QC @ CERN: trainability issues and summary of QC4HEP applications

Workshop on
Quantum Computing 4 HEP
IPPP Durham – 19-20 September 2023

Michele Grossi, PhD

CERN QTI
Quantum Computing Sen. Fellow

QUANTUM
TECHNOLOGY
INITIATIVE

CERN QTI & Motivation

Trainability issues with GM

Summary of QC4HEP applications

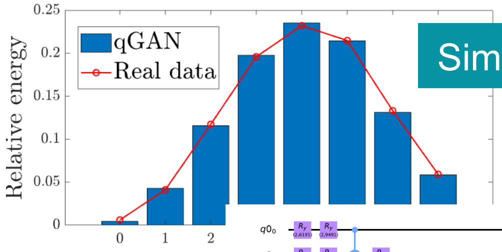
CERN QTI & Motivation

Trainability issues with GM

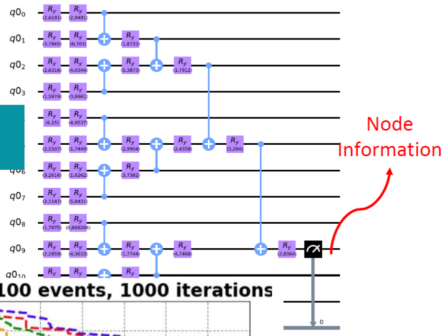
Summary of QC4HEP applications

CERN QTI 1 - Areas of Investigation

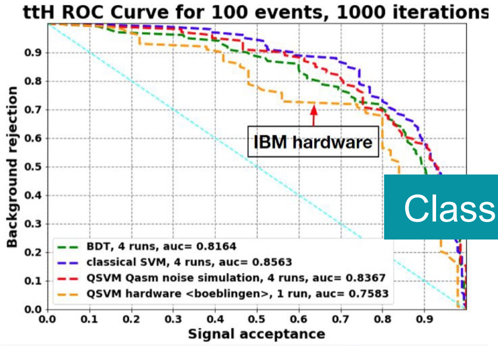
Computing



Simulation

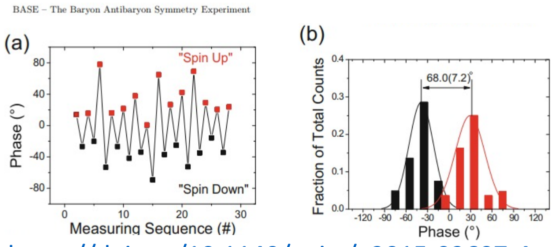


Reconstruction



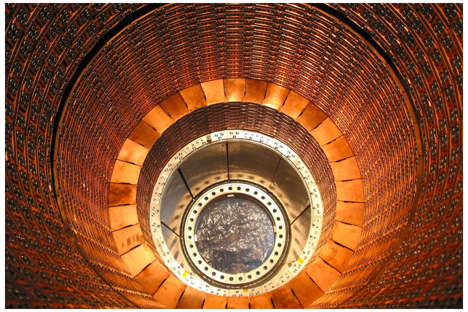
Classification

Sensing



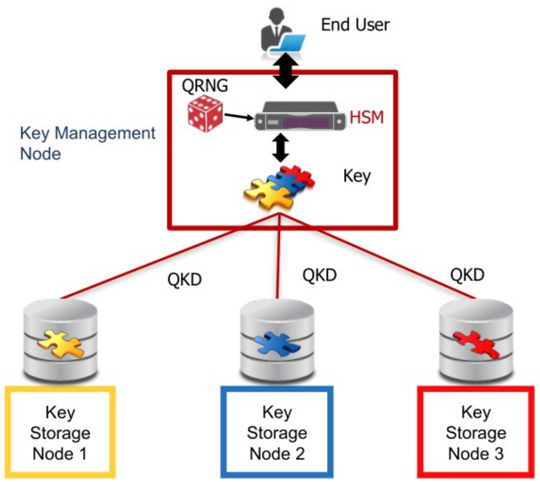
https://doi.org/10.1140/epjst/e2015-02607-4

Low-energy experiments, quantum states measurements, nano-technologies



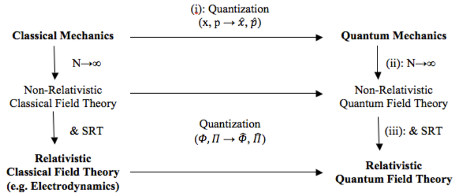
Future HEP Detectors

Networks/Comms

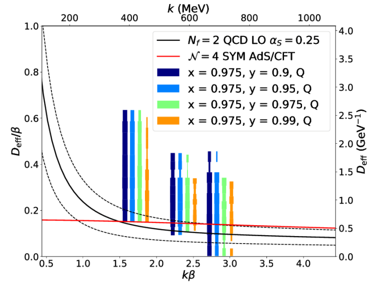


Quantum networks, QKD applications

Theory



Quantum Field Theory



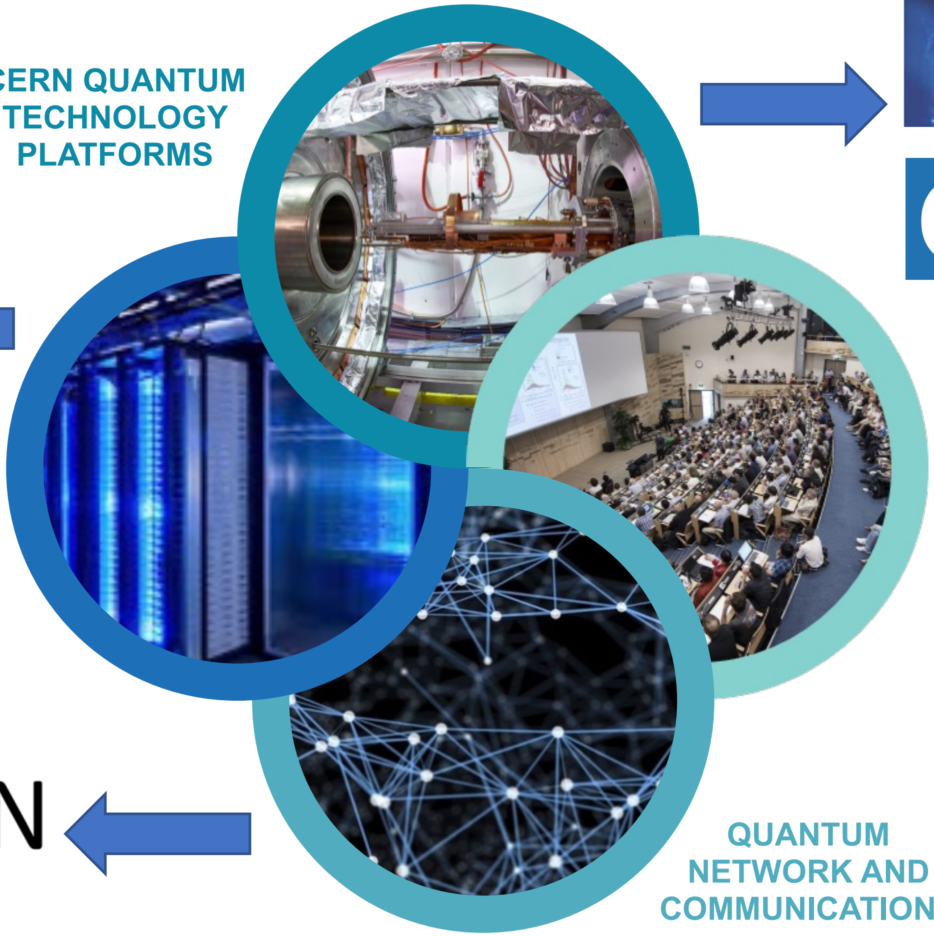
https://cds.cern.ch/record/2703396

Lattice QCD

CERN QTI Phase 2 – Expected Impact (high-level)

