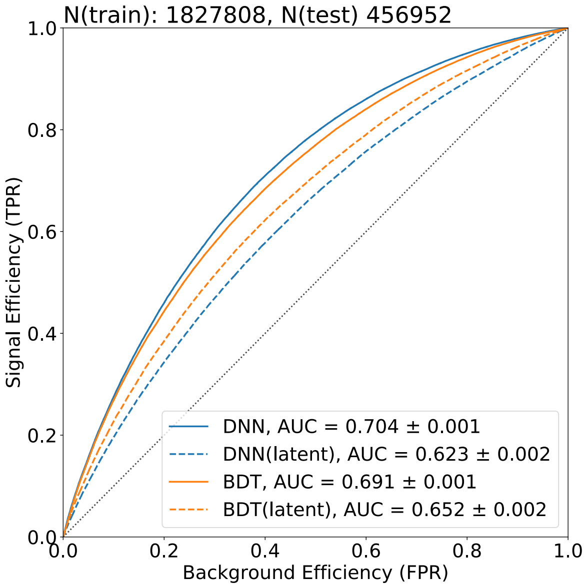
Classical models trained on 67 features

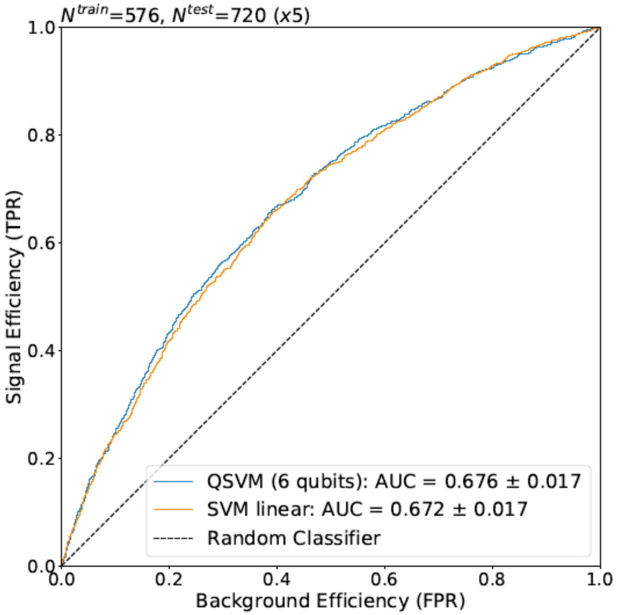
Test several dimensionality reduction strategies (PCA, AutoEncoder, Kmeans..)



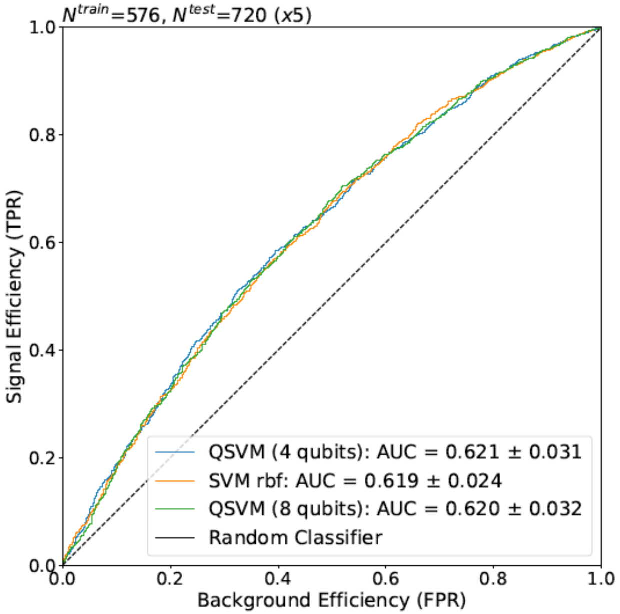








(b) Models trained on the original input features (67), discarding the 3 least informative ones (64).

