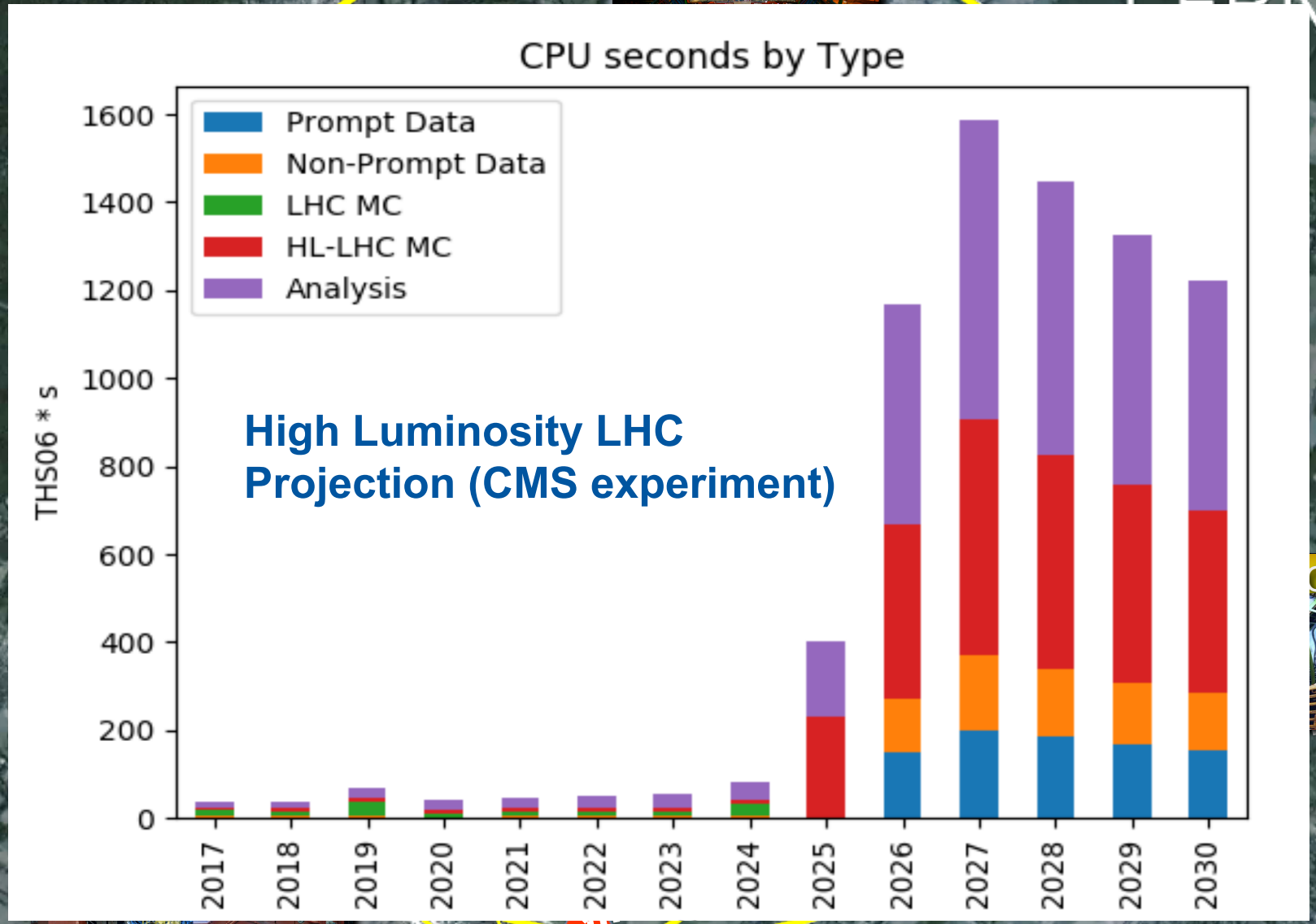
CERN

The CERN Quantum Technology Initiative

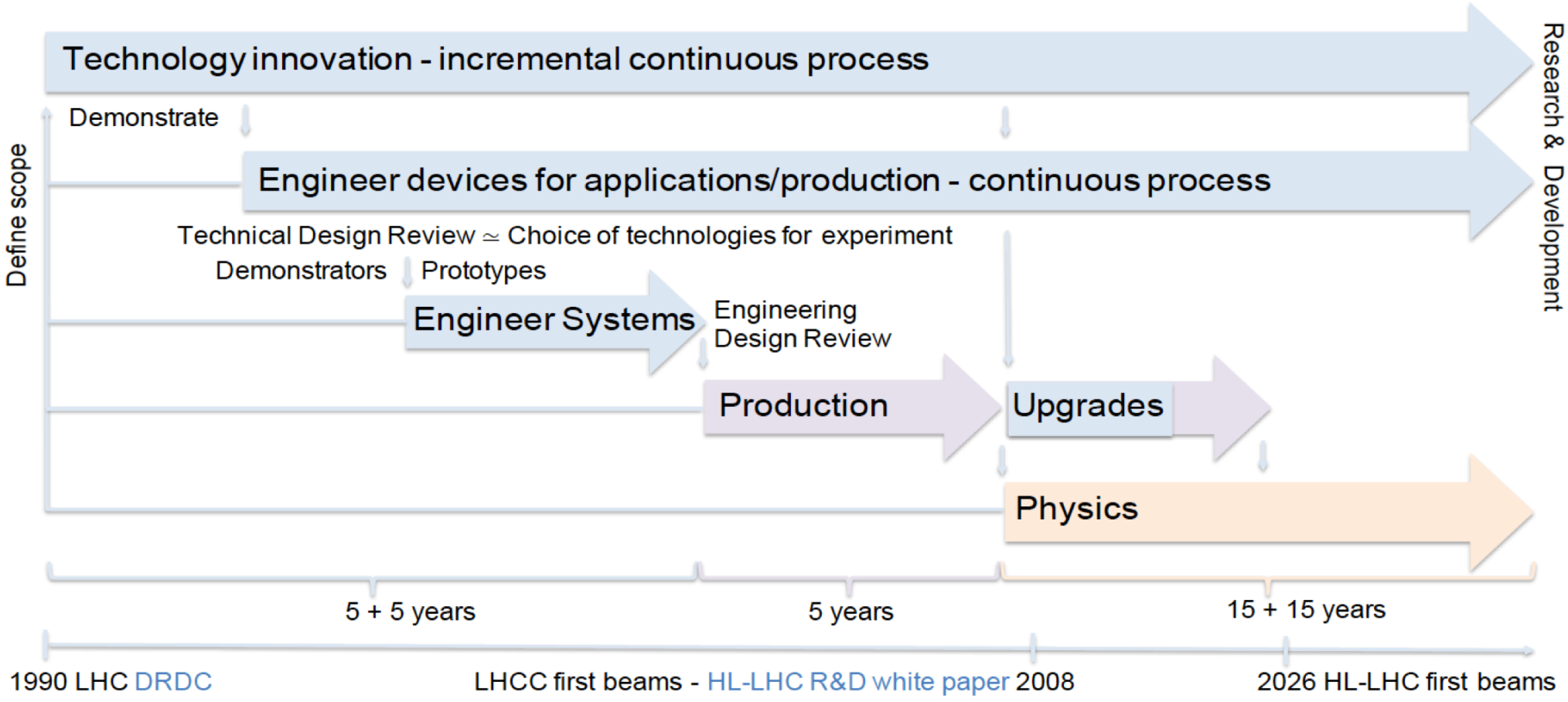
Quantum Machine Learning at CERN openlab

Sofia Vallecorsa
AI and Quantum Research - CERN openlab
CERN IT Department

06/11/2020


CPU seconds by Type
Prompt Data
Non-Prompt Data
LHC MC
HL-LHC MC
Analysis
THS06 * s
1600
1400
1200
1000
800
600
400
200
0
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
High Luminosity LHC
Projection (CMS experiment)

Typical HEP program timeline (R&D in blue)

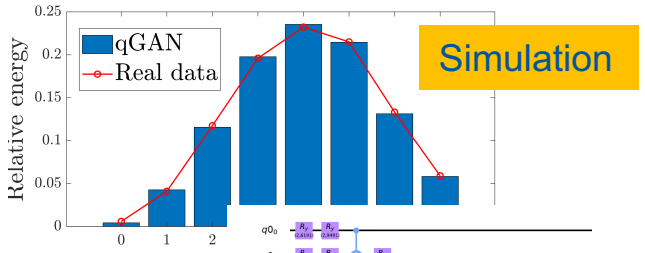


R&D timescale is ≃ 10 years, approved experimental programs are natural drivers for techniques and resources, technology innovation can suffer cycle/migration effects

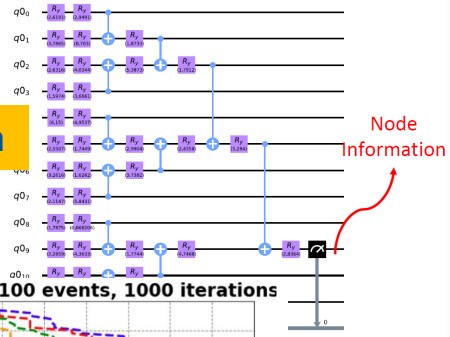
Arxiv:1910.11775

Quantum Expertise and Activities @CERN

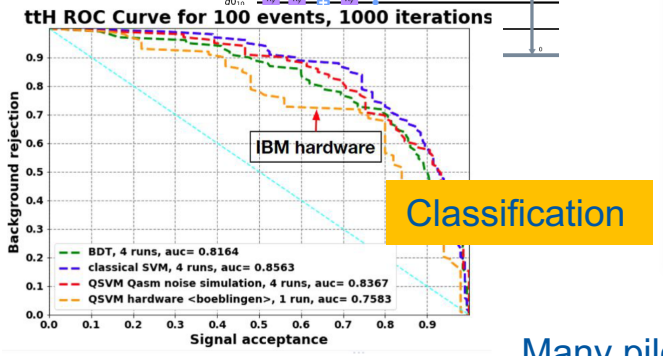
Computing



Simulation

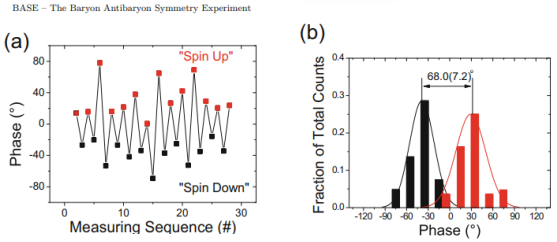


Reconstruction



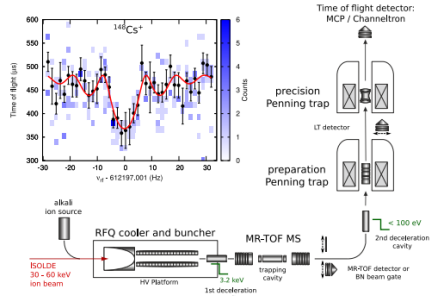
Classification

Sensing



https://doi.org/10.1140/epjst/e2015-02607-4

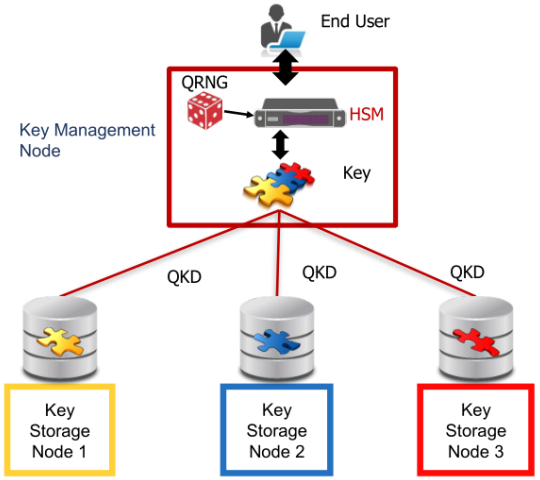
BASE phase-sensitive measurement of spin allowing very precise magnetic field drift measurements



https://doi.org/10.1088/1361-6471/aa5a20

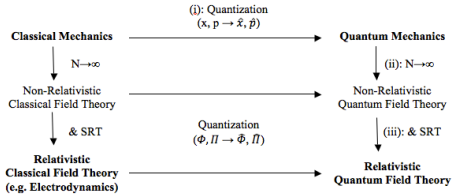
ISOLTRAP Mass-Spec

Communications

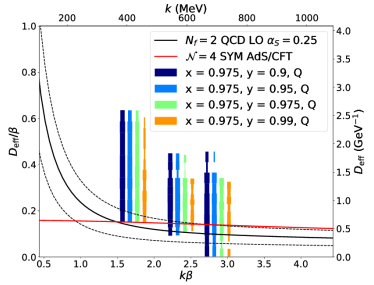


QKD node Secure Data

Theory



Quantum Field Theory



https://cds.cern.ch/record/2703396

Lattice QCD

Many pilot projects already started as part of the **CERN openlab quantum** programme (https://openlab.cern/quantum)

CERN Quantum Technology Initiative

Strategy

Coordination

R&D

Capacity building

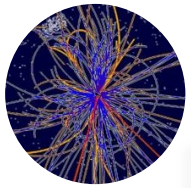
Joint HEP R&D Programme

CERN Management

QT Advisory Board (Member States)

QT Initiative Management

Sensing, Detectors R&D



Computing & Engineering

Communication

Simulation, Information Processing

Academic Programmes / Industrial Collaborations / Knowledge Transfer

High-Level Objectives

Computing

- Assess QC potential in HEP
- Build expertise in state-of-the-art software stack
- Develop quantum systems simulators
- Set up a distributed QC emulation platform

Communications

- Explore possible applications of QKD
- Comms+sensing for detectors?
- Explore Quantum Network/Internet

Sensing

- Ion/particle traps as computing and sensing devices
- Mass/charge/gravity sensors
- Quantum clocks
- Nanowires/nanodots for particle tracking/calorimetry
- Rydberg calorimetry

Theory

- Simulation of quantum systems
- Quantum gravity
- Information processing, error mitigation/correction strategies

QC Examples Applications

Quantum **Support Vector Machines** for ~~Higgs boson~~ **classification**

Pattern recognition

Quantum **Graph Neural Networks** for ~~particle trajectory~~ ~~**reconstruction**~~

Quantum **Generative Adversarial Networks** ~~for detector~~ **simulation**

Reinforcement learning

Quantum **Boltzman Machines** ~~for ALICE Grid optimisation~~

data

Hybrid **quantum-classical** ~~**tracking hits**~~ **embedding**

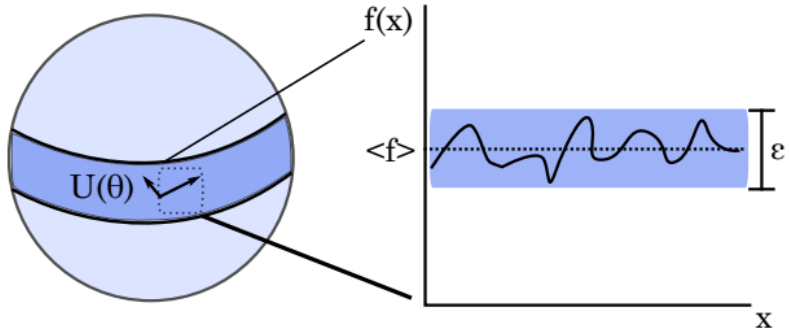
Collaborations with institutes (DESY, U. Aachen, INFN, U.Tokyo, U. Wisconsin, METU, P. U. Bucharest, U. Oviedo, …) and companies (IBM, Google, Intel, CQC, ...)