CERN QUANTUM TECHNOLOGY PLATFORMS

HYBRID QUANTUM COMPUTING AND ALGORITHMS



COLLABORATION FOR IMPACT

QUANTUM NETWORK AND COMMUNICATIONS

Physics Beyond Colliders

FCC hh ee he

HiLumi HL-LHC PROJECT

WLCG
Worldwide LHC Computing Grid

LHCOPN

OPEN QUANTUM INSTITUTE

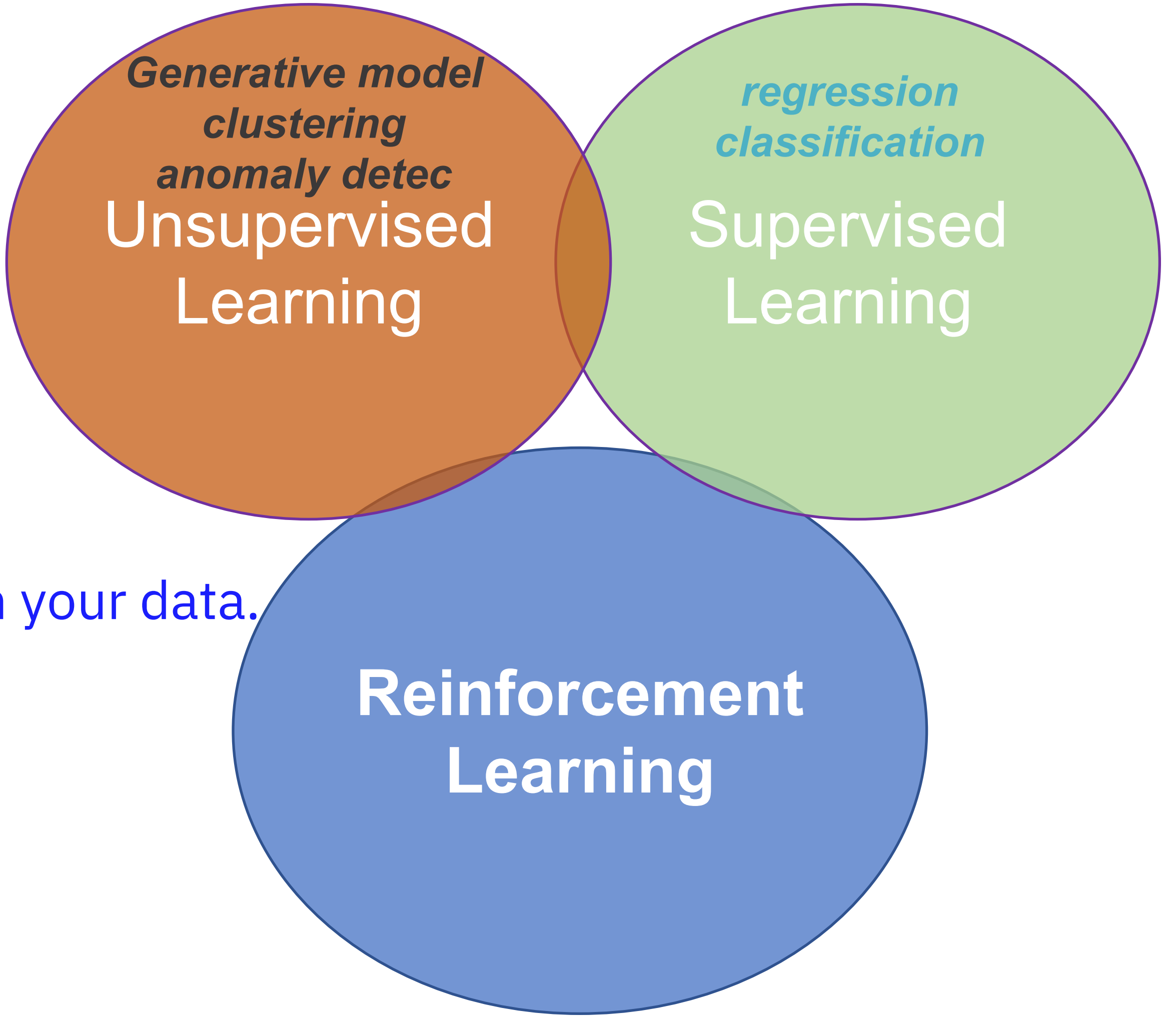
CERN | QUANTUM TECHNOLOGY INITIATIVE

Studying ~~Deep~~ Learning in physics
Quantum Machine

- **High quality labelled training data** from realistic MC simulation
- Large **experimental datasets**
- Interestingly **structured data** at multiple scales
- Detailed understanding of **systematic uncertainties**

M. Erdmann, J. Glombitza,G. Kasieczka, U. Klemradt, Deep Learning for physics research

Machine Learning + QC



Unsupervised ML

Unlabeled data.

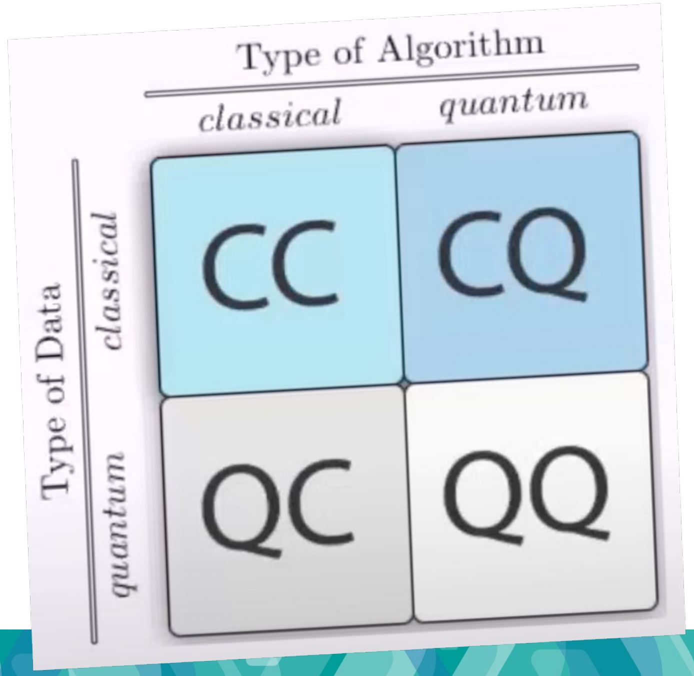
ML finds patterns in your data.

Indirect evaluation.

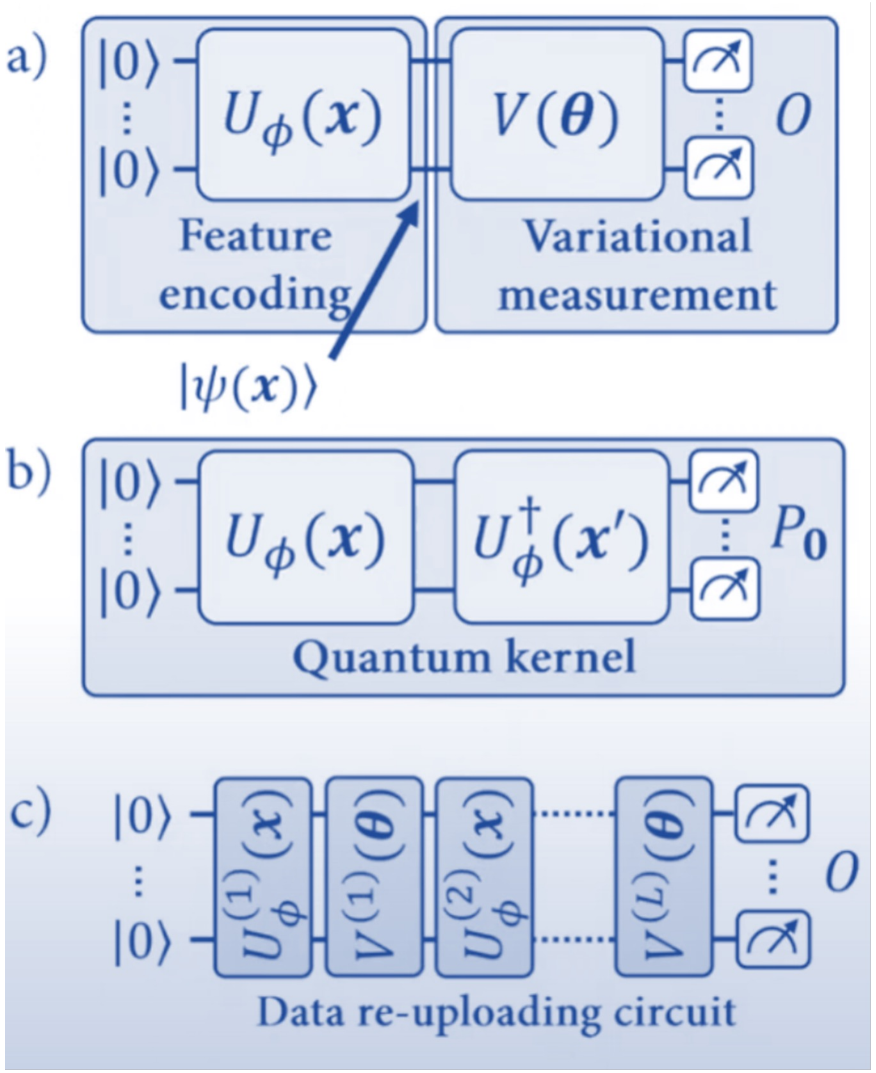
Supervised ML

Labeled data, i.e., data with defined output.