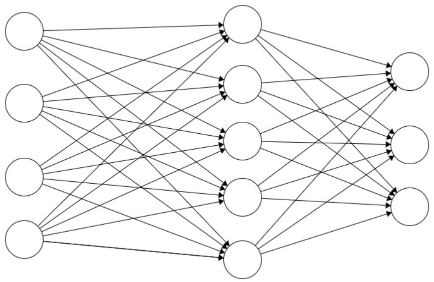


(a) Models trained on the AE latent space features (16).

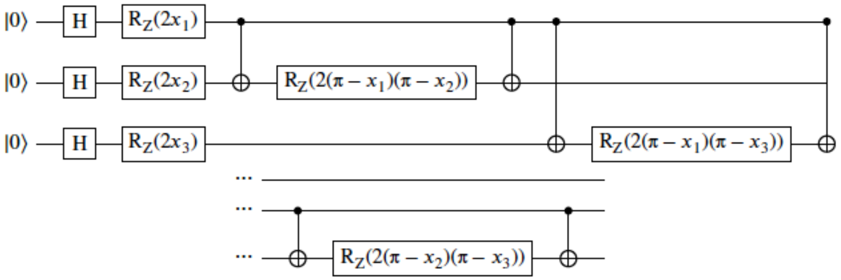
Variational Quantum Classifiers

Classical dense neural network to reduce dimensionality

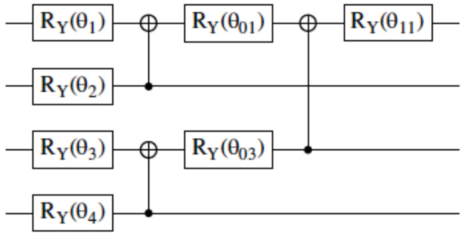
- 4 qubits, 8 variables



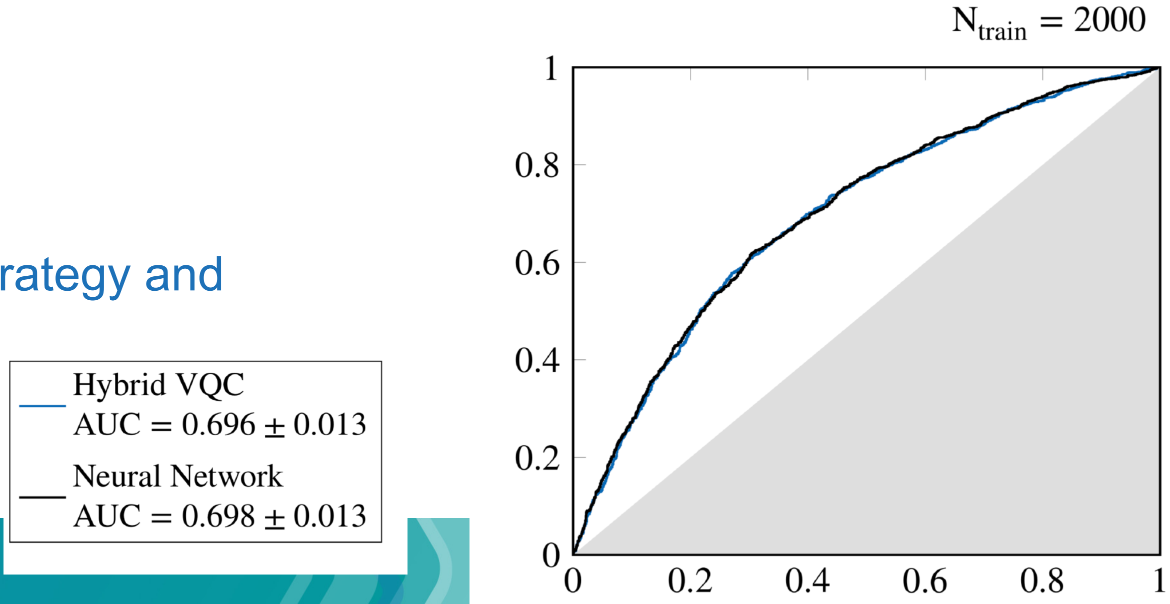
ZZ feature map with data-re-uploading



2-local variational form



Simultaneous training of classical feature extraction strategy and quantum classifier improves the accuracy



CERN

QUANTUM
TECHNOLOGY
INITIATIVE