Quantum Machine Learning

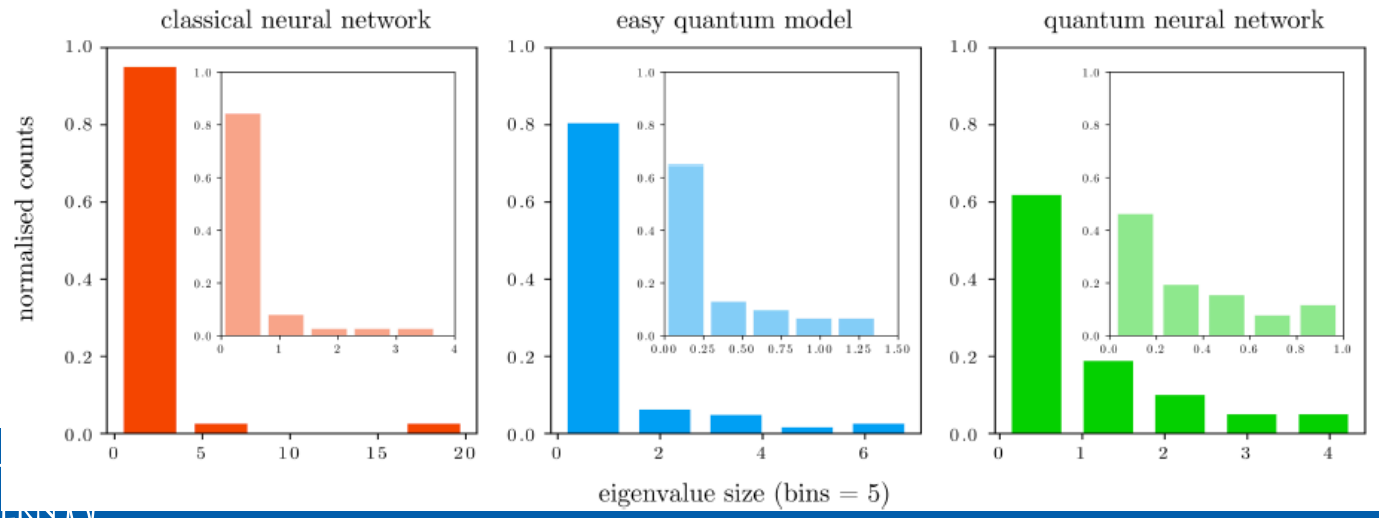
Exploit **quantum effects** for machine learning but

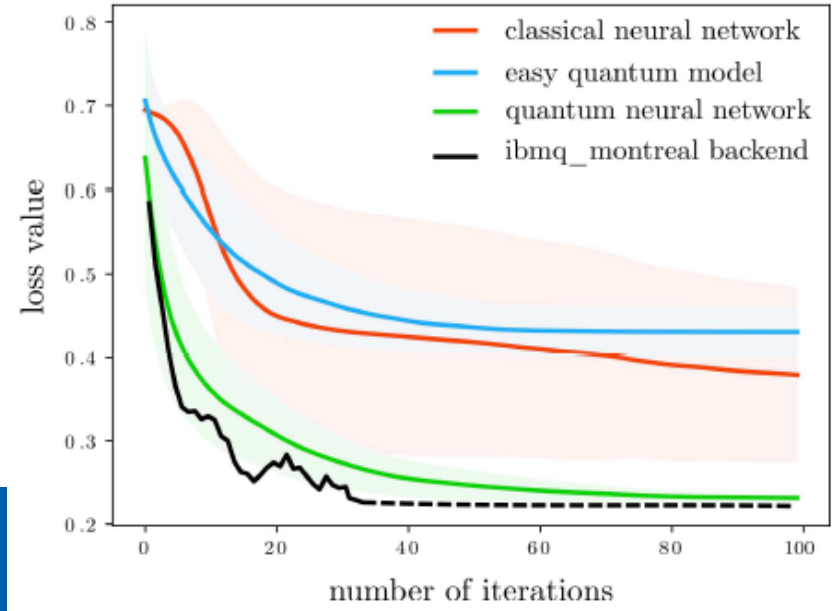
- How do we represent **learning rules**?
- **Convergence?**
- How do we address **data loading**?

Quantum Advantage?



J. McClean at al. arXiv:1803.11173





A. Abbas et al. arxiv:2011.00027

QML implementation

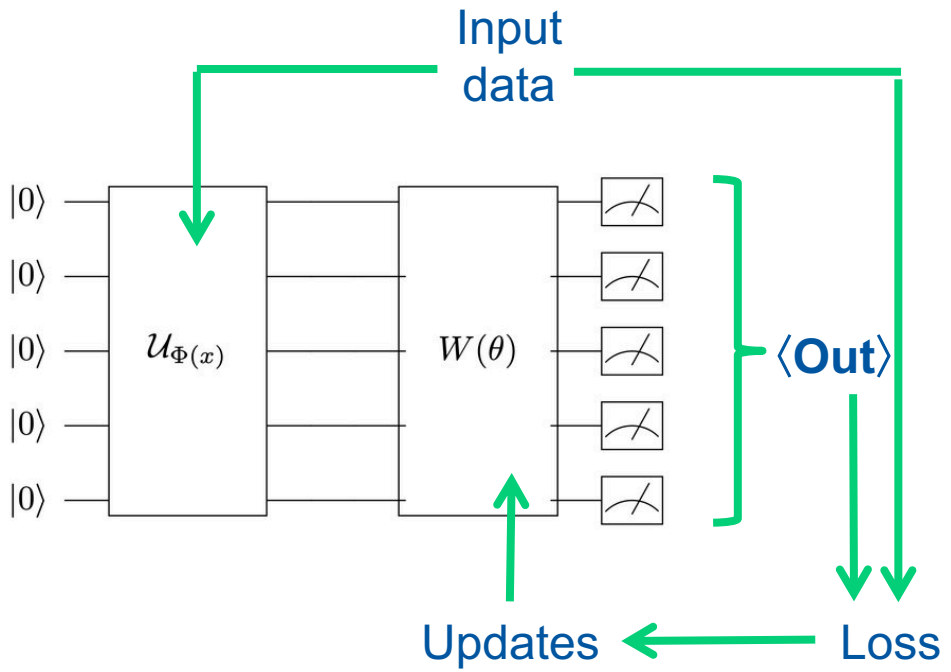
Variational approach:

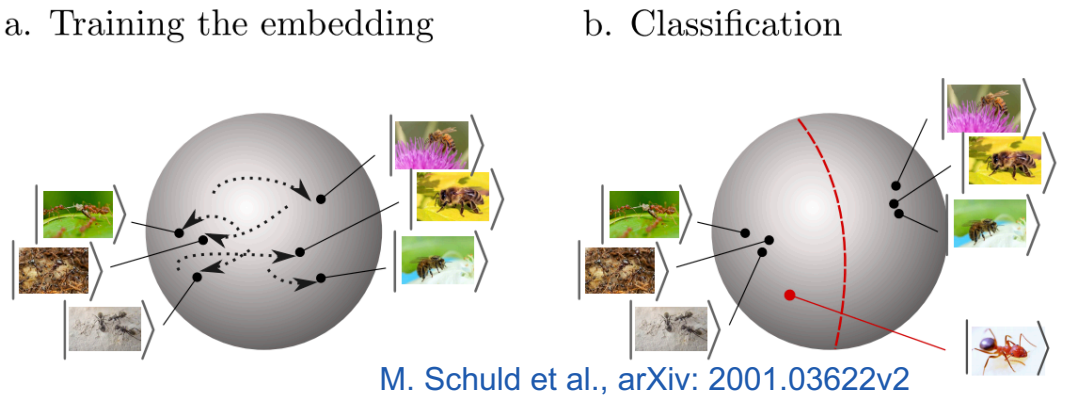
Quantum circuits with **adjustable continuous parameters**

Can use **gradient-free** methods

or **stochastic gradient-descent** optimisers

Data Embedding can be **learned**





M. Schuld et al., arXiv: 2001.03622v2

QC Examples Applications

QuantHEP Seminar https://quanthep-seminar.org/session/session-2/

Pattern recognition

Quantum **Graph Neural Networks** for ~~particle trajectory~~ **~~reconstruction~~**

Quantum **Generative Adversarial Networks** ~~for detector~~ **simulation**

Reinforcement learning

Quantum **Boltzman Machines** ~~for ALICE Grid optimisation~~

data

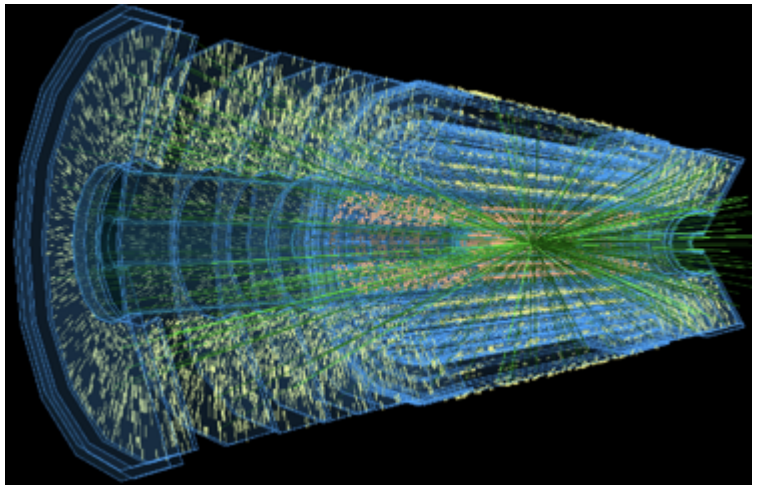
Hybrid **quantum-classical** ~~**tracking hits**~~ **embedding**

Collaborations with institutes (DESY, U. Aachen, INFN, U.Tokyo, U. Wisconsin, METU, P. U. Bucharest, U. Oviedo, …) and companies (IBM, Google, Intel, CQC, ...)

Charged particle tracking

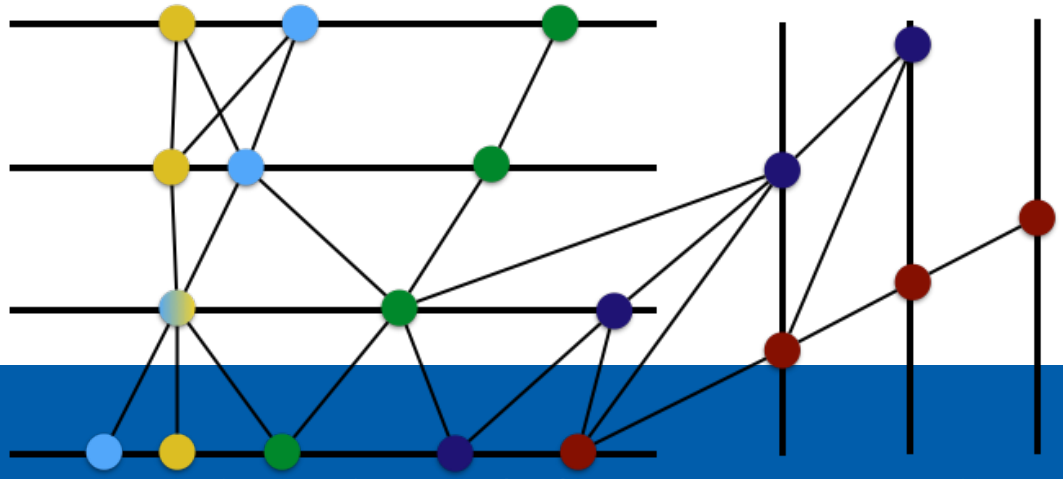
Hep.TrkX project introduces **Graph Neural Networks** for particle trajectory reconstruction

- Data as a **graph of connected hits**
- Connect plausibly-related hits using geometric constraints
- Full event embedding requires **large graphs** (~10^5 nodes)



HepTrkX GNN is a cascade of **Input**, **Edge** and **Node** Networks

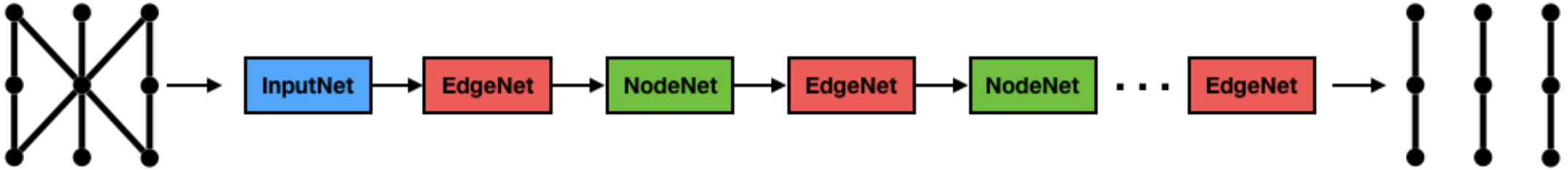
- **Edge network** outputs edge features, using the start and end nodes
- **Node network** classifies nodes using all connected nodes features on the previous and next layers

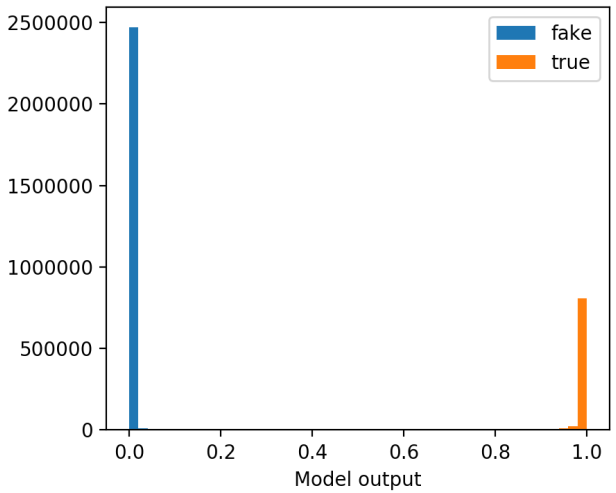


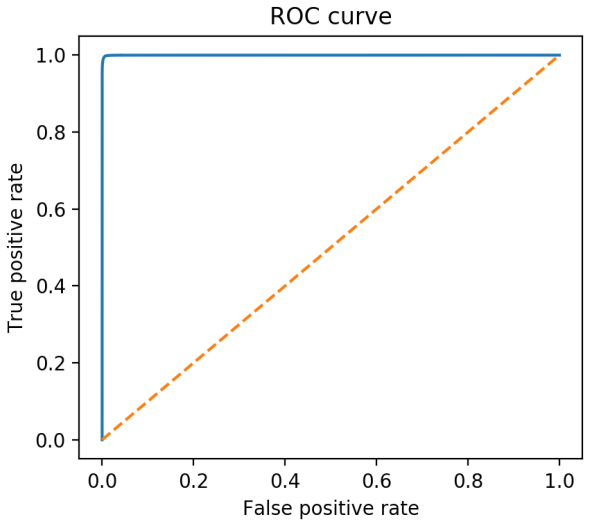
Now in the **ExaTrkX** project

arXiv:1810.06111, arXiv:2007.00149

GNN for particle tracking

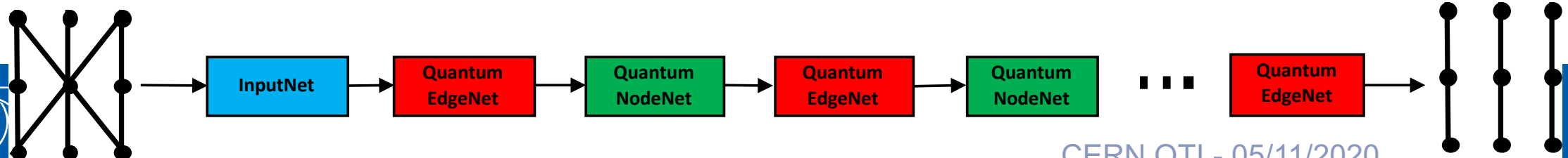






HEP.TrkX GNN Scores:
Purity: 99.5%
Efficiency: 98.7%
Overall Accuracy: 99.5% with 0.5 threshold

Quantum: replace Edge and Node networks with **quantum variational circuits**

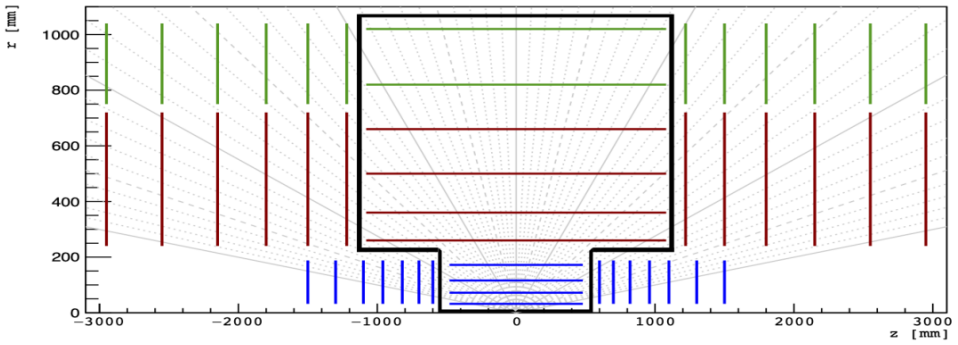


arXiv:2003.08126

Input data

TrackML challenge dataset

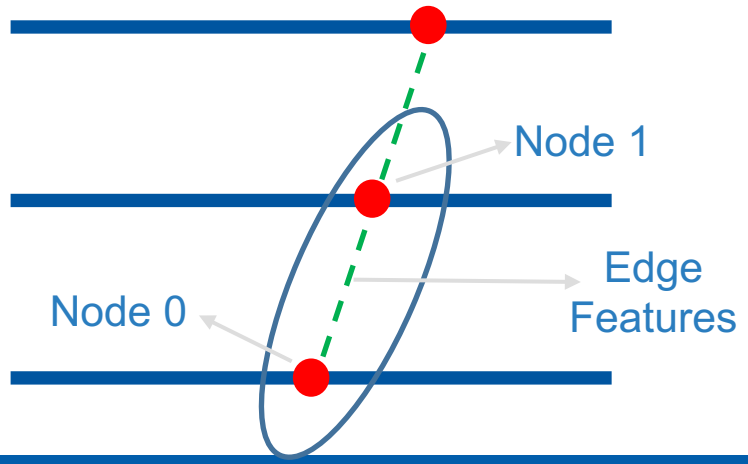
Preprocessing reduces track ambiguity and data dimensionality