A model is trained giving this data and you have direct evaluation.



QML models



a) Explicit quantum model:

$$f_{\theta}(x) = \mathrm{Tr}[\rho(x) O_{\theta}]$$

$$\rho(x) = |\psi(x)\rangle\langle\psi(x)|$$

$$O_{\theta} = V^{\dagger}(\theta) O V(\theta)$$

A linear model with a restricted w

b) Implicit quantum model:

$$f_{\alpha}(x) = \mathrm{Tr}[\rho(x) O_{\alpha,\mathcal{D}}] \qquad O_{\alpha,\mathcal{D}} = \sum_{m=1}^{M} \alpha_m \rho(x^{(m)})$$

A kernel linear model

c) Data re-uploading model:

$$f_{\theta}(x) = \mathrm{Tr}[\rho(x, \theta) O_{\theta}]$$

S.Jerbi at al., Quantum Machine Learning Beyond Kernel Methods – Nature Communications 14, 517 (2023)

CERN QTI & Motivation

Trainability issues with GM

Summary of QC4HEP applications

Generative Model

unsupervised learning problem

Explicit

- definition of explicit density form that allows likelihood inference
- VAE

Implicit

- flexible transformation from random noise to generated samples
  (a stochastic process to draw samples from the underlying data distribution)
- no distribution specified/required
- no tractable likelihood function required
- GAN

Trainability barriers and opportunities in quantum generative modeling

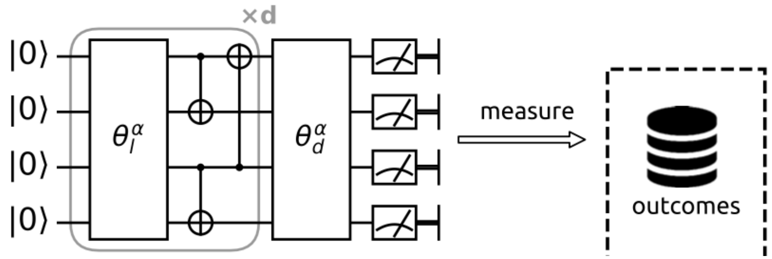
Manuel S. Rudolph,[1,*] Sacha Lerch,[1,*] Supanut Thanasilp,[1,2,*]
Oriel Kiss,[3,4] Sofia Vallecorsa,[3] Michele Grossi,[3] and Zoë Holmes[1]

[1] *Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[2] *Chula Intelligent and Complex Systems, Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, 10330*
[3] *European Organization for Nuclear Research (CERN), Geneva 1211, Switzerland*
[4] *Department of Nuclear and Particle Physics, University of Geneva, Geneva 1211, Switzerland*

(Dated: May 5, 2023)

Quantum Generative Models

1 **Discrete data**
→ Use quantum random source



Characteristics :

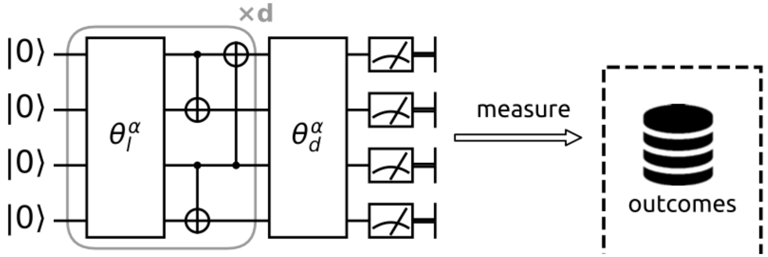
- Discrete
- Uses quantum randomness.
- 1 shot = 1 sample.
- Needs more qubits.

Examples

- Quantum circuit Born machines (QCBM) (*Phys. Rev. A* **98**, 062324, 2018)
- Discrete Quantum GAN for learning random distribution (*npj Quantum Inf* **5**, 103, 2019)
- Quantum GAN for Bar and Stripes generation (*Phys. Rev. A* **99**, 5, 2019)

Quantum Generative Models

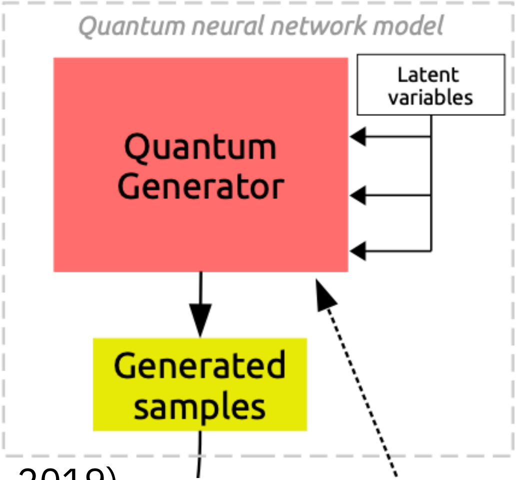
① **Discrete data**
→ Use quantum random source



Characteristics :
- Discrete
- Uses quantum randomness.
- 1 shot = 1 sample.
- Needs more qubits.

Examples
- Quantum circuit Born machines (QCBM) (*Phys. Rev. A* **98**, 062324, 2018)
- Discrete Quantum GAN for learning random distribution (*npj Quantum Inf* **5**, 103, 2019)
- Quantum GAN for Bar and Stripes generation (*Phys. Rev. A* **99**, 5, 2019)

② **Continuous data**
→ Use classical random source.

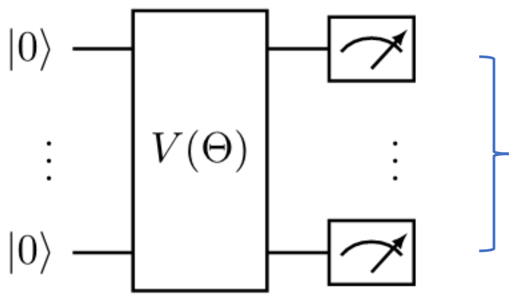


Characteristics :
- Continuous.
- Requires low number of qubits.
- High number of shots.
- 1 sample = many shots.

Examples
- Variational Quantum Generator (*arXiv:1901.00848*, 2019)
- Style-based quantum GAN for MC event generation (*Quantum* **6**, 777, 2022)

Quantum Circuit Born machine (QCBM) in a nutshell

1. ***Sample** from a variational pure state $|\psi(\theta)\rangle$ by projective measurement with probability given by the **Born rule**:* $\boldsymbol{p_\theta(x)} = |\langle \boldsymbol{x} | \boldsymbol{\psi(\theta)} \rangle|^2$.



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

2. ***Training*** (Hybrid loop):

- KL divergence — Delgado and Hamilton, arXiv:2203.03578.
- Adversarial (QGAN) — Zoufal, et al., *npj Quantum Inf* **5**, 103 (2019).
- In the phase space — Kyriienko, et al., arXiv: 2202.08253.
- Maximum Mean Discrepancy — Rudolph et al, arXIV: 2305.02881.

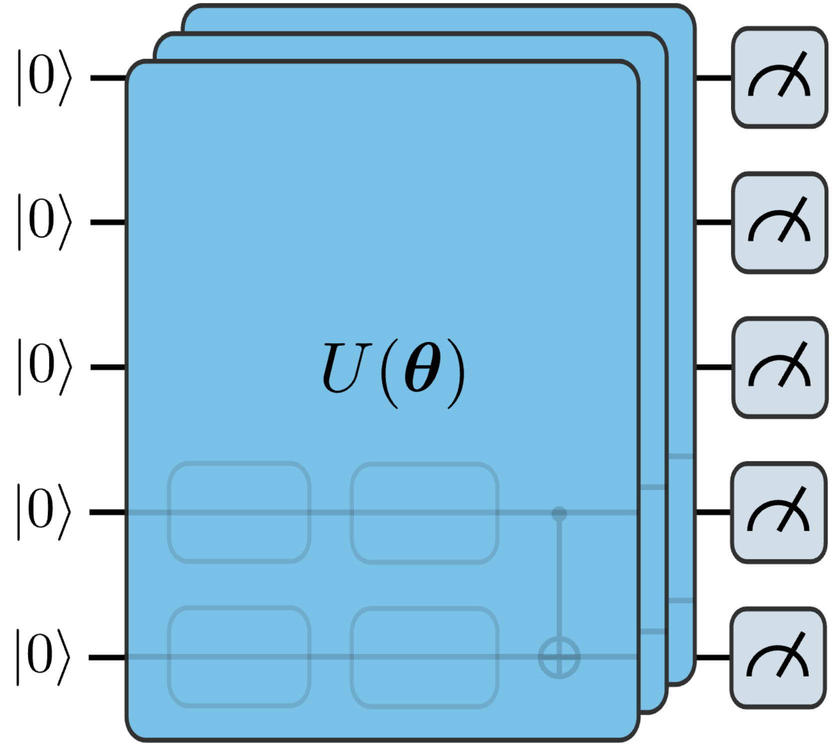
$$\text{MMD(P,Q)} = \mathbb{E}_{\substack{X\sim P\\ Y\sim P}}[K(X,Y)] + \mathbb{E}_{\substack{X\sim Q\\ Y\sim Q}}[K(X,Y)] - 2\mathbb{E}_{\substack{X\sim P\\ Y\sim Q}}[K(X,Y)]$$

3. **Why** the MMD ?