10k collision events (200 soft QCD interactions) arXiv: 1904.06778

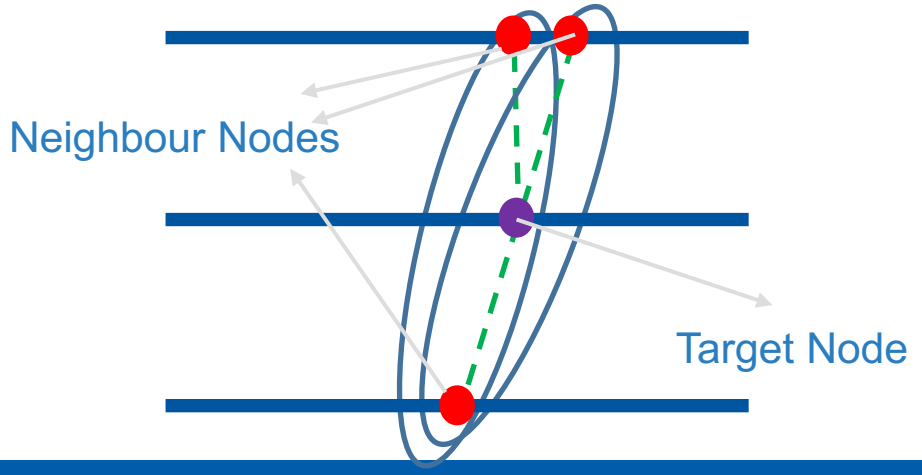
Edge Network Input

Node information of each edge



Node Network Input

Triplet node information

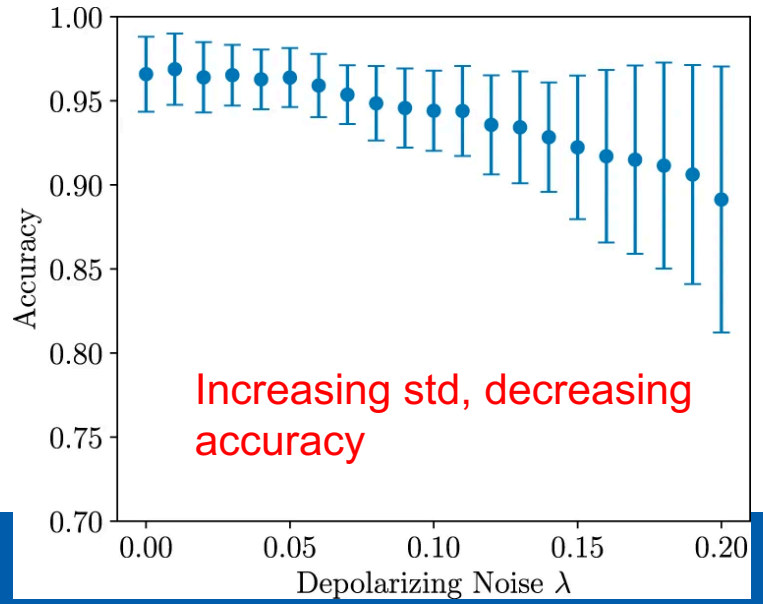


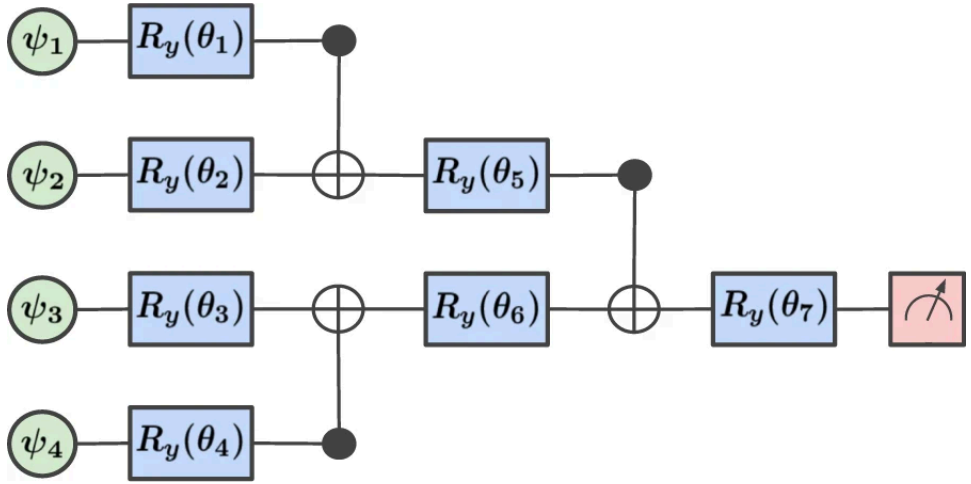
Hierarchical Quantum Classifiers

Hierarchical quantum circuits that can perform binary classification
- Tree Tensor Networks (TTN)
- MERA graphs (additional entangling, reproduces long-range correlations)

Robust against noise

Test TTN on the Iris dataset (IBM Q experience – 5 qubits)

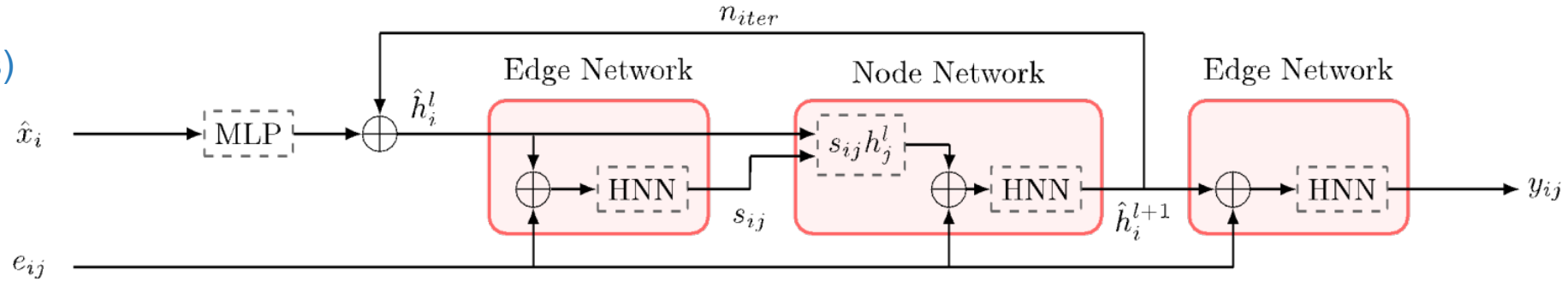




npj Quantum Information **4**, 65 (2018)

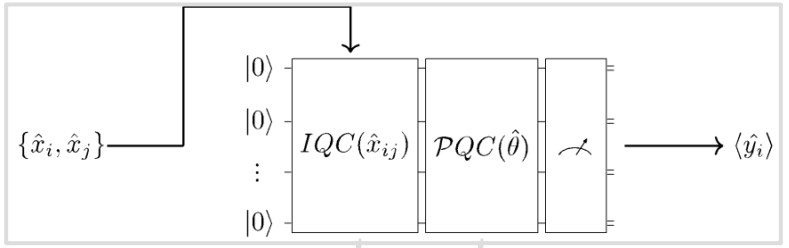
Quantum Graph Neural Network

Node information
(3D cylindrical coordinates)

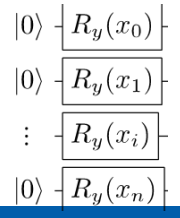


Graph connectivity matrix

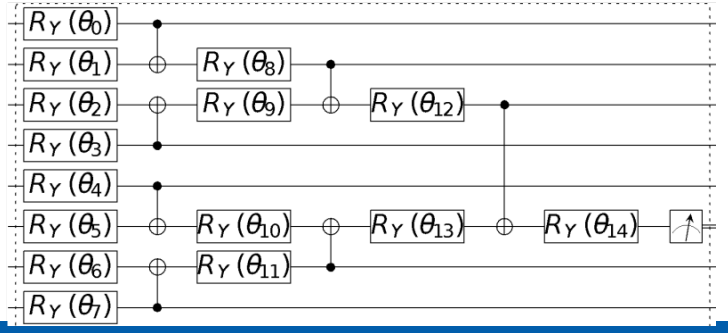
HNN (Hybrid Neural Network):



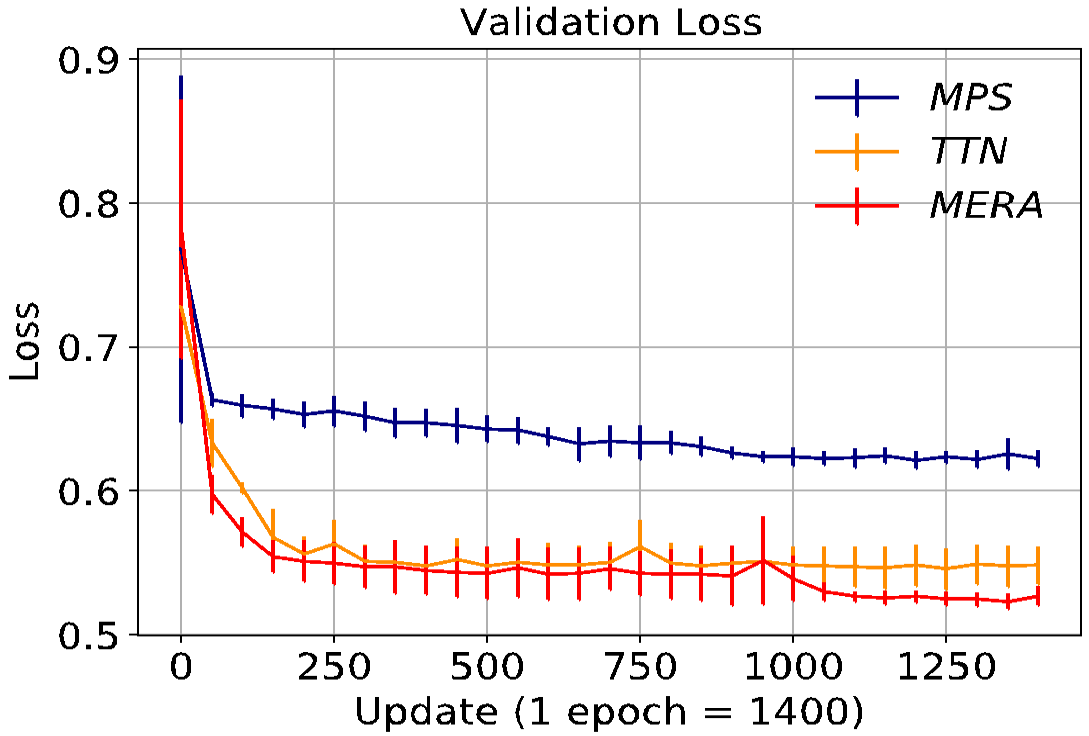
IQC
(Information Encoding
Quantum Circuit)

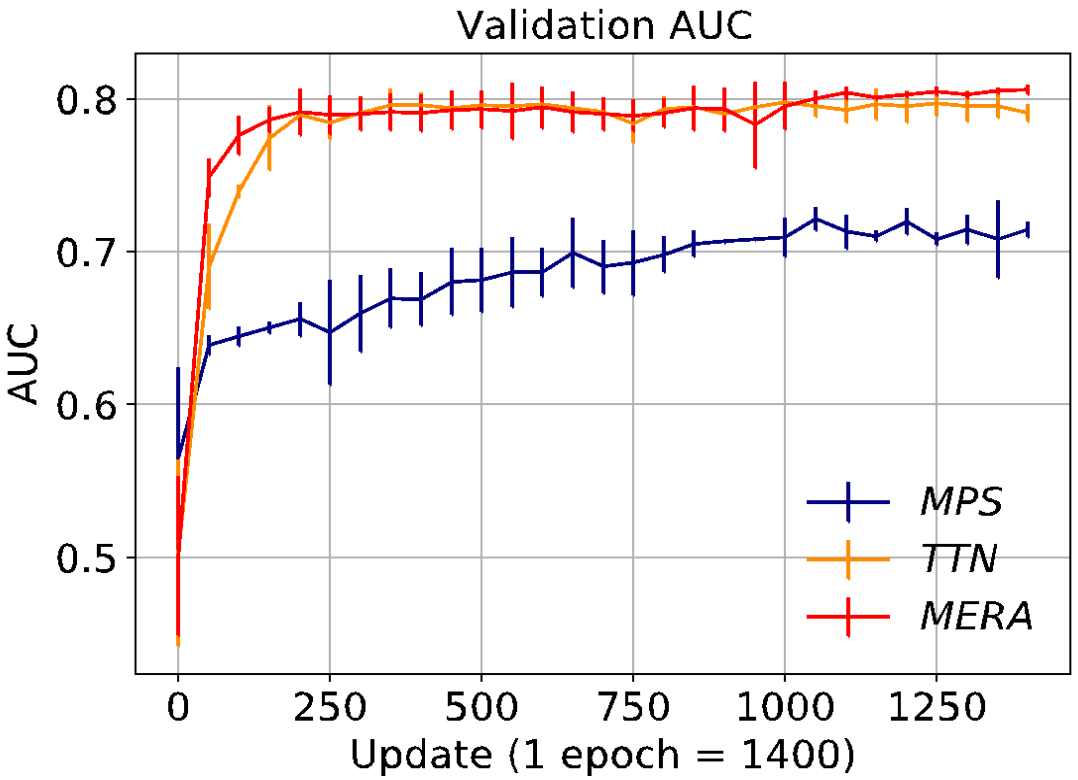


PQC (Parametrized Quantum Circuit)



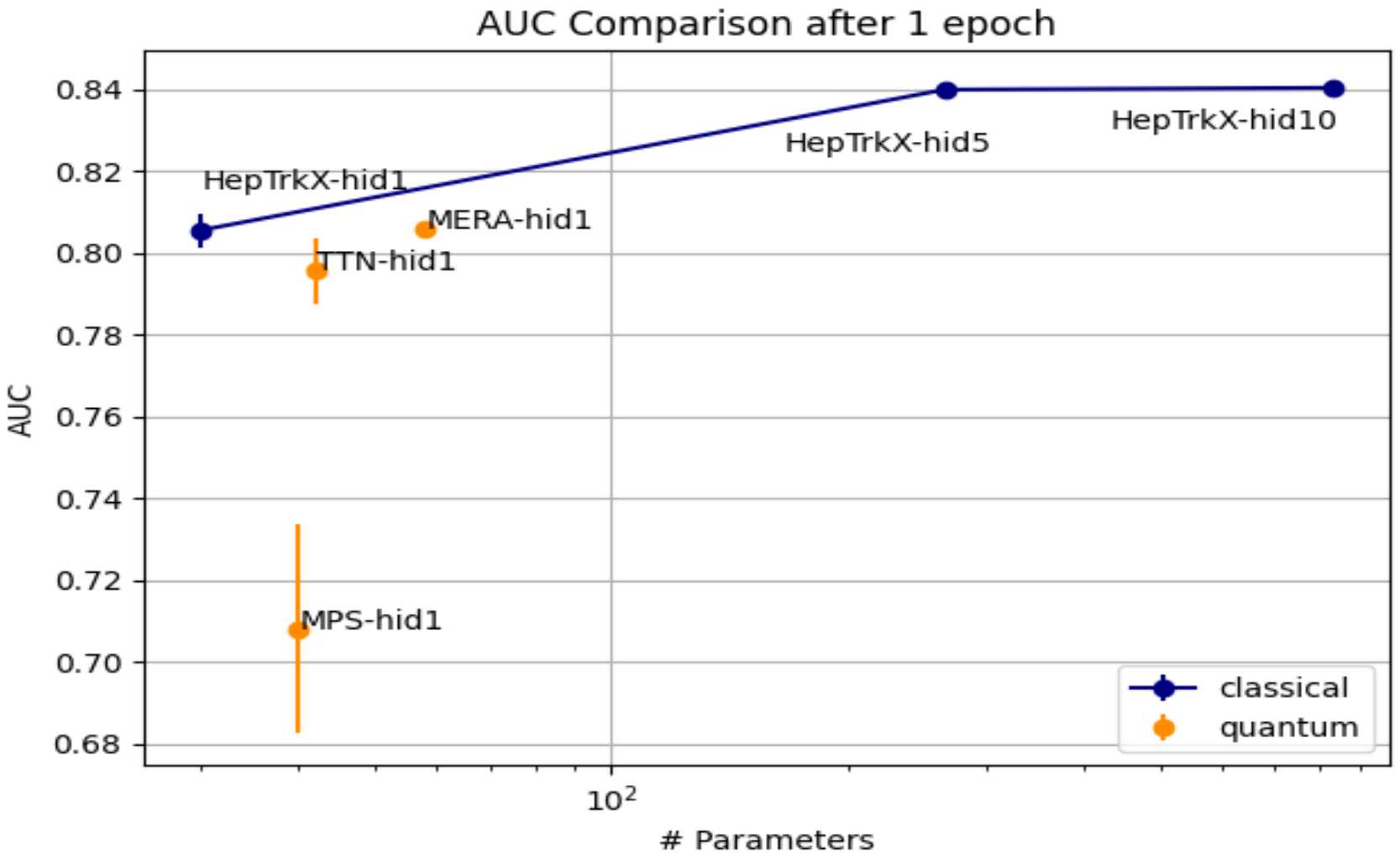
Training Results





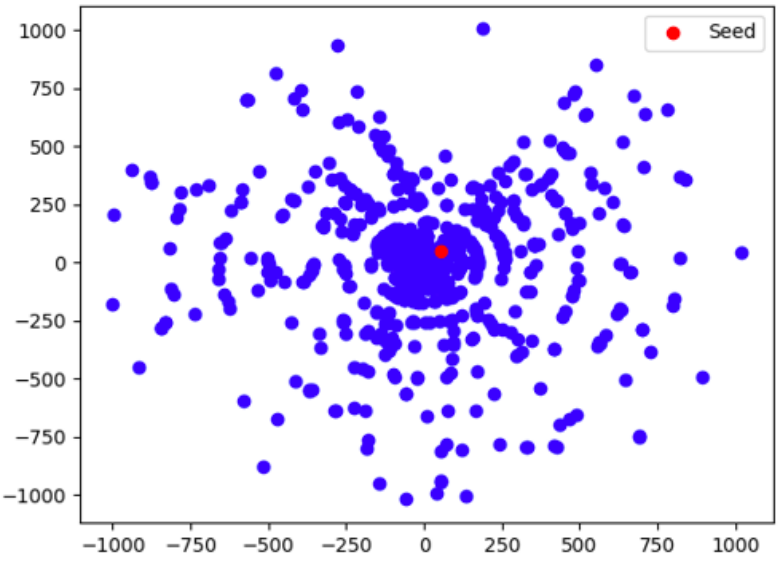
Training set: 1400 subgraphs, Validation set: 200 subgraphs,
using ADAM, binary cross entropy, lr = 0.03, shots =1000. **Hidden Dimension Size = 1**.

Comparison to HEP-TrkX

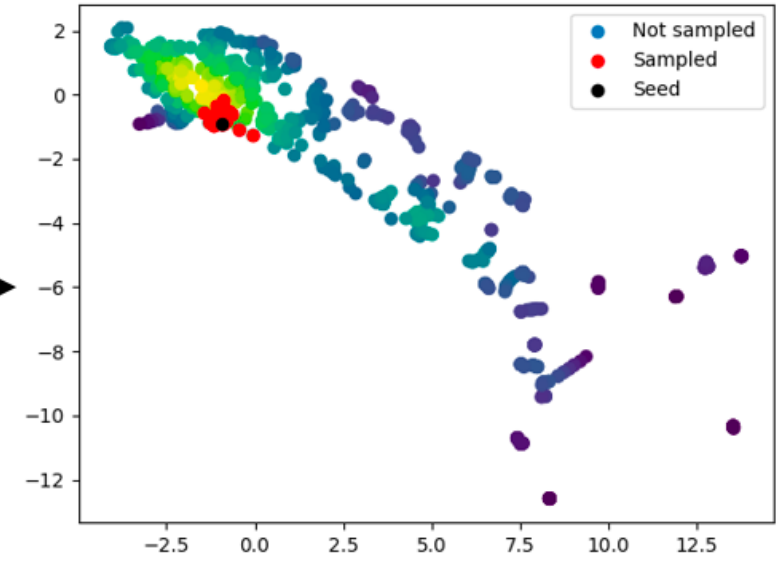


Learning the embedding

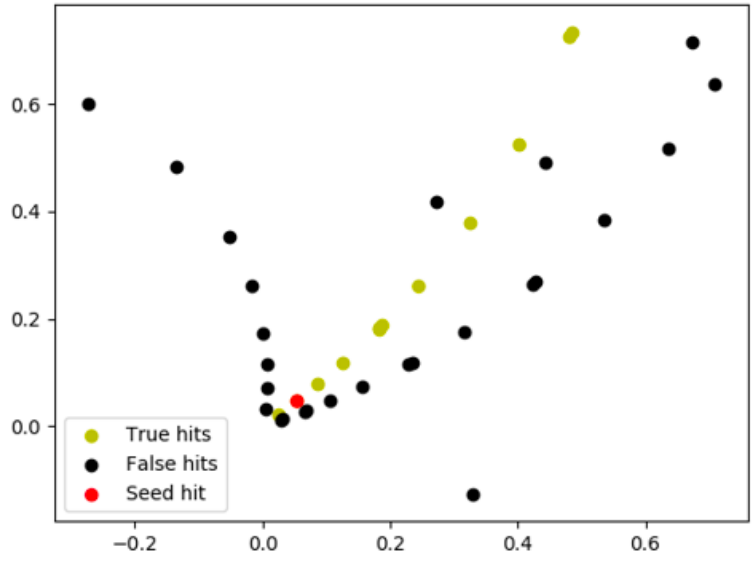
ExaTrkX: Improving accuracy by representing the input data in an embedded state



Original x-y space



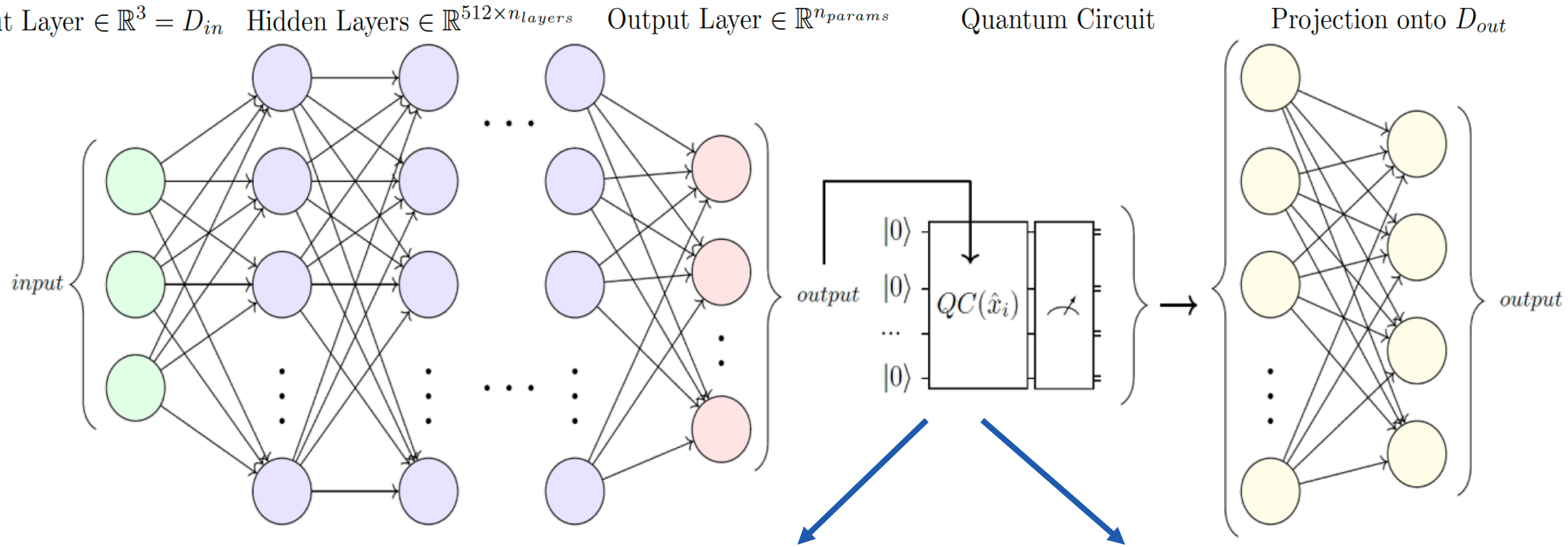
Embedded space:
Hinge-loss based MLP



Selected hits in original space

Hybrid embedding

C. Rieger
IOP Quantum 2020



Fixed quantum circuit

Fully trainable feature map

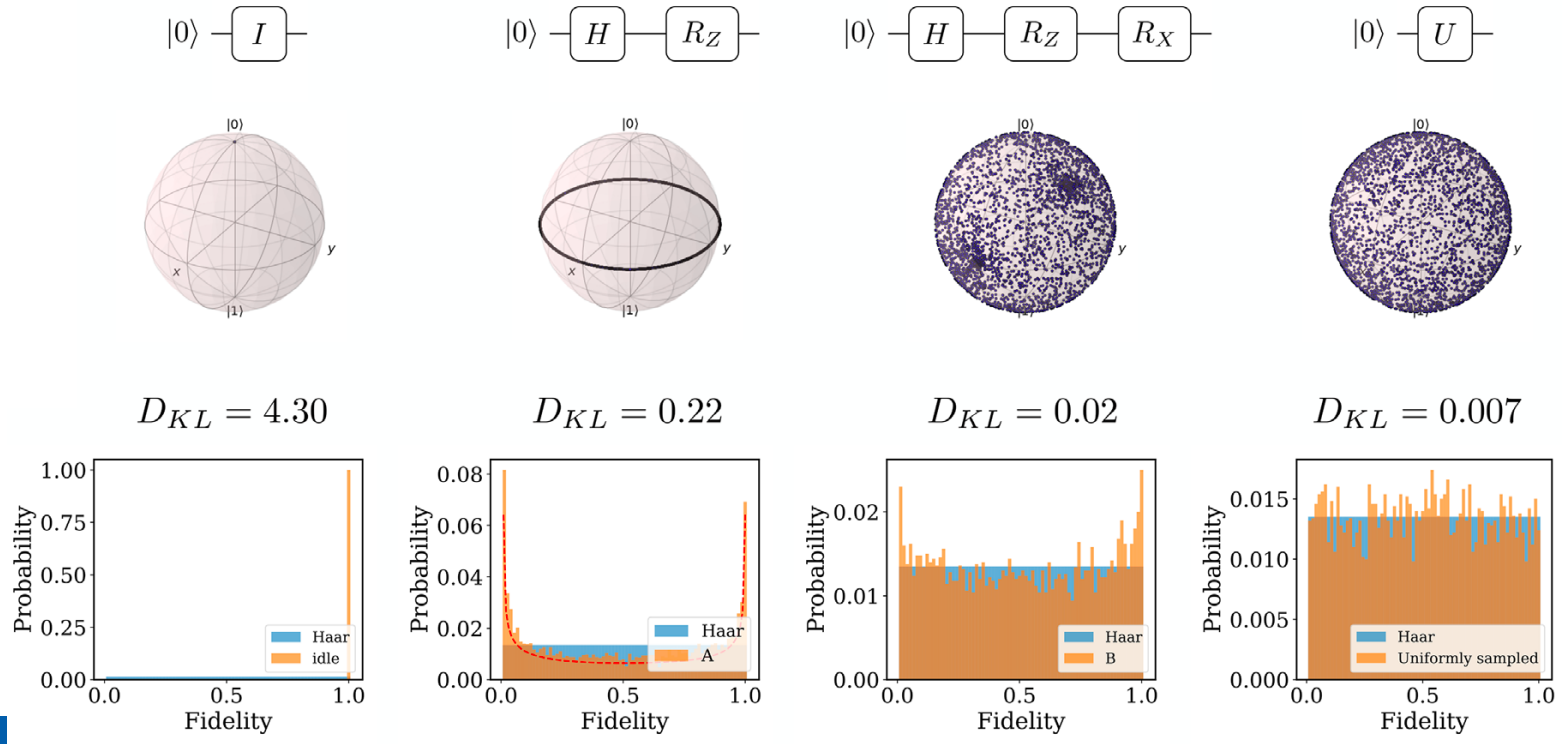
Choosing a quantum circuit?

Expressibility: ability to generate states that are well representative of the Hilbert space.

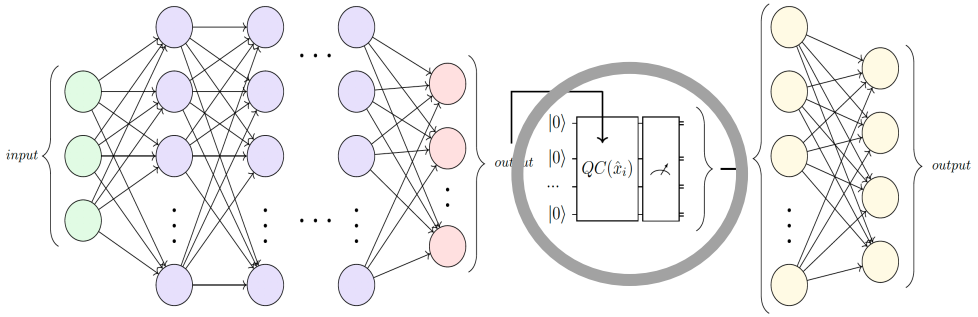
$$\mathrm{Expr} = D_{\mathrm{KL}}\left(\hat{P}_{\mathrm{PQC}}(F;\boldsymbol{\theta}) \,||\, P_{\mathrm{Haar}}(F)\right)$$

Entanglement : average Meyer–Wallach entanglement of states generated by the circuit.

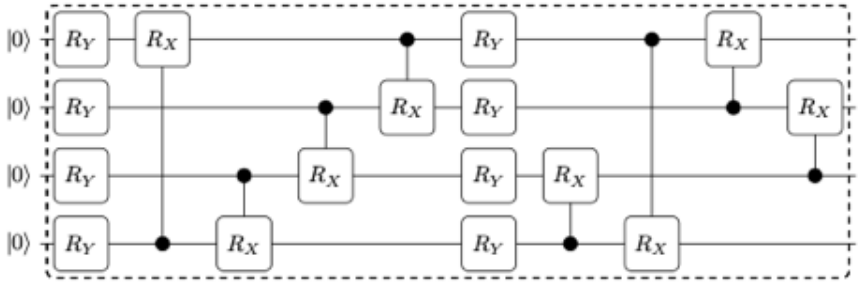
$$\mathrm{Ent} = \frac{1}{|S|}\sum_{\theta_i \in S} Q\left(|\psi_{\theta_i}\rangle\right)$$



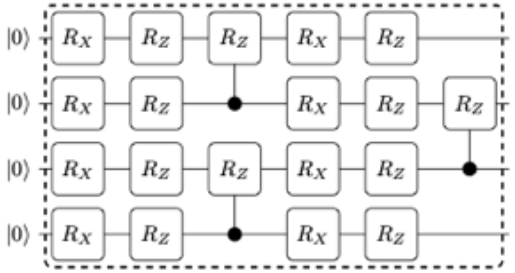
Example circuits



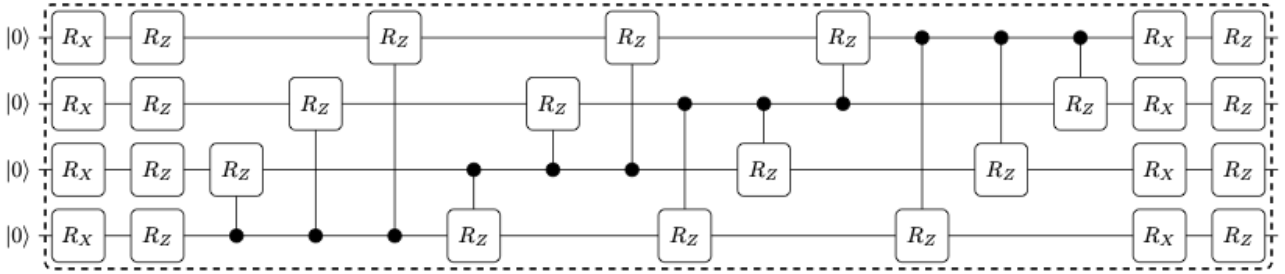
Circuit number	Number of parameters $n_{parameters}$	Entanglement (the higher the better) [1]	Expressibility (the lower the better) [1]	Average training time per batch (= 100 samples)
5	28	0.29	0.05	37 s $\pm$ 8 s
7	19	0.21	0.10	20 s $\pm$ 4 s
11	12	0.54	0.13	14 s $\pm$ 4 s
14	16	0.49	0.02	16 s $\pm$ 4 s



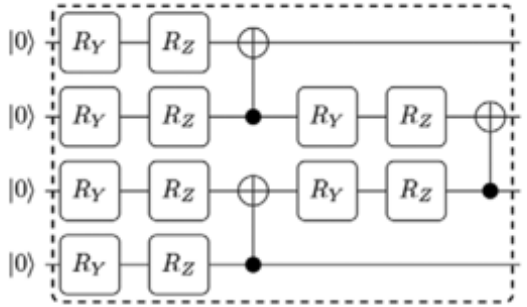
Circuit 14



Circuit 7



Circuit 5

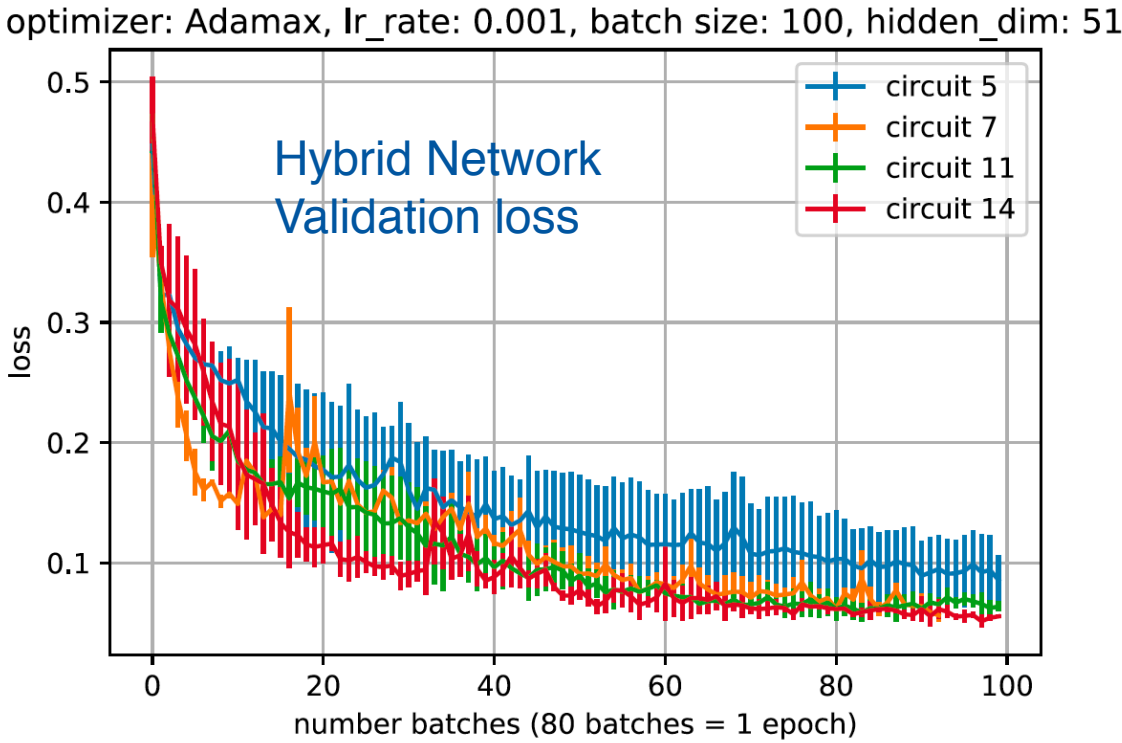


Circuit 11

[1] S. Sim at al, http://dx.doi.org/10.1002/qute.201900070

Fixed quantum circuit

Circuit number	Number of parameters $n_{parameters}$	Entanglement (the higher the better) [1]	Expressibility (the lower the better) [1]	Average training time per batch (= 100 samples)
5	28	0.29	0.05	37 s ± 8 s
7	19	0.21	0.10	20 s ± 4 s
11	12	0.54	0.13	14 s ± 4 s
14	16	0.49	0.02	16 s ± 4 s