- Resource efficient for NISQ devices.
- Stable.
- However, empirically less performant.

Quantum Circuit Born Machine (QCBM)

Benedetti et al., npj Quantum Inf 5, 45 (2019)

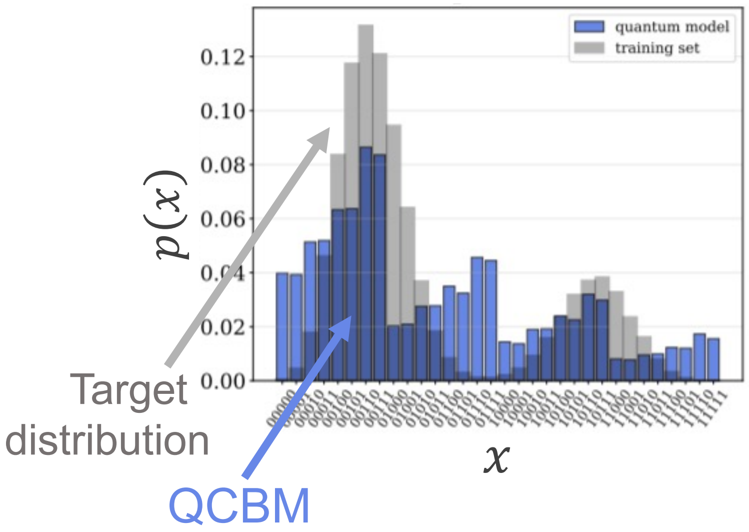


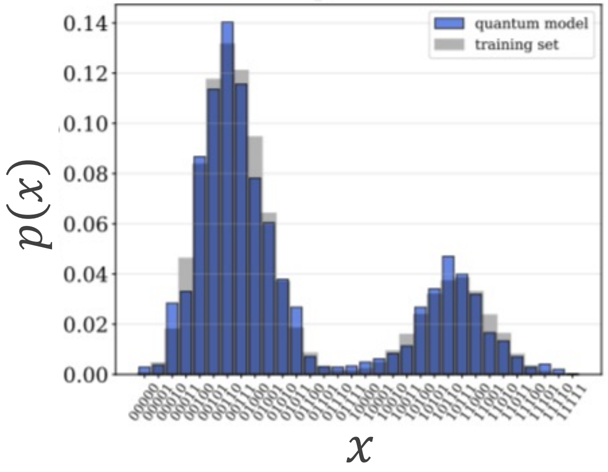
Probability for each sample:

$$p(x) = \left|\boldsymbol{\alpha}_{x_1 x_2 \dots x_n}\right|^2 = |\langle x|U(\theta)|0\rangle|^2$$

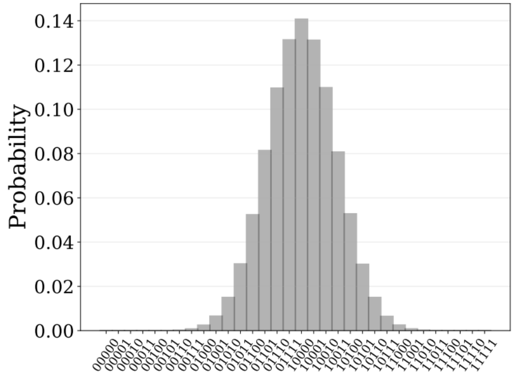
$$|\psi\rangle = \begin{pmatrix} \alpha_{0\dots0} \\ \alpha_{0\dots1} \\ \vdots \\ \alpha_{1\dots1} \end{pmatrix}$$

A deeper circuit gives more flexibility!

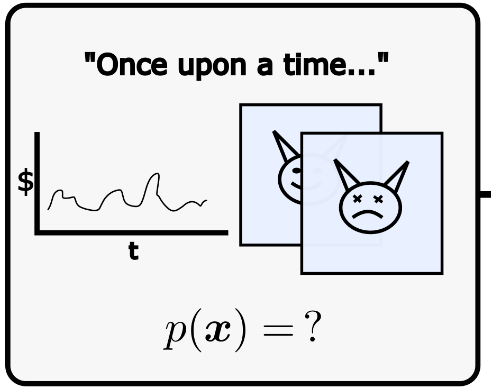


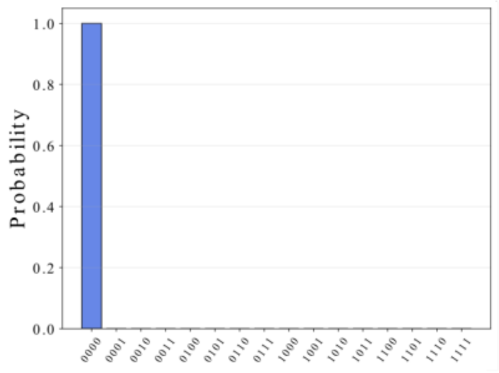


Quantum Circuit Born Machine (QCBM)