Long training times for simulation: O(days)

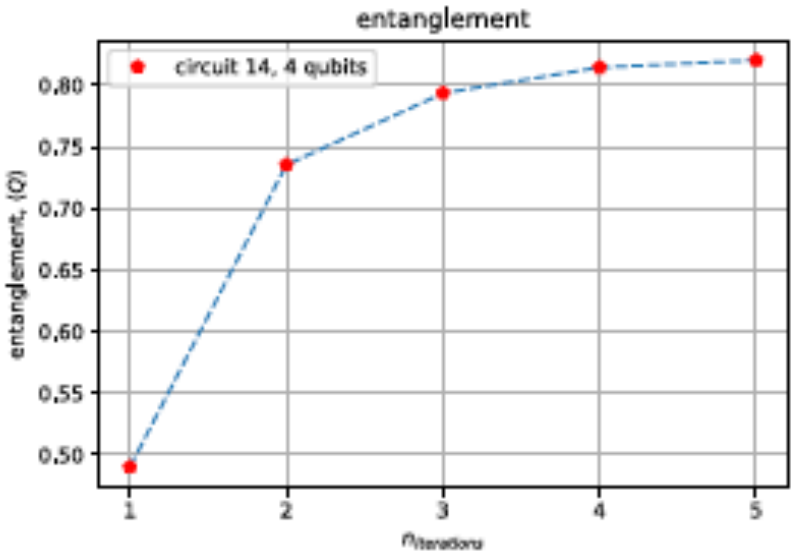
Effect of parameter initialization and circuit depth

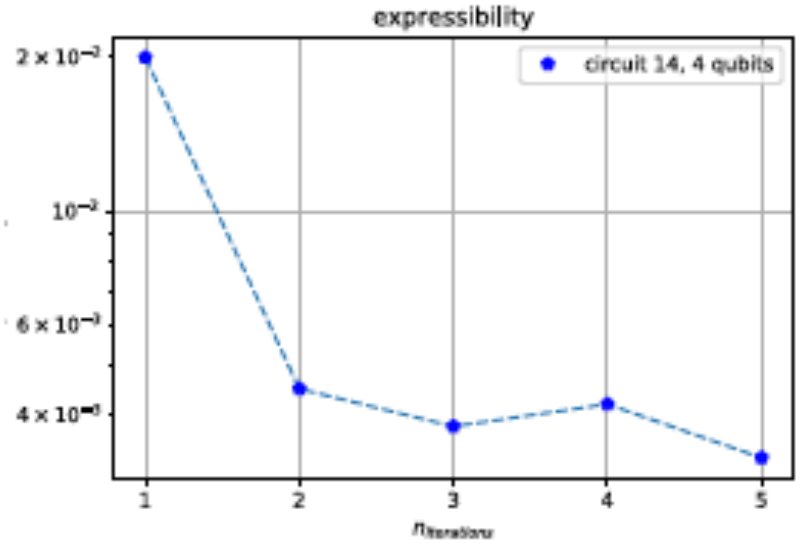
[1] Sukin Sim at al, http://dx.doi.org/10.1002/qute.201900070

Fixed quantum circuit (II)

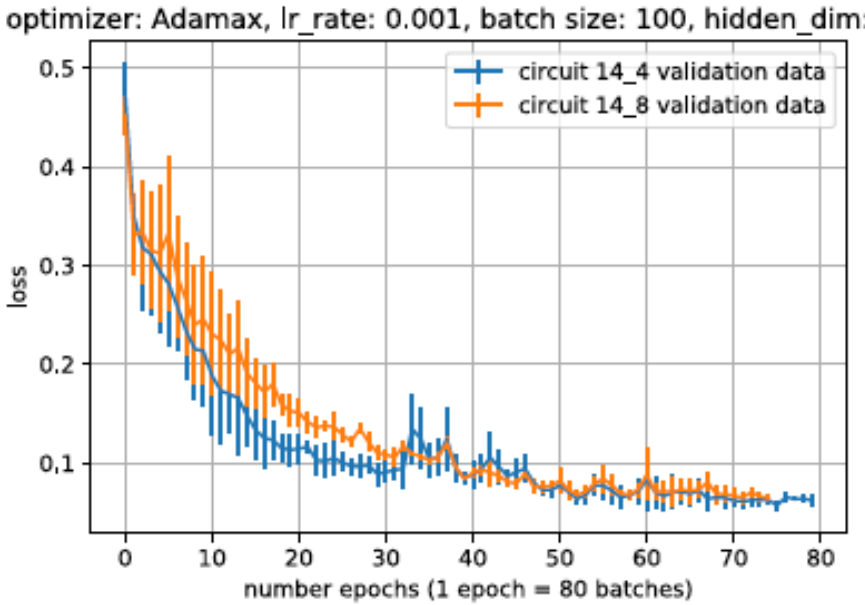
C. Rieger
IOP Quantum 2020

Increasing depth

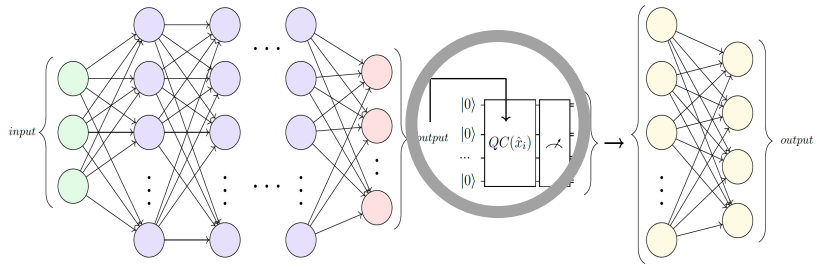


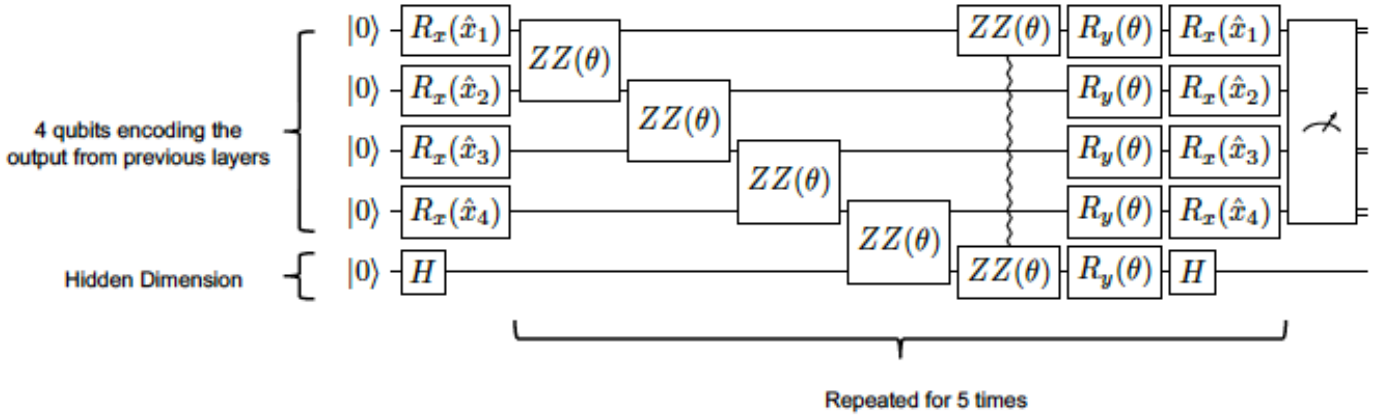


Increasing N_{qubits}



Trainable feature map





Quantum circuit	Number of parameters $n_{parameters}$	Entanglement (the higher the better)	Expressibility (the lower the better)	Average training time per batch (= 100 samples)
QFM ($n_{iteration}$ = 5)	74	0.79	0.003	5 min 38 s ± 8 s
14 ($n_{iteration}$ = 1)	16	0.50	0.02	16 s ± 4 s

Generative Models

$$V = \min_G \max_D \underbrace{\mathrm{E}_{\boldsymbol{x}\sim p_{real}(x)}[\log D_\theta(x)]}_{\text{Likelihood to assign correct label to real samples}} + \underbrace{\mathrm{E}_{\boldsymbol{z}\sim p_z}[\log(1 - D_\theta(G_\varphi(\boldsymbol{z})))]}_{\text{Likelihood to assign correct label to generated samples}}$$

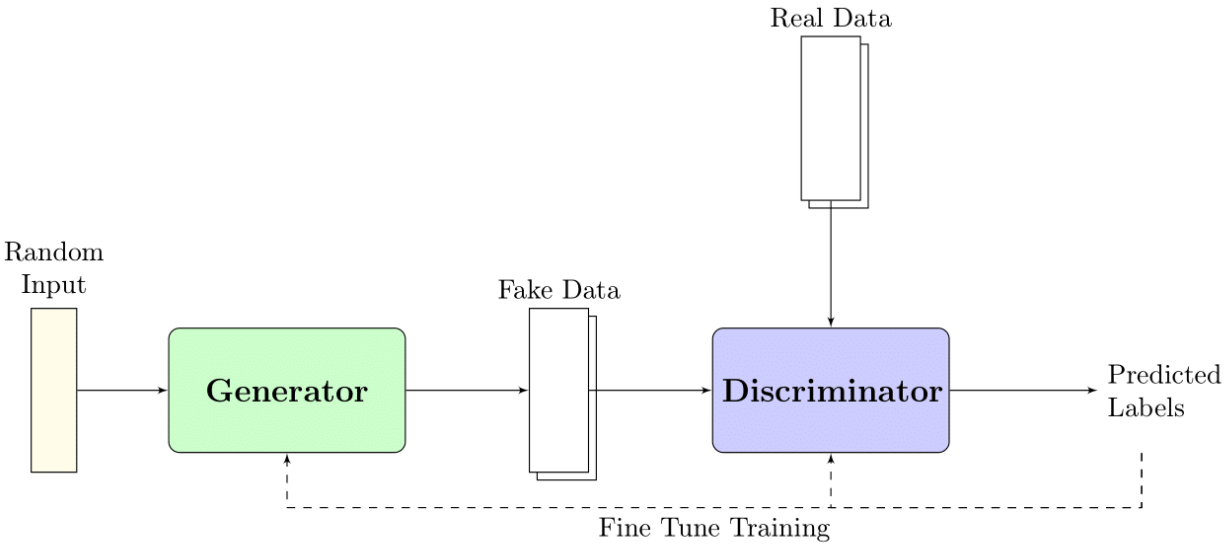
Classical Generative Models can replace Monte Carlo simulation

- Quantum Generative Models might have **larger representational power**

Useful in different applications

Generative Adversarial Networks:

- Generator : Generates a fake output from a random noisy input
- Discriminator : Classify fake and real data
- Vanilla GAN: play minmax game with value function V



S. Y. Chang
IOP Quantum 2020

Quantum Generative Models

Classical use case 3DGAN prototype for calorimeter simulation

Detector output interpreted as a 3D image.

Quantum GAN investigations

Down sample 3DGAN use case to **manageable number of pixels**

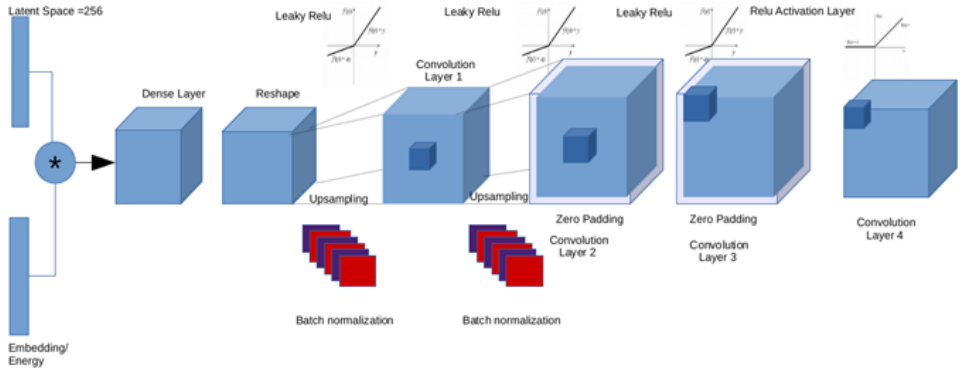
Use **compressed data representation** in quantum states.

Qubits or **Continuous Variables**

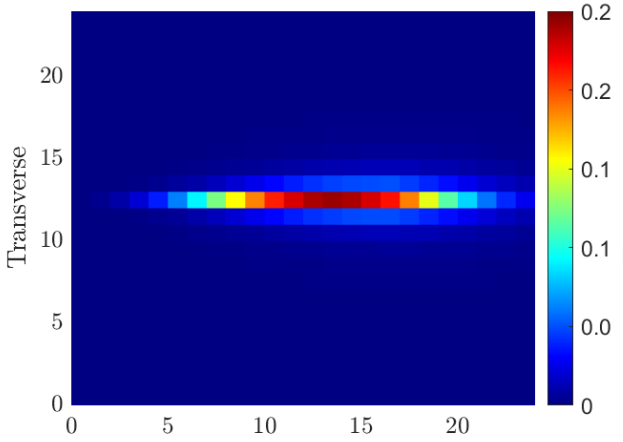
Different **hybrid** classical-quantum combinations

3DGAN generator

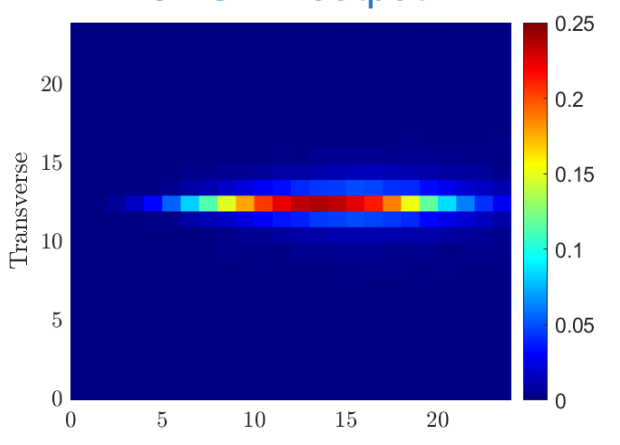
https://doi.org/10.1051/epjconf/201921402010



Real image



3DGAN output



Hybrid Classical-Quantum GAN

IBM qGAN can load probability distributions in quantum states

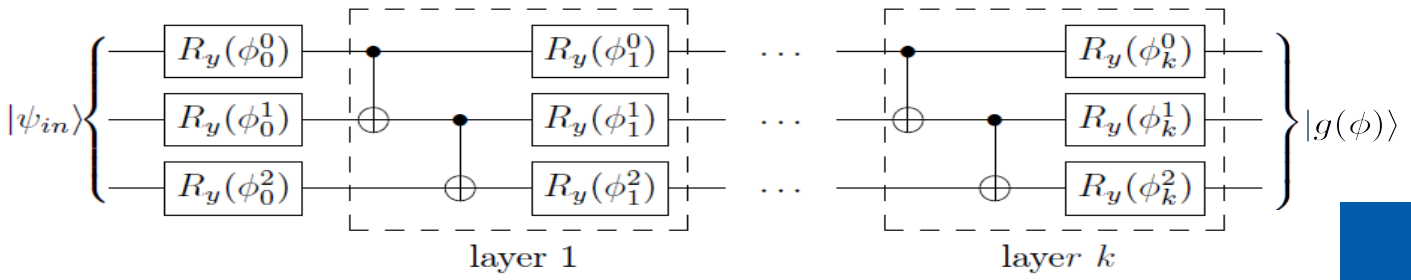
Simplify dataset: extract 1D energy profiles from 3DGAN images

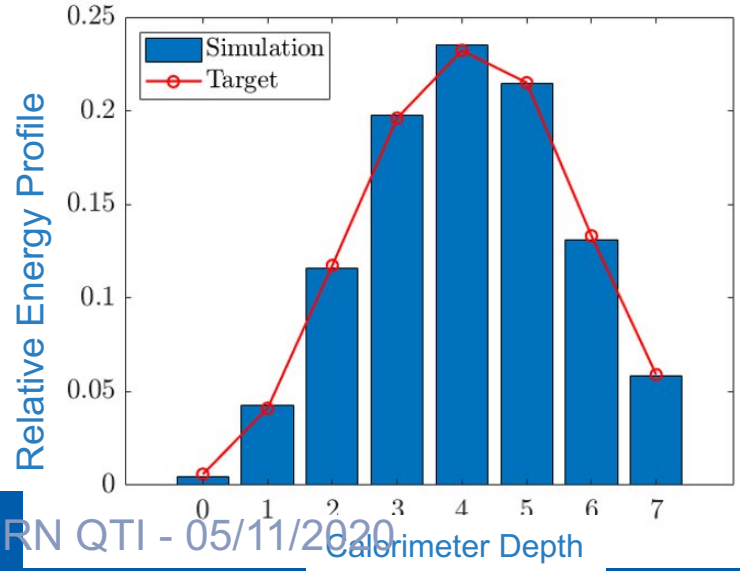
- 2^n classical pixels expressed by n qubits
- Probability of getting state $|k\rangle$ = (Relative) Energy at pixel k

Train a hybrid classical-quantum GAN to generate **few-pixels image**

- **Classical Discriminator (pyTorch):** 512 nodes + Leaky ReLU → 256 nodes + Leaky ReLU → single-node + sigmoid
- **Quantum Generator (Qiskit): 3 R_y layers**

https://doi.org/10.1038/s41534-019-0223-2





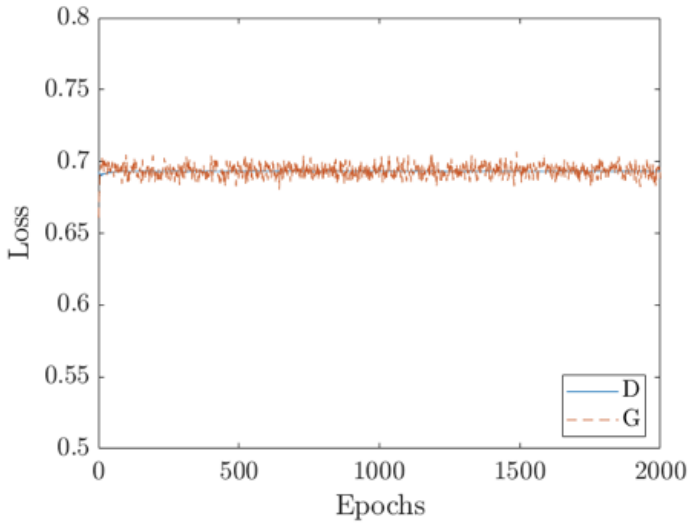
qGAN for calorimeter simulation

Low resolution 2D energy profile

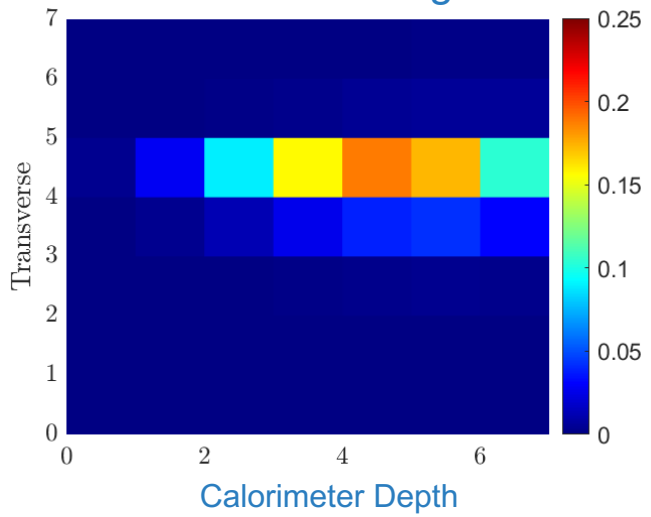
Increase generator and discriminator depth (6 qubits – 64 pixels)
- Realistic 2D images
- Stable training losses

IBM qGAN **is not a real generative model**
- It uses the adversarial training approach to **embed** image probability distribution in a quantum circuit

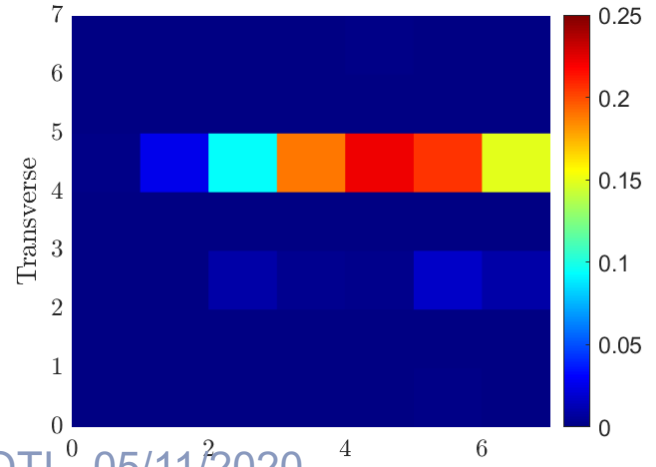
Need a way to sample single images



Real image



qGAN image



Extending the qGAN model

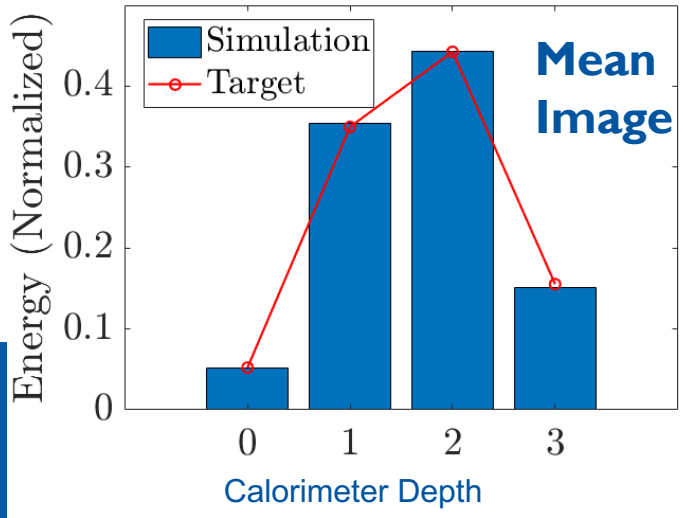
Collaboration with Cambridge Quantum Computing

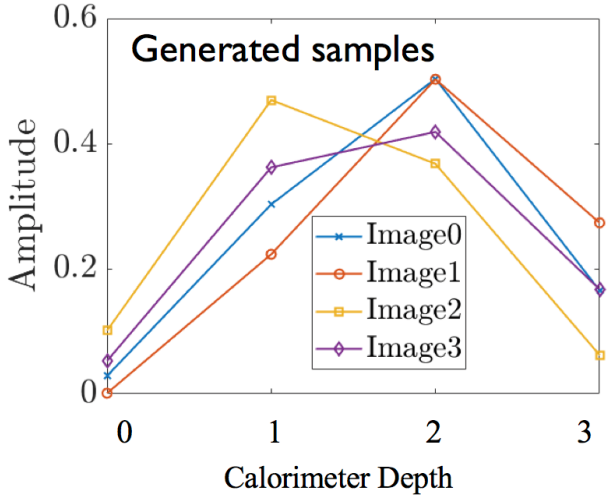
Two-steps quantum generator to learn the average distribution and sample images from it

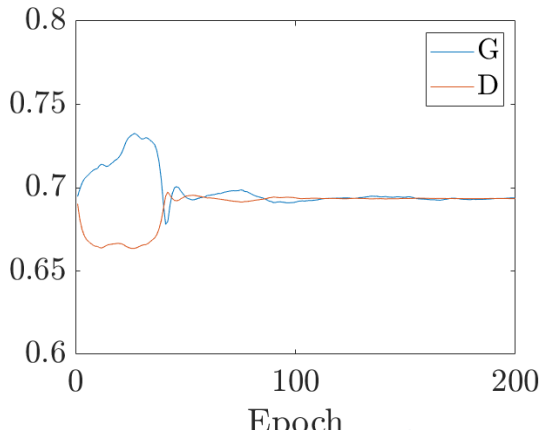
- Ry variational form implemented using **qiskit & t|ket**
- Uniform parameter initialization
- AMSGRAD optimizer with lr = 10^{-4}, lr2 = 10^{-3}

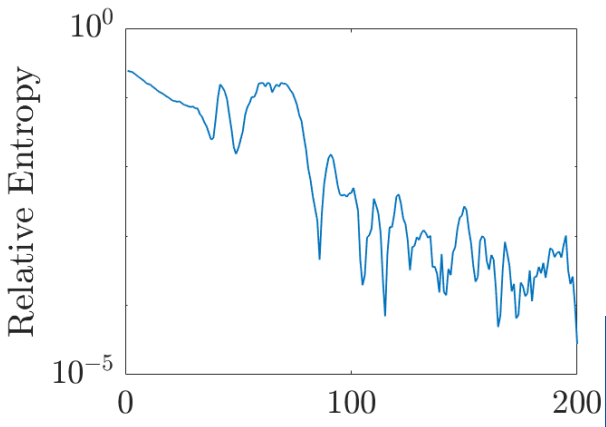
Classical discriminator (pyTorch) 4 nodes → 512 nodes → 256 nodes → 1 node

- Leaky ReLu between hidden layers + sigmoid
- AMSGRAD optimizer + Gradient penalty for stability and convergence









S. Y. Chang
IOP Quantum 2020

Photonic QC qGAN

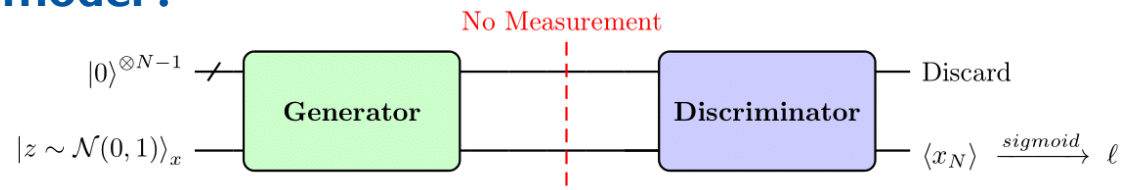
Use the same 1D electron energy deposition pattern in calorimeter volume

Image data encoded as a displaced state $|\boldsymbol{x}\rangle = \bigotimes_{i=1}^{N} |x_i\rangle = \bigotimes_{i=1}^{N} D_i(x_i)|0\rangle$

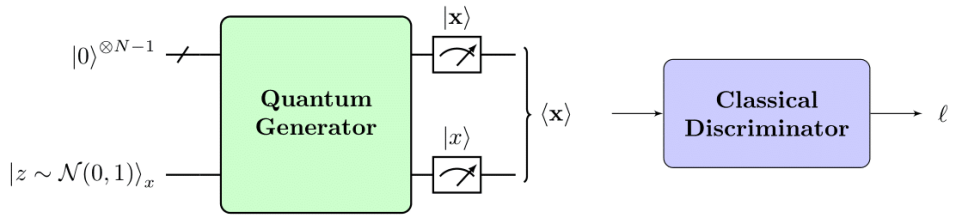
Fundamental information-carrying units = Qumodes

→ Work with **continuous variables**

Fully Quantum model :



Hybrid model :



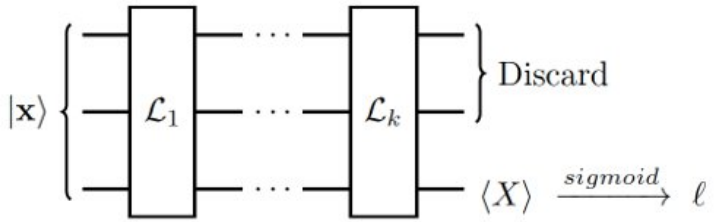
https://doi.org/10.1103/PhysRevLett.82.1784

S. Y. Chang
IOP Quantum 2020

A simple classifier

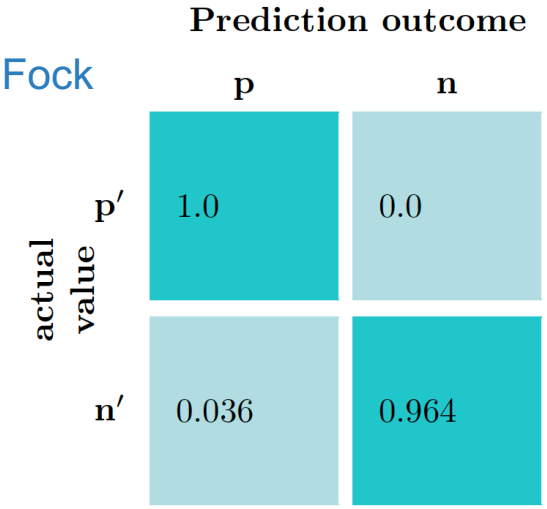
Test quantum discriminator network n=3 depth =3

StrawberryFields for network implementation.
Pennylane for autodifferentiation

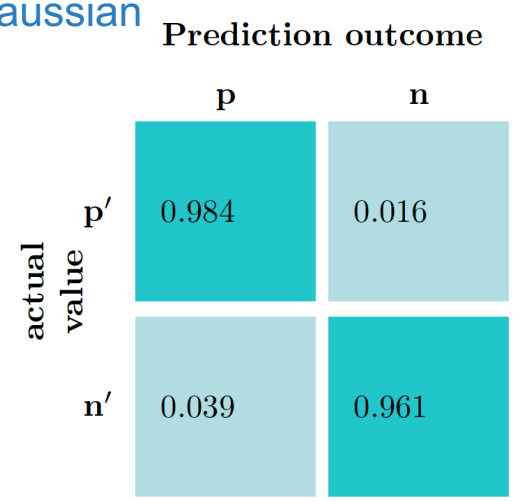


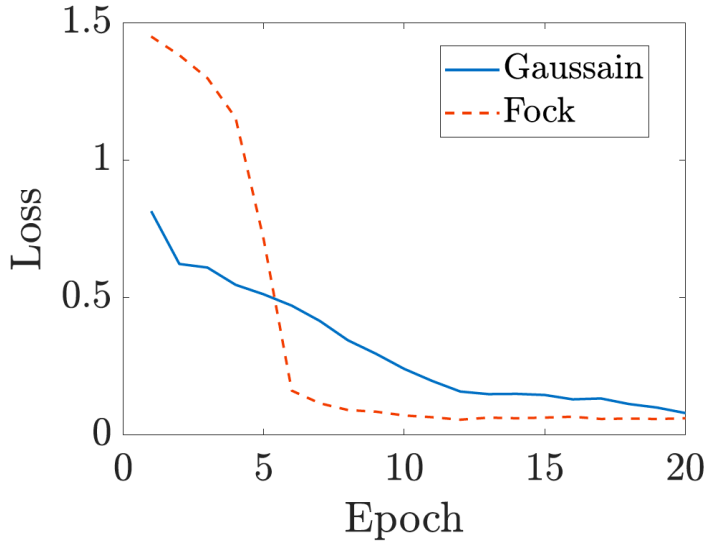
Compare 1D energy profile to flat distribution