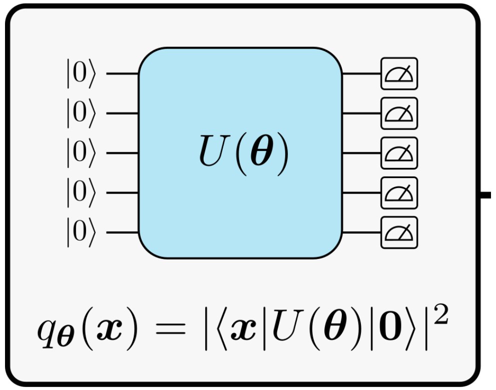
Real-world



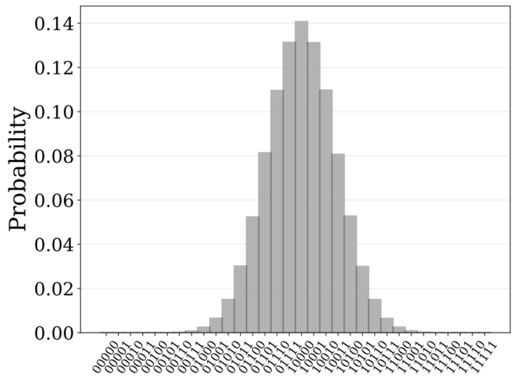


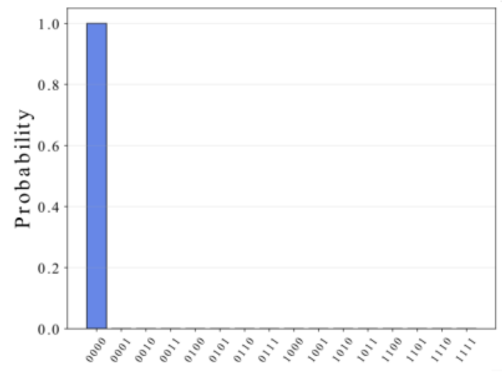
Overlap 0%

Quantum Circuit Born Machine

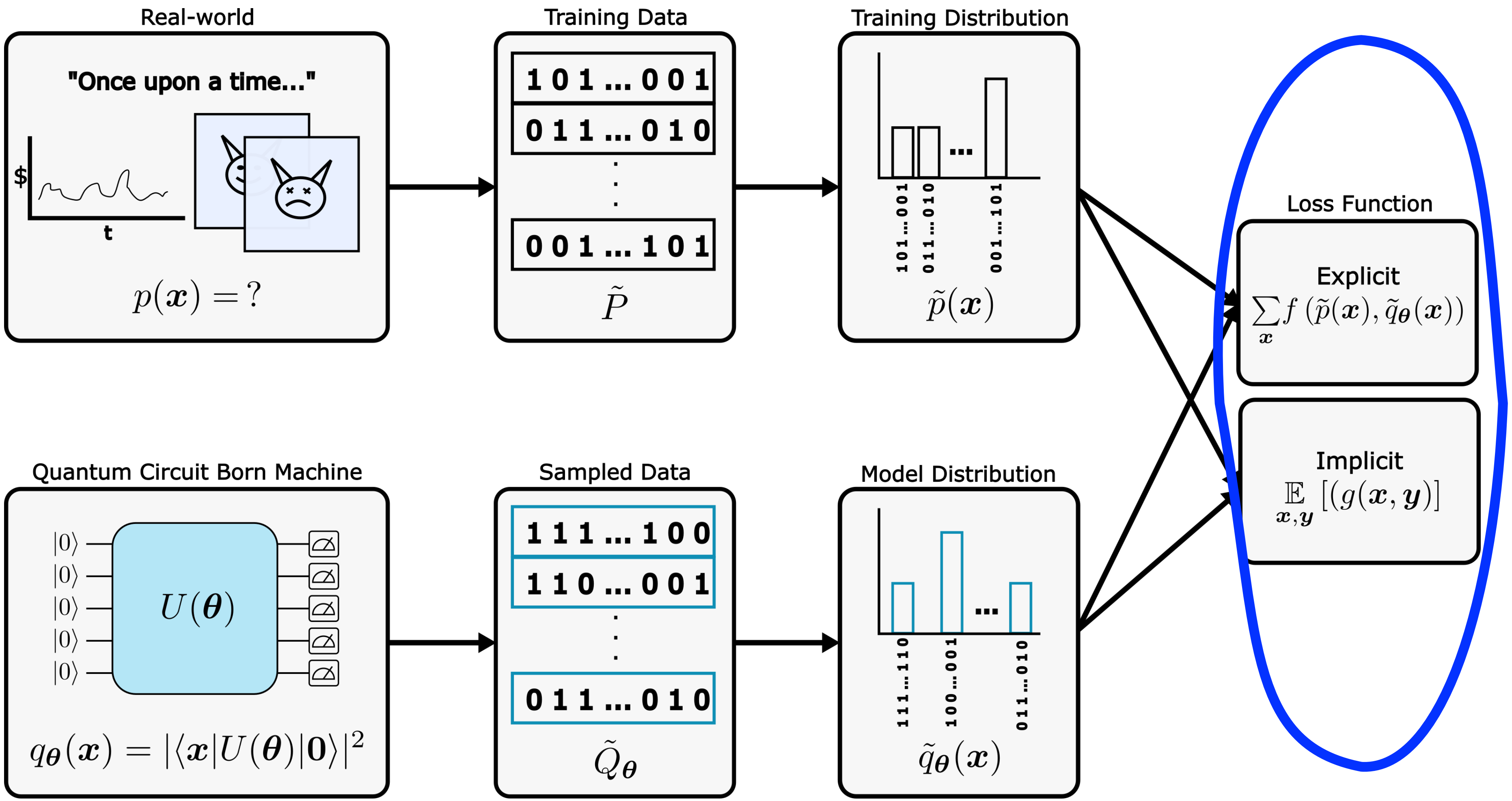


Quantum Circuit Born Machine (QCBM)

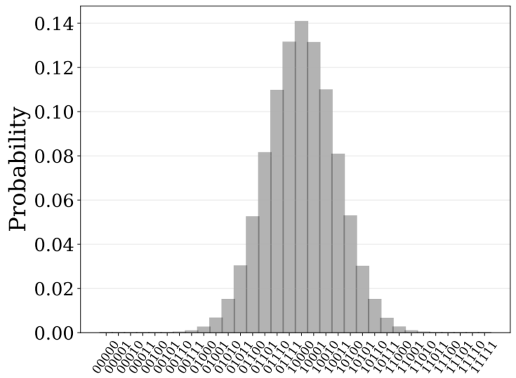


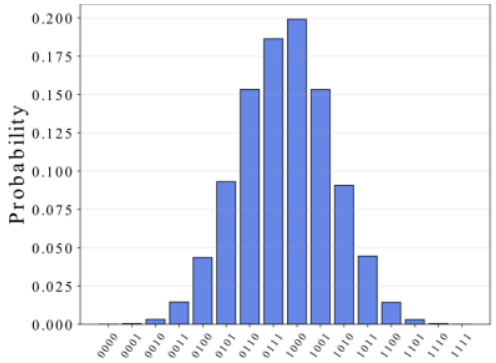


Overlap 0%



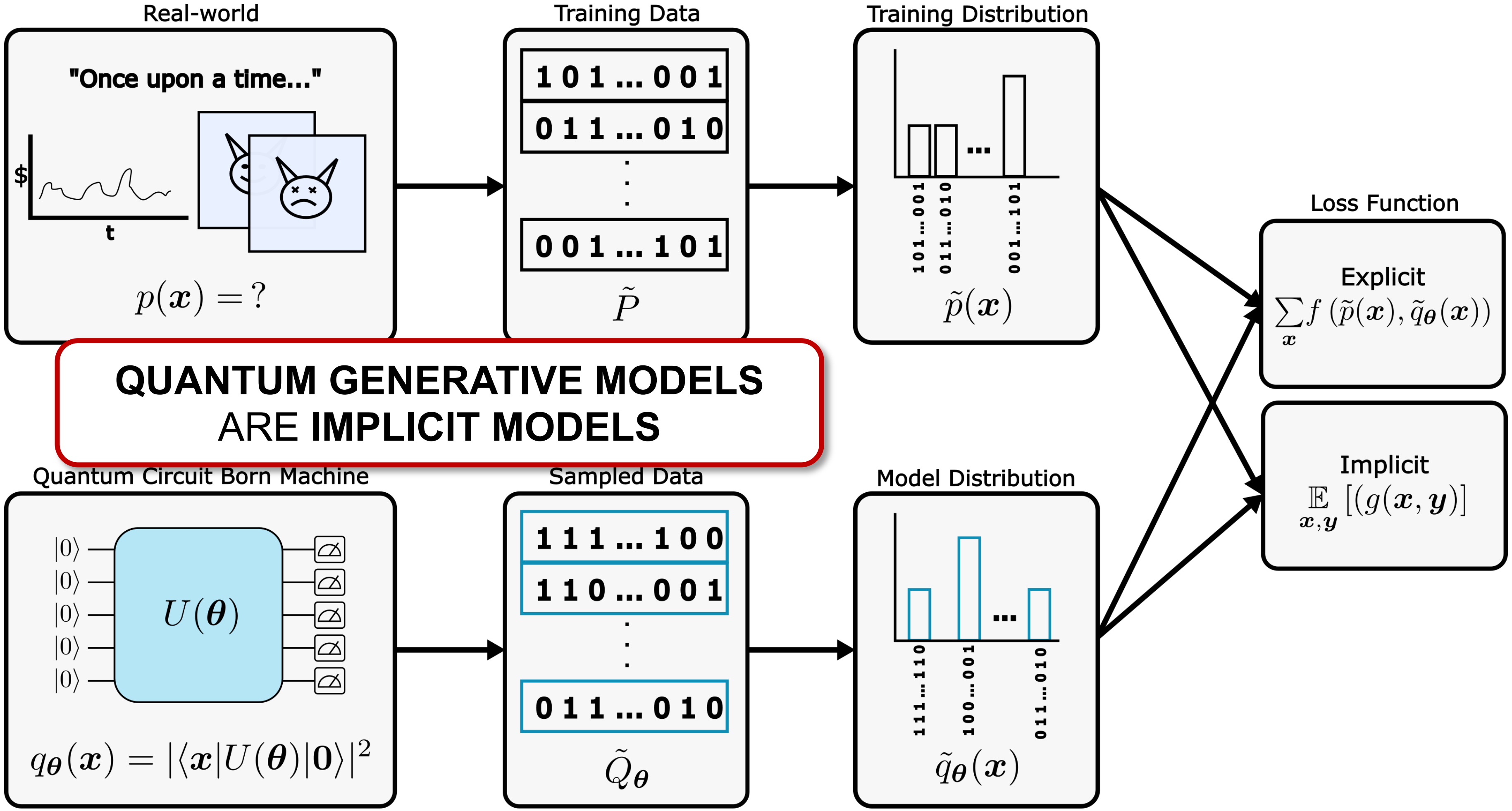
Quantum Circuit Born Machine (QCBM)



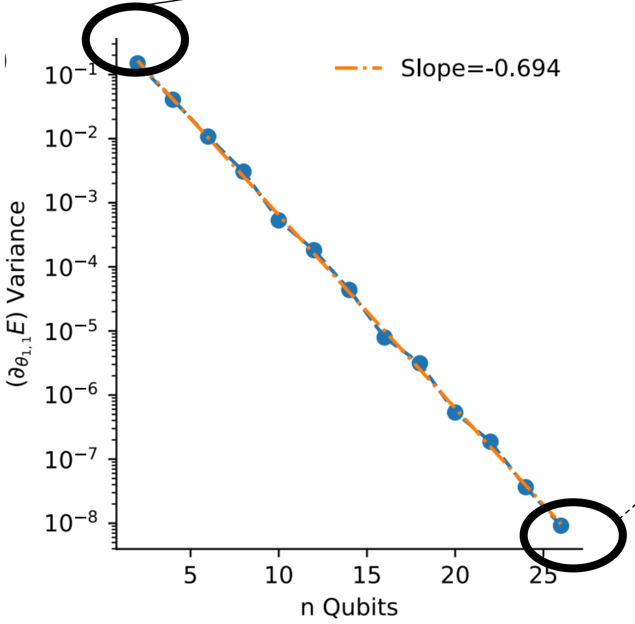


Overlap 99.7%

QUANTUM GENERATIVE MODELS
ARE **IMPLICIT MODELS**



Barren plateaus



Small gradients

High precision required to find cost-minimizing direction

Resource intensive

($\frac{1}{\epsilon^2}$ measurements are required estimate a cost to precision ϵ)

Barren plateaus

Choice in circuit

- Too expressive
- Too entangling

Choice in target learning problem

Choice in cost function (loss)

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications **9**, Article number: 4812 (2018) | Cite this article

Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus

Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles
PRX Quantum **3**, 010313 – Published 24 January 2022

Entanglement-Induced Barren Plateaus

Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe
PRX Quantum **2**, 040316 – Published 25 October 2021

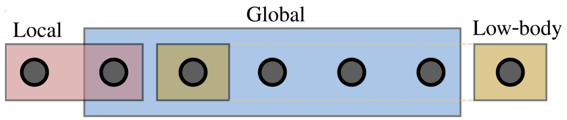
Barren Plateaus Preclude Learning Scramblers

Zoë Holmes, Andrew Arrasmith, Bin Yan, Patrick J. Coles, Andreas Albrecht, and Andrew T. Sornborger
Phys. Rev. Lett. **126**, 190501 – Published 12 May 2021

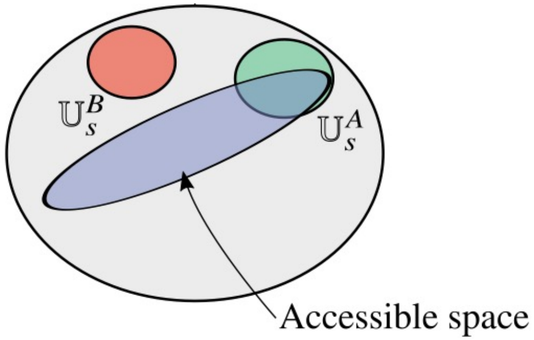
Cost function dependent barren plateaus in shallow parametrized quantum circuits

M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio & Patrick J. Coles

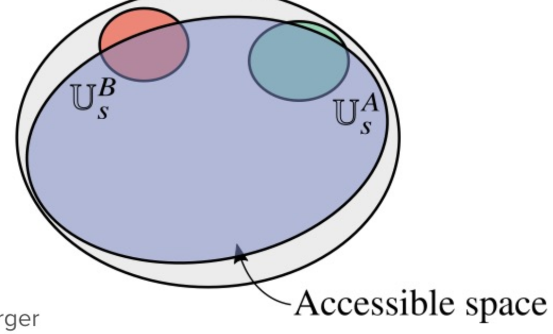
Nature Communications **12**, Article number: 1791 (2021) | Cite this article



Inexpressive



Expressive



Global

$$H = \sigma_1^z \otimes \sigma_2^z \otimes \cdots \otimes \sigma_n^z$$

$$\langle \psi(\theta)|H|\psi(\theta)\rangle$$

Local

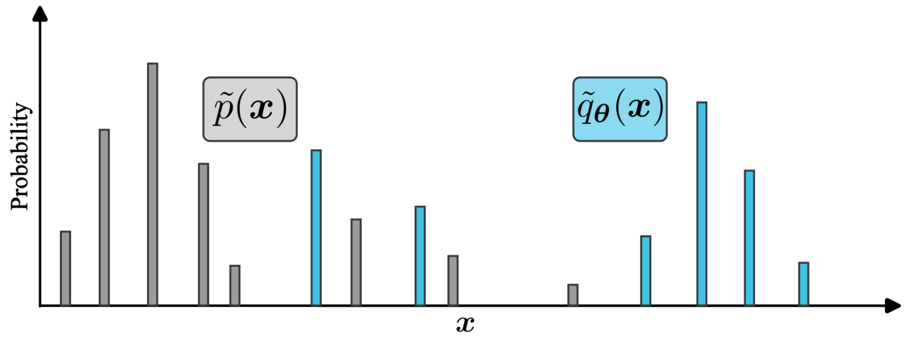
$$H = \sigma_1^z \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_n$$

Generative Loss Function

Explicit

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

Sample quantum state and build the empirical distribution q to be used in the loss



Generative Loss Function

Explicit

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

KL Divergence

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

Reverse KL Divergence

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

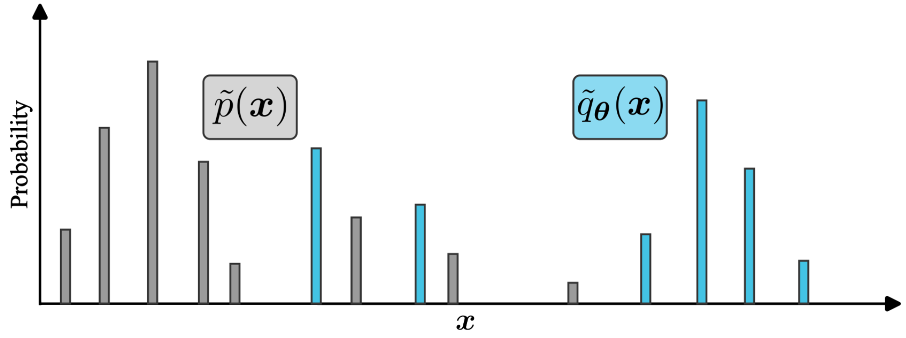
Jensen-Shannon Divergence

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

Total Variation Distance

$$\mathcal{L}^{\mathrm{TVD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} |p(\boldsymbol{x}) - q_{\boldsymbol{\theta}}(\boldsymbol{x})|$$

Problem: It is very rare you sample relevant data point



Generative Loss Function

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

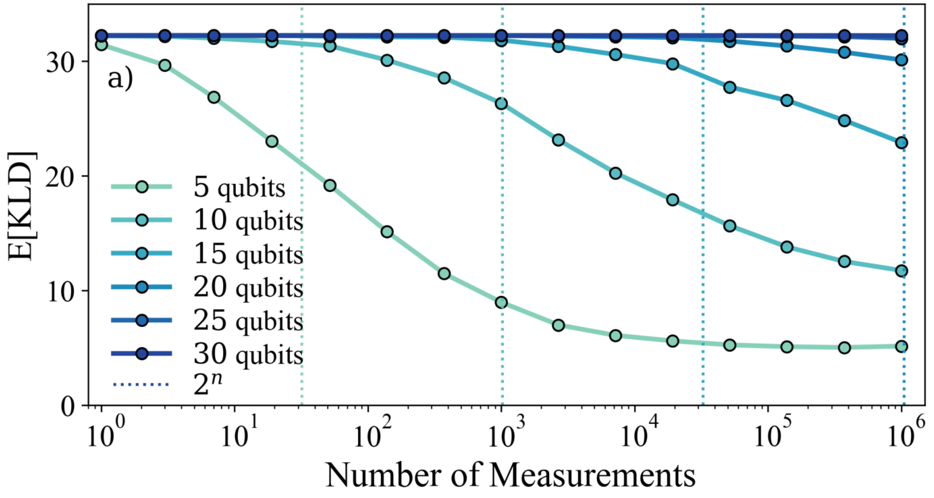
KL Divergence
$$\mathcal{L}^{\mathrm{KLD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} p(\boldsymbol{x}) \log\left(\frac{p(\boldsymbol{x})}{q_{\boldsymbol{\theta}}(\boldsymbol{x})}\right)$$