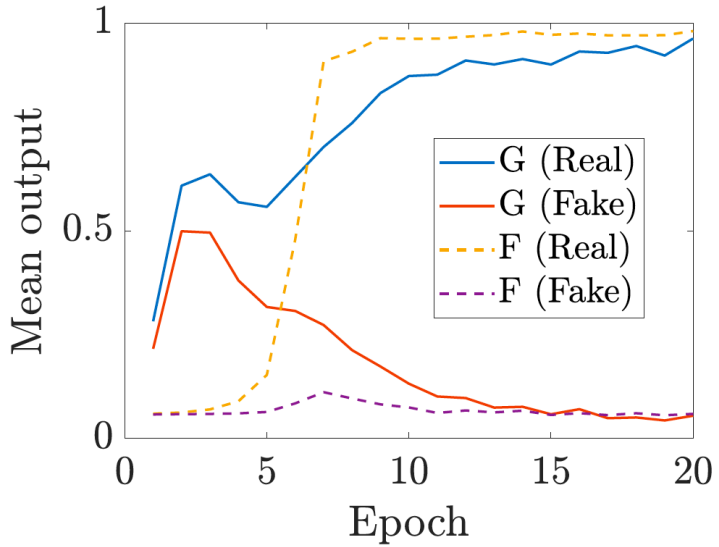
Fock



Gaussian





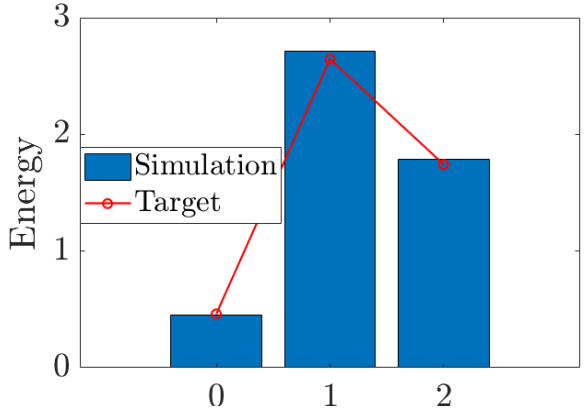


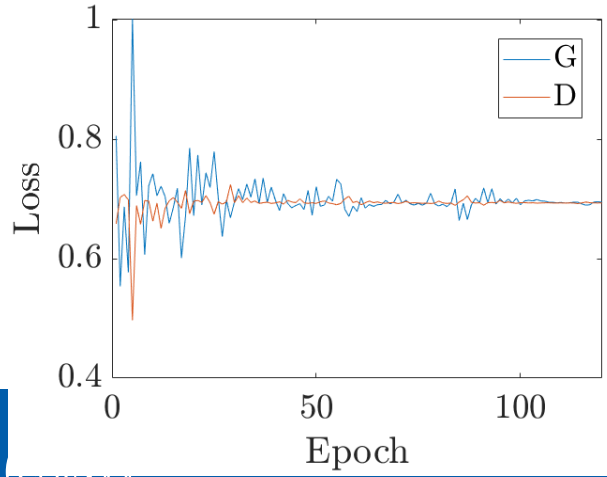
https://doi.org/10.1103/PhysRevLett.82.1784

cvGAN preliminary results

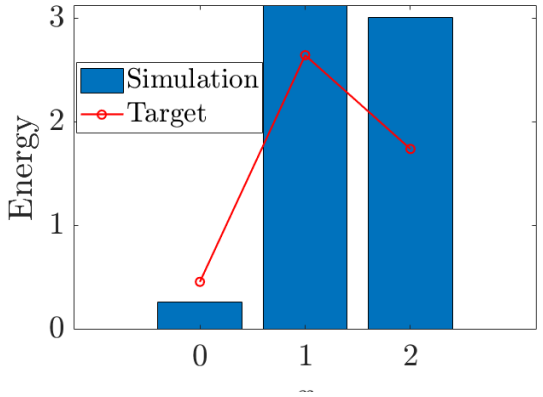
WORK IN PROGRESS

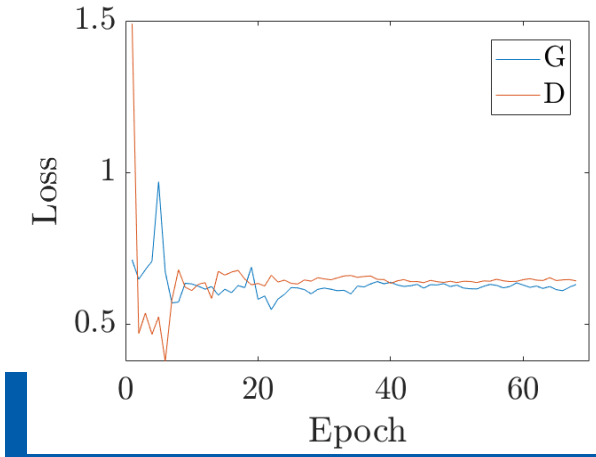
Hybrid



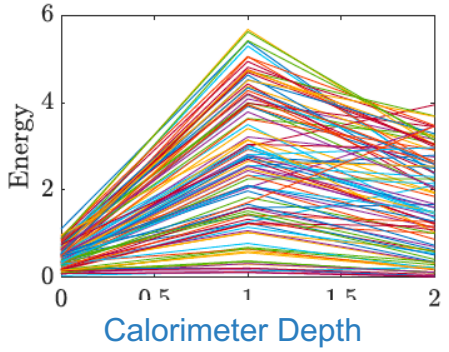


Fully Quantum

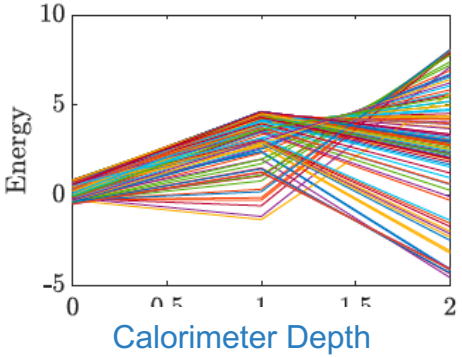




Real



Fully quantum qGAN



Fully quantum approach doesn't converge

- Mode collapse

Training process is time-consuming and limited by **computing power**

- multiprocessing with 50 cores yields
  - **160 min/epoch (quantum)**
  - **80 min/epoch (hybrid)**

QC Simulation Platform

Enable building skills and starting R&D work, both as a preparation to real H/W and to explore “quantum-inspired” computational models

“Standardized” access to different simulators, hardware, tools, libraries (e.g. pre-packaged containers, Jupyter notebooks, GitHub, etc.)

Multiple participating sites, capitalizing on CERN world-level expertise in operating distributed infrastructures

Summary

- **CERN Quantum Technology Initiative** established
  - **Assess the potential of QC** in the time scale of the High Luminosity LHC runs and beyond
  - **Build skills** for future programmes
  - Provide a thematic **HEP focus** for international collaborations
- **CERN openlab** has started investigating opportunities in QC and across QC and other relevant fields (AI, HPC)
  - Foster **collaborations between scientists and industry**
  - Major focus on **education and skills development**
- Initial studies on QML **yield very exciting results**
  - Hints of quantum advantage ?
  - More work in understanding trainability, capacity and effective dimensions is needed

https://quantum.cern/

Thanks!

Sofia.Vallecorsa@cern.ch

Quantum Sensing and Low-Energy Physics

Scope

Low-Energy Physics: antimatter, dark matter searches, symmetries, EDM's (AD, AeGIS, ISOLDE, etc.)

Strategy

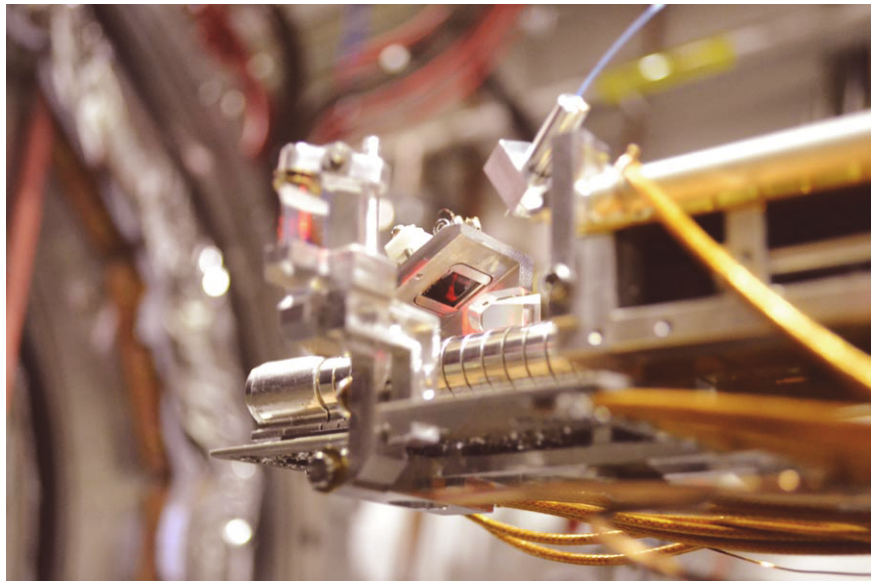
Discrete processes, changes of quantum states

Applications

Novel devices: nanowires, photon upconverters, microwaves, magnetic junctions, SQUIDs, TES

Measurements of properties of trapped, atoms, ions, molecules, Rydberg atoms, neutral systems

Correlations of entangled systems: e.g. e^+ e^- 3γ decay: simultaneous measurement of E, polarization and direction



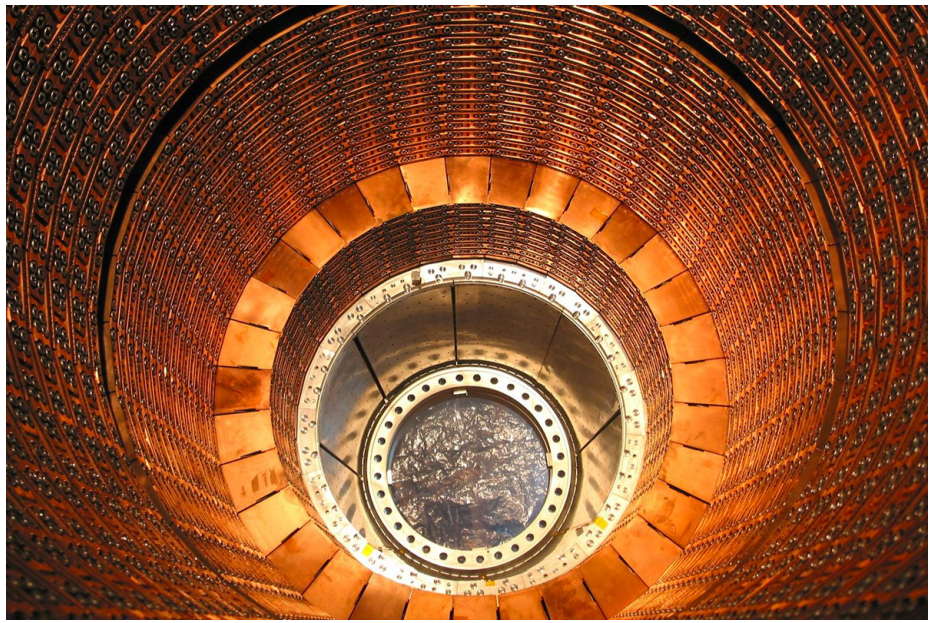
Quantum Sensing for High-Energy Physics

Scope

High-Energy Physics, particle tracking, calorimetry, identification in HEP detectors

Strategy

Quantum “priming” of detectors before measurement, signal enhancement
by laser excitation, quantum effects due to size, cryogenics



Applications

Chromatic particle trackers composed of arrays of nanodots of varying size,
Calorimeters and low-energy single-particle (photons, mip’s, ions,...) detectors made of arrays of nanowires (SNSPD)
“Rydberg-amplified” calorimeters with high dE/dx

Numerical Methods and Simulations in Particle Theory

D. Grabowska, CERN

Modern day HEP requires high performance computing, relying on Monte Carlo simulations

- Mass Spectrum and Scattering in Low Energy Nuclear Physics
- Hadronic contributions to BSM Experimental Searches
- Event generation for Particle Collisions

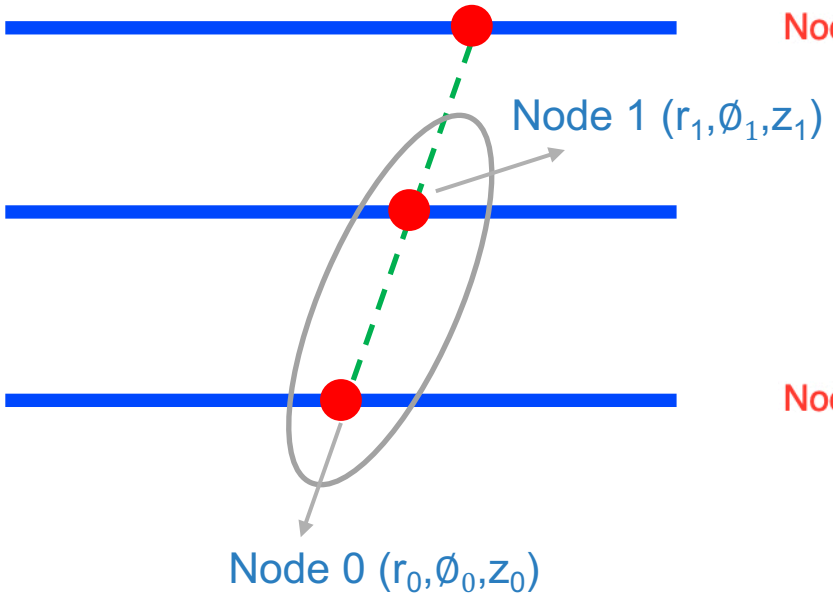
Main focus is developing methodologies and algorithms that would allow us to address these questions using quantum computers, without relying on importance sampling

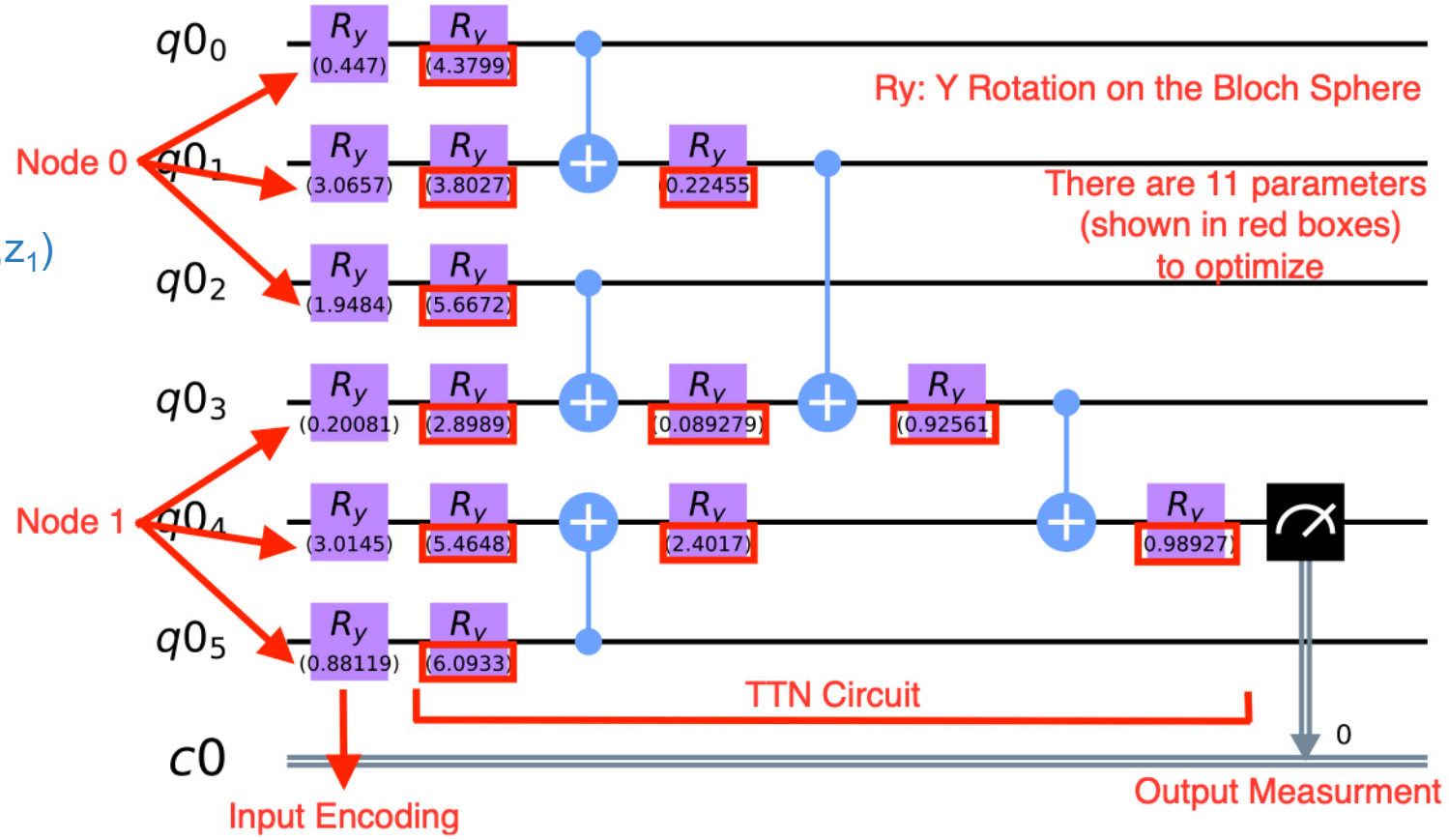
But not every physics problem is amenable to Monte Carlo simulation

- Nuclear Physics at Finite Density (sign problem)
- Interference effects in Parton Showers (must work at amplitude level)
- Transfer Phenomena (must work in real time)
- Baryonic Physics (signal to noise problem)

Quantum Networks (I)