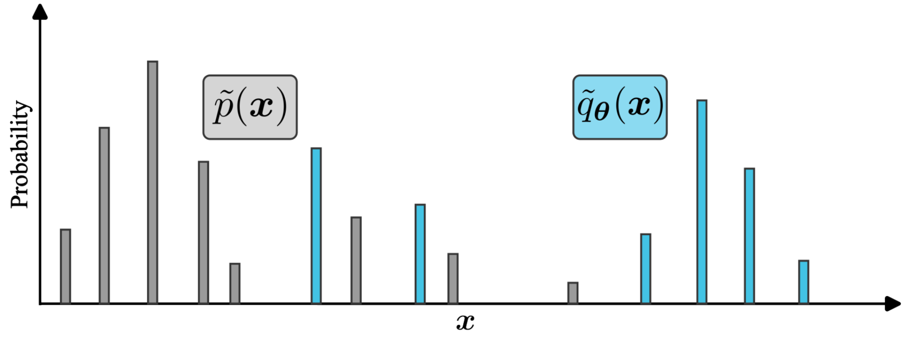
Reverse KL Divergence
$$\mathcal{L}^{\mathrm{rev-KLD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} q_{\boldsymbol{\theta}}(\boldsymbol{x}) \log\left(\frac{q_{\boldsymbol{\theta}}(\boldsymbol{x})}{p(\boldsymbol{x})}\right)$$

Jensen-Shannon Divergence
$$\mathcal{L}^{\mathrm{JSD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} \left[p(\boldsymbol{x}) \log\left(\frac{p(\boldsymbol{x})}{p(\boldsymbol{x}) + q_{\boldsymbol{\theta}}(\boldsymbol{x})}\right) + q_{\boldsymbol{\theta}}(\boldsymbol{x}) \log\left(\frac{q_{\boldsymbol{\theta}}(\boldsymbol{x})}{p(\boldsymbol{x}) + q_{\boldsymbol{\theta}}(\boldsymbol{x})}\right)\right]$$

Total Variation Distance
$$\mathcal{L}^{\mathrm{TVD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} |p(\boldsymbol{x}) - q_{\boldsymbol{\theta}}(\boldsymbol{x})|$$

Loss concentration





Generative Loss Function

Explicit

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

ARE **NOT TRAINABLE** FOR **GENERIC CIRCUITS**

NO GO THM

KL Divergence

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

Reverse KL Divergence

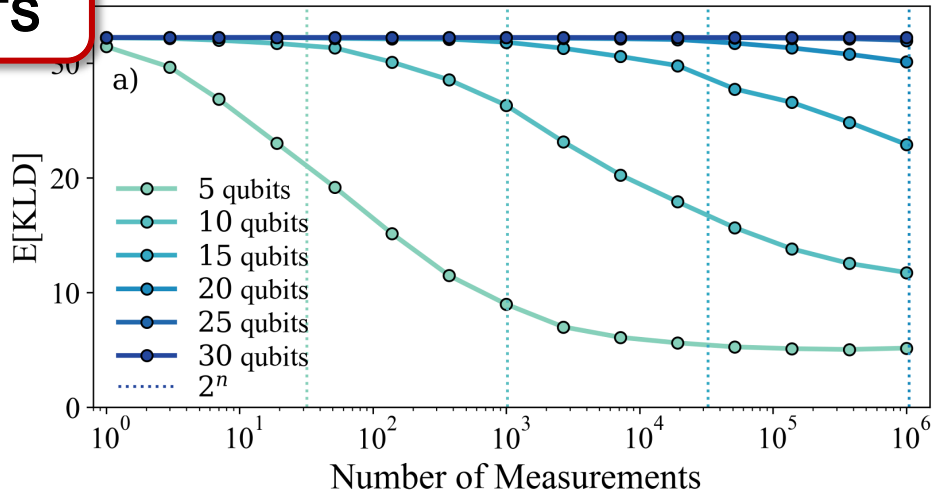
$$\mathcal{L}^{\mathrm{rev-KLD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} q_{\boldsymbol{\theta}}(\boldsymbol{x}) \log\left(\frac{q_{\boldsymbol{\theta}}(\boldsymbol{x})}{p(\boldsymbol{x})}\right)$$

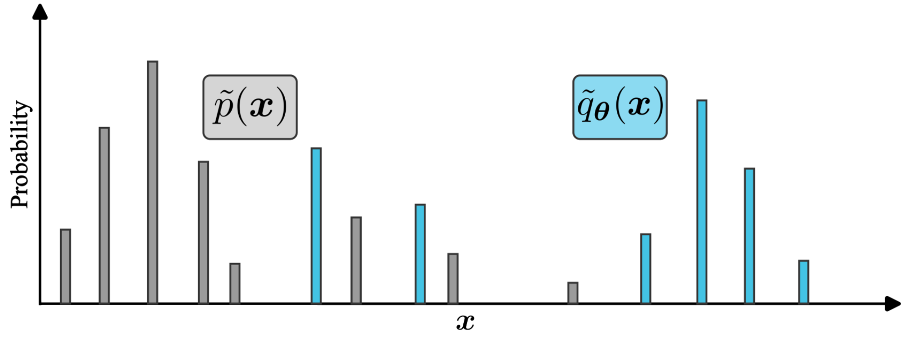
Jensen-Shannon Divergence

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

Total Variation Distance

$$\mathcal{L}^{\mathrm{TVD}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{x}\in\mathcal{X}} |p(\boldsymbol{x}) - q_{\boldsymbol{\theta}}(\boldsymbol{x})|$$





Generative Loss Function

Implicit

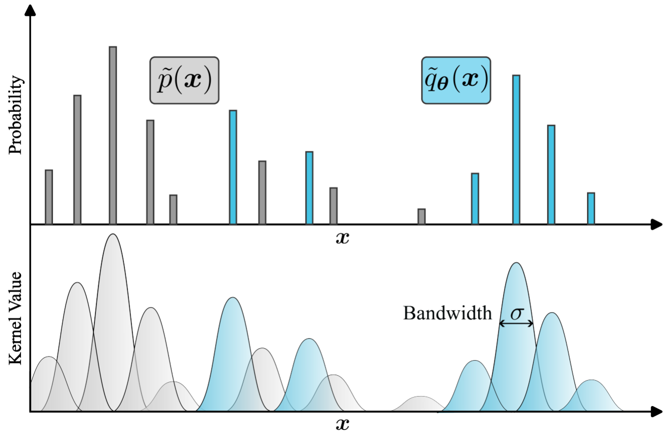
$$\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{y})\right]$$

Maximum Mean Discrepancy

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

with

$$K_\sigma(\boldsymbol{x},\boldsymbol{y}) = e^{-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}{2\sigma}} = \prod_{i=1}^{n} e^{-\frac{(x_i-y_i)^2}{2\sigma}}$$



Generative Loss Function

Implicit

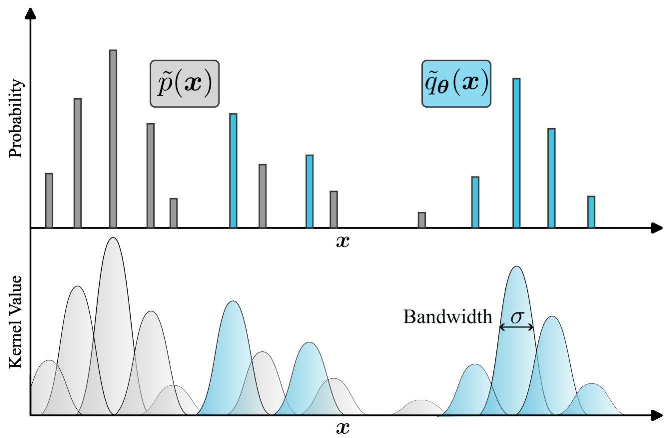
$$\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[g(\boldsymbol{x},\boldsymbol{y})\right]$$

Maximum Mean Discrepancy

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

with

$$K_\sigma(\boldsymbol{x},\boldsymbol{y}) = e^{-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}{2\sigma}} = \prod_{i=1}^{n} e^{-\frac{(x_i-y_i)^2}{2\sigma}}$$



Equivalent representation as observable ML vs QC

$$\mathcal{M}(\rho,\rho') = \mathrm{Tr}\left[O^{(\sigma)}_{\mathrm{MMD}}(\rho\otimes\rho')\right]$$

$$O^{(\sigma)}_{\mathrm{MMD}} := \sum_{\boldsymbol{x},\boldsymbol{y}} K_\sigma(\boldsymbol{x},\boldsymbol{y})|\boldsymbol{x}\rangle\langle\boldsymbol{x}| \otimes |\boldsymbol{y}\rangle\langle\boldsymbol{y}|\,.$$

to Pauli basis

$$O^{(\sigma)}_{\mathrm{MMD}} = \sum_{l=0}^{n}\binom{n}{l}(1-p_\sigma)^{n-l}p_\sigma^l D_{2l}$$

Generative Loss Function

Implicit

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

Equivalent representation as observable ML vs QC

$$\mathcal{M}(\rho,\rho') = \mathrm{Tr}\left[O^{(\sigma)}_{\mathrm{MMD}}(\rho\otimes\rho')\right]$$

$$O^{(\sigma)}_{\mathrm{MMD}} := \sum_{\boldsymbol{x},\boldsymbol{y}} K_\sigma(\boldsymbol{x},\boldsymbol{y})|\boldsymbol{x}\rangle\langle\boldsymbol{x}|\otimes|\boldsymbol{y}\rangle\langle\boldsymbol{y}|\,.$$

to Pauli basis

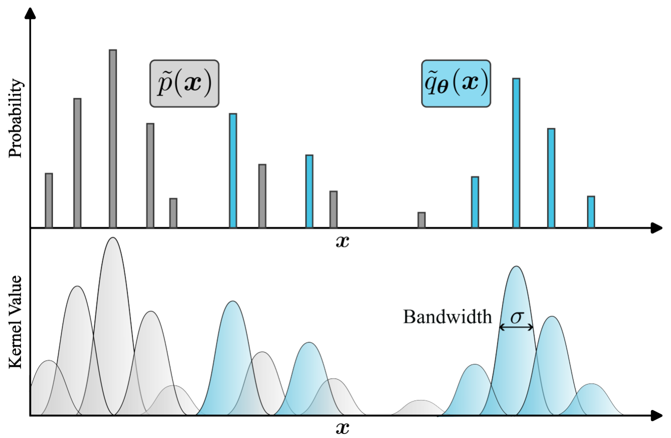
$$O^{(\sigma)}_{\mathrm{MMD}} = \sum_{l=0}^{n}\binom{n}{l}(1-p_\sigma)^{n-l}p_\sigma^l D_{2l}$$

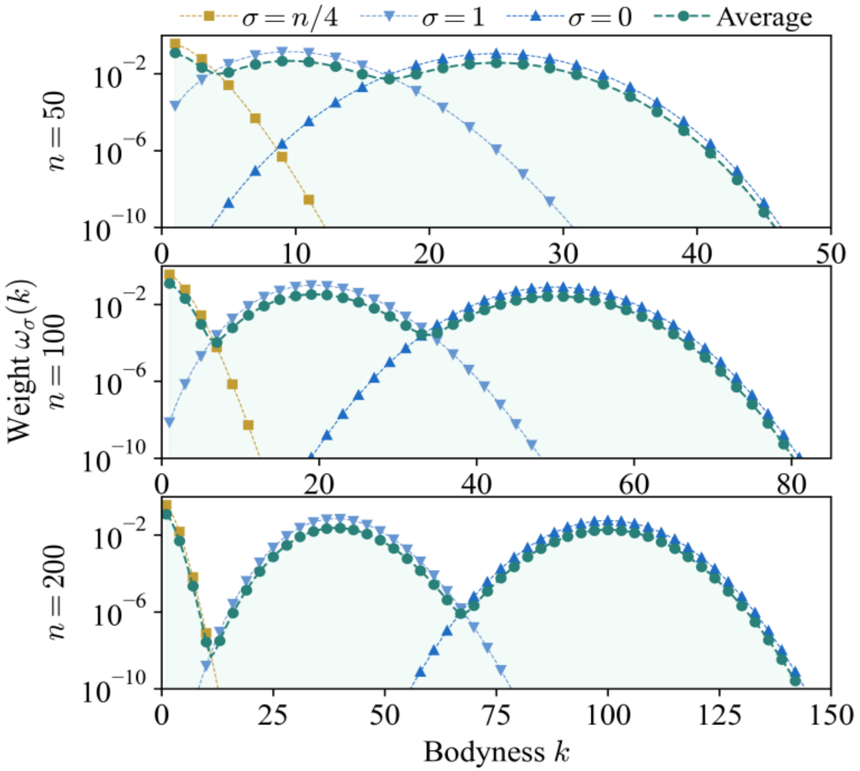
Maximum Mean Discrepancy

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

with

$$K_\sigma(\boldsymbol{x},\boldsymbol{y}) = e^{-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}{2\sigma}} = \prod_{i=1}^{n} e^{-\frac{(x_i-y_i)^2}{2\sigma}}$$





Generative Loss Function

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

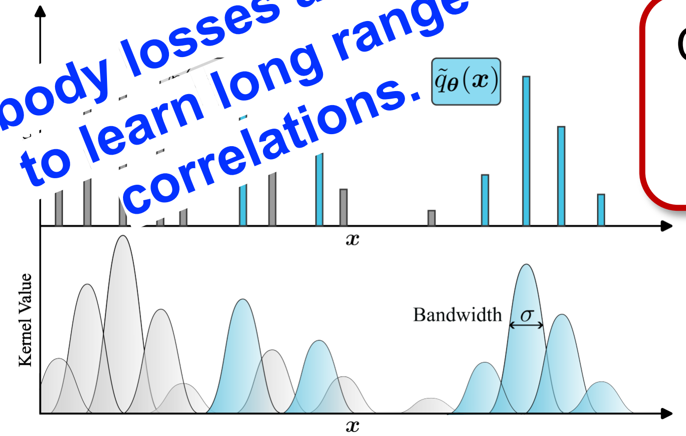
Maximum Mean Discrepancy

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

with

$$K_\sigma(\boldsymbol{x},\boldsymbol{y}) = e^{-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}{2\sigma}} = \prod_{i=1}^{n} e^{-\frac{(x_i-y_i)^2}{2\sigma}}$$

Full-body losses are needed to learn long range correlations.



Equivalent representation as observable ML vs QC

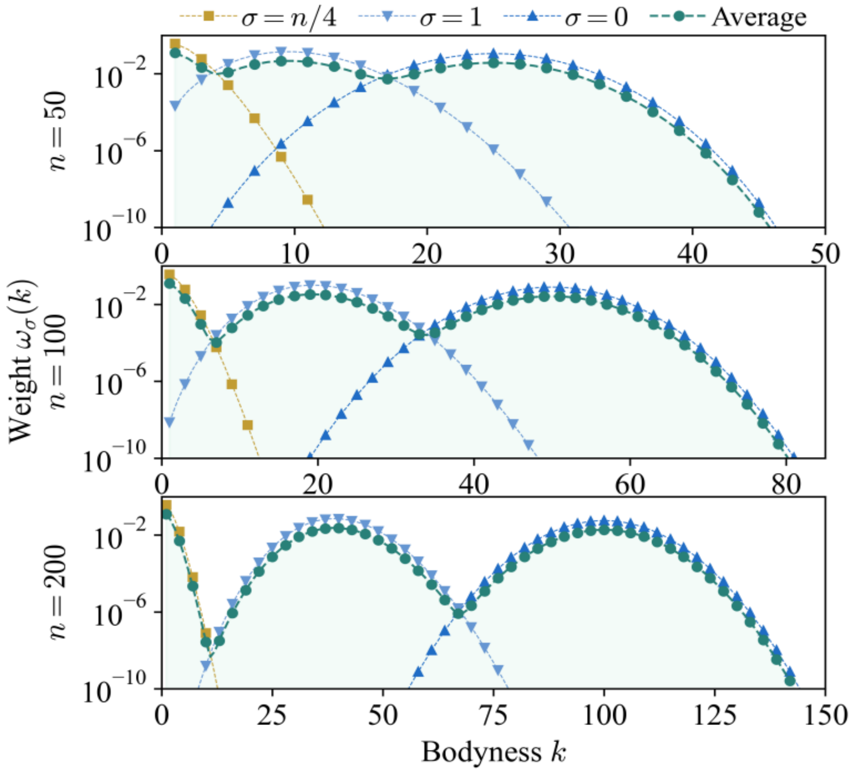
$$\mathcal{M}(\rho,\rho') = \mathrm{Tr}\left[O^{(\sigma)}_{\mathrm{MMD}}(\rho\otimes\rho')\right]$$

$$O^{(\sigma)}_{\mathrm{MMD}} := \sum_{\boldsymbol{x},\boldsymbol{y}} K_\sigma(\boldsymbol{x},\boldsymbol{y})|\boldsymbol{x}\rangle\langle\boldsymbol{x}|\otimes|\boldsymbol{y}\rangle\langle\boldsymbol{y}|\,.$$

to Pauli basis