Edge and Node NN as Tree Tensor Networks

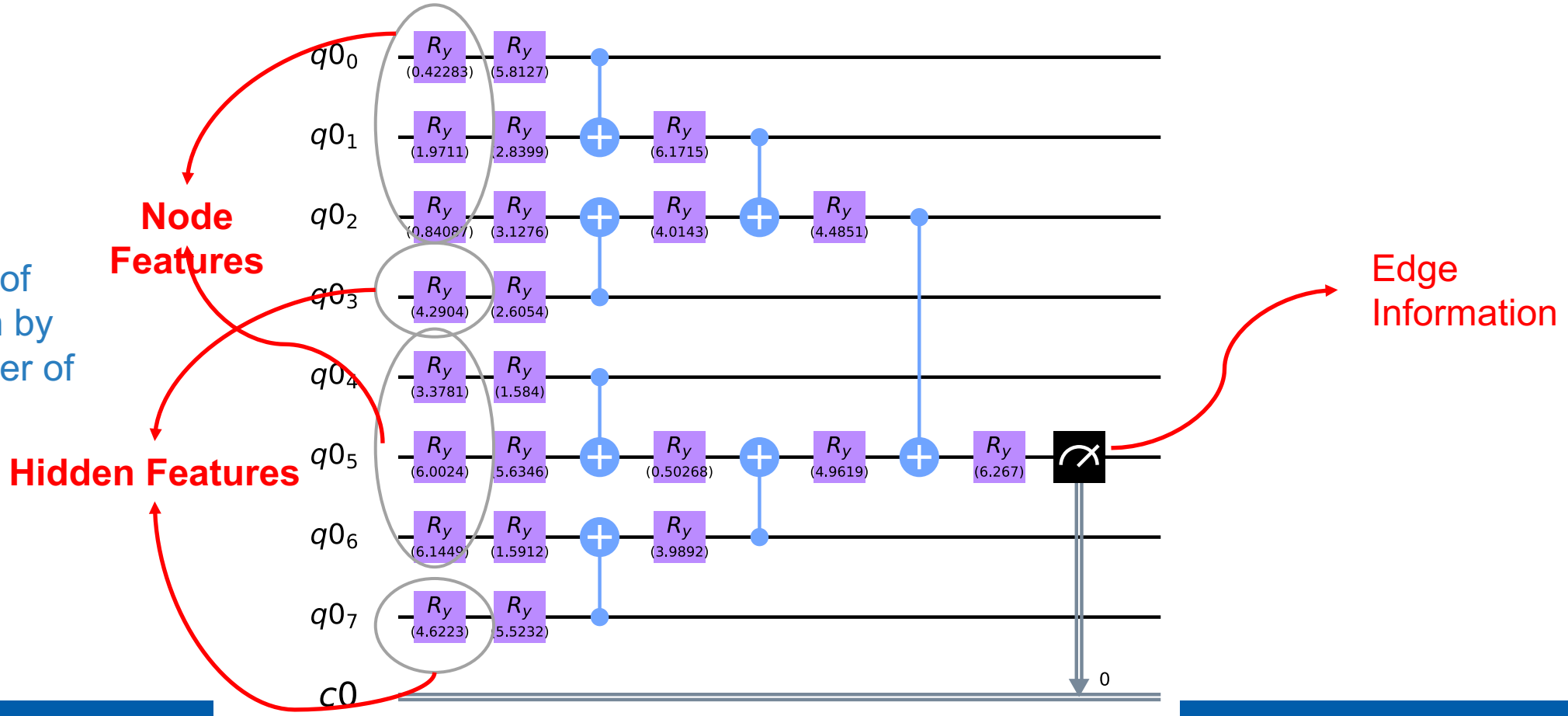




Quantum Networks (II)

Quantum Edge Network

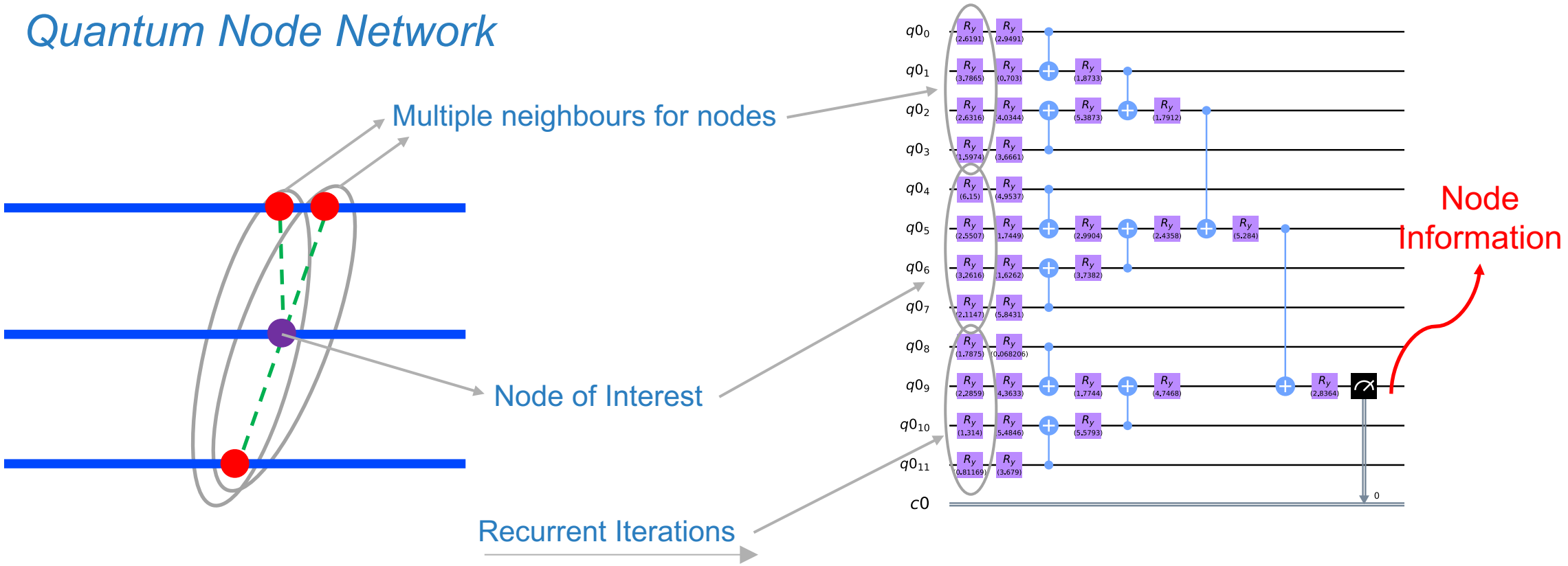
Increase the size of **hidden dimension** by increasing the number of qubits



Quantum Networks

A circuit is setup for each possible neighbor.
2 independent circuits are required for this example.

Quantum Node Network



CV QC simulation

StrawberryFields

- **Gaussian backend** : Maintain an internal covariance matrix
  - Can manipulate exact values
  - **Cannot simulate non-Gaussian gates**
    - Replace by squeezing gates

- **Fock backend** : Use Fock basis representation for circuit
  - Can simulate non-Gaussian gates
  - Approximate the state $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle \approx \sum_{n=0}^{cutoff\ \dim -1} c_n |n\rangle$
    - Cutoff dimension limited by computing resources
    - Accuracy of results decreases as the measured value digresses from zero
  - **Slower than gaussian backend**

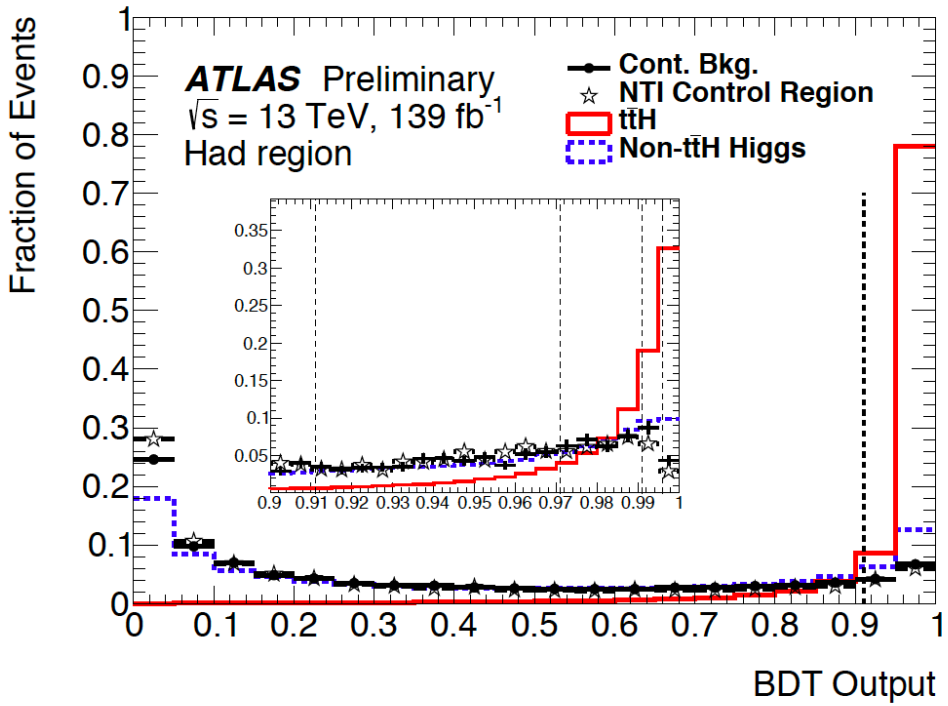
https://arxiv.org/abs/1811.04968

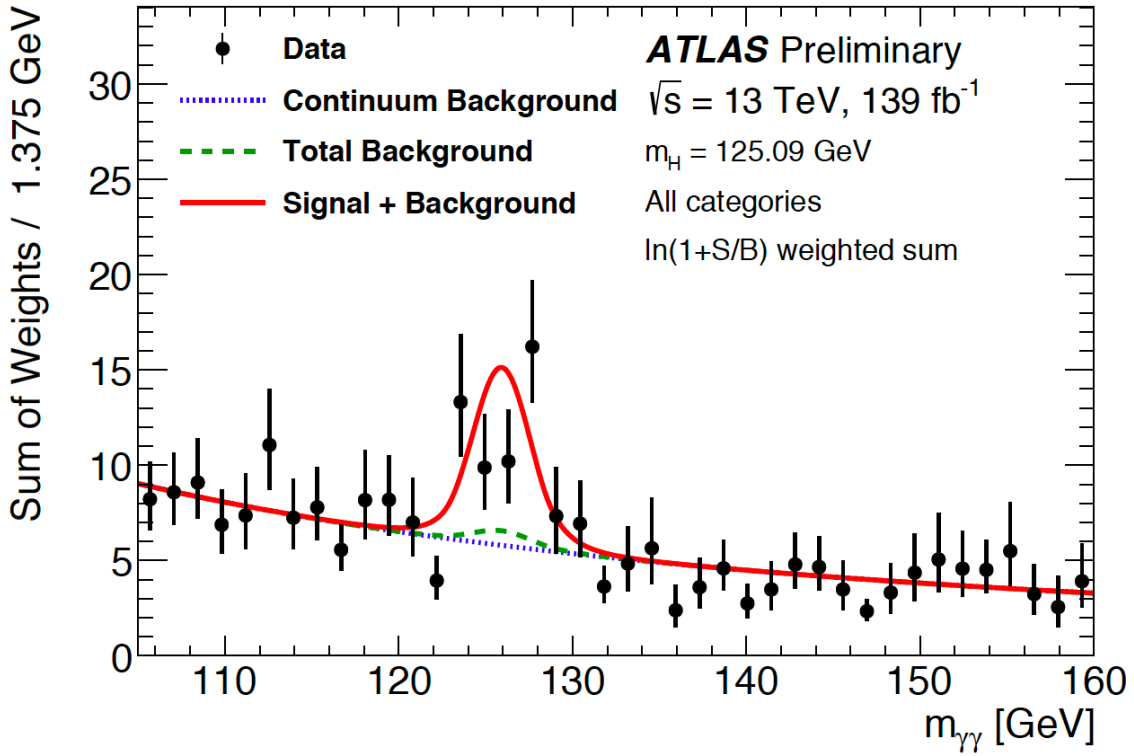
https://doi.org/10.1103/PhysRevLett.82.1784

Higgs boson identification

Classical analysis

Train a XGBoost BDT on ~40 high level features describing the kinematics and topology of the event





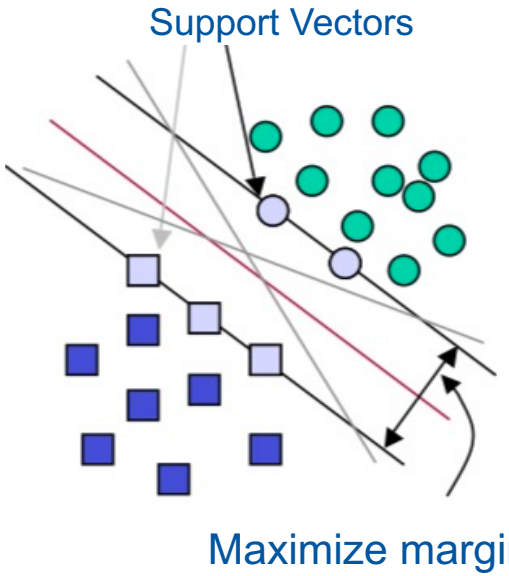
Phys. Lett. B 784 (2018) 173
ATLAS-CONF-2019-004

Quantum Support Vector Machine

Prof. Sau Lan Wu

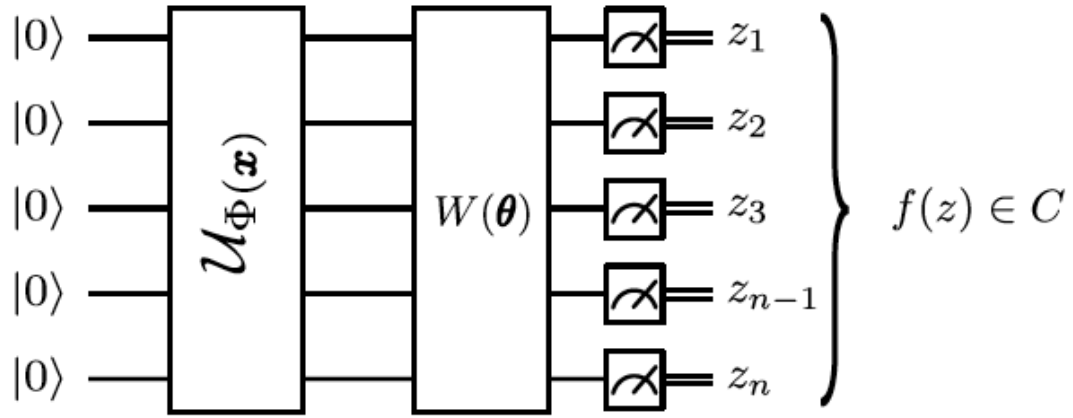
A quantum variational classifier

SVM are **kernel methods**:
Feature maps in high dimensionality space improve separation power



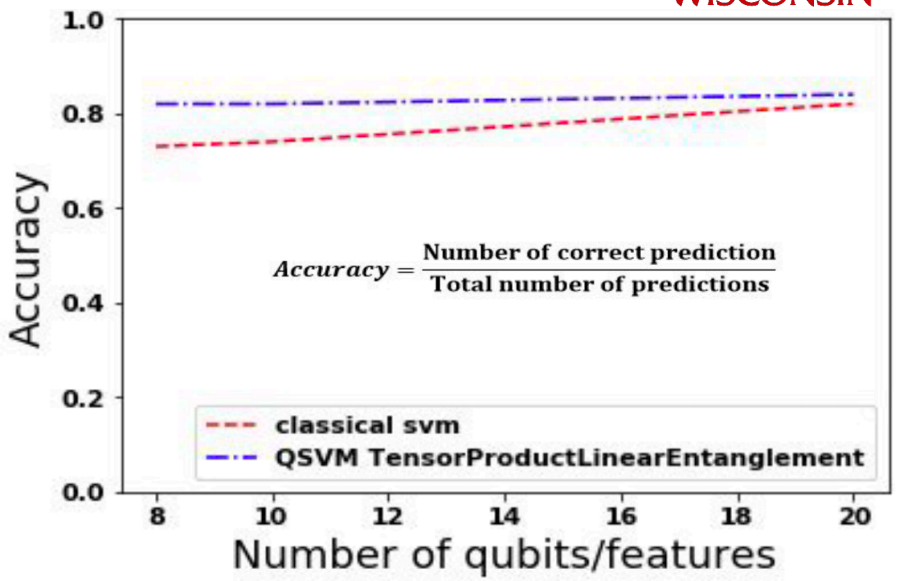
Define a quantum **SVM in Qiskit**

- Test different input **encoding (qubit and amplitude)**
- Reduce number of input features using PCA : 8,10,20 features (N qubits)
- Test different entanglement configurations
- **Shallow variational classifier + Binary measurement**



Quantum SVM

Quantum SVM for ttH(H → γγ) classification

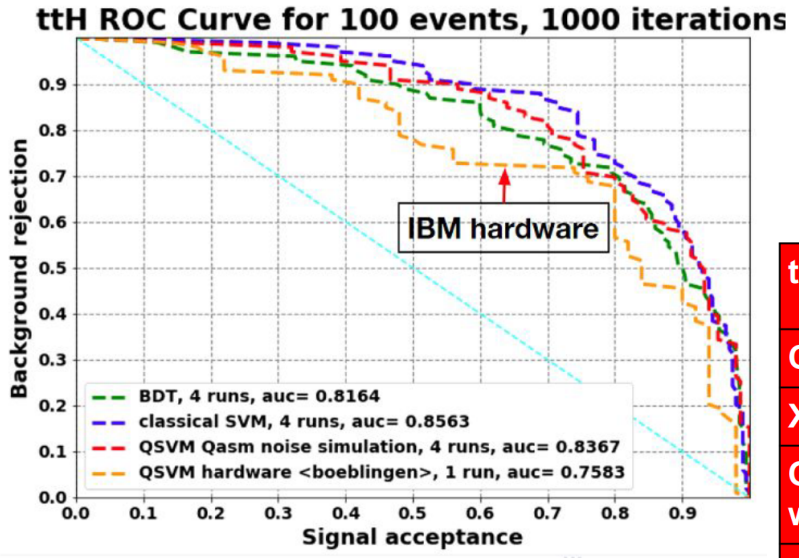


Comparison to BDT and classical SVM

- 100 training events, 100 test events, and 5 qubits
- 1000 iteration on IBM boeblingen

Running full training with quantum simulators requires **large computing resources**

- **Memory increases** with qubit, training events and complexity



ttH(H->γγ) AUC	AUC
Classical SVM	0.856
XGBoost BDT	0.816
QSVM Simulation with Noise	0.837
QSVM Hardware	0.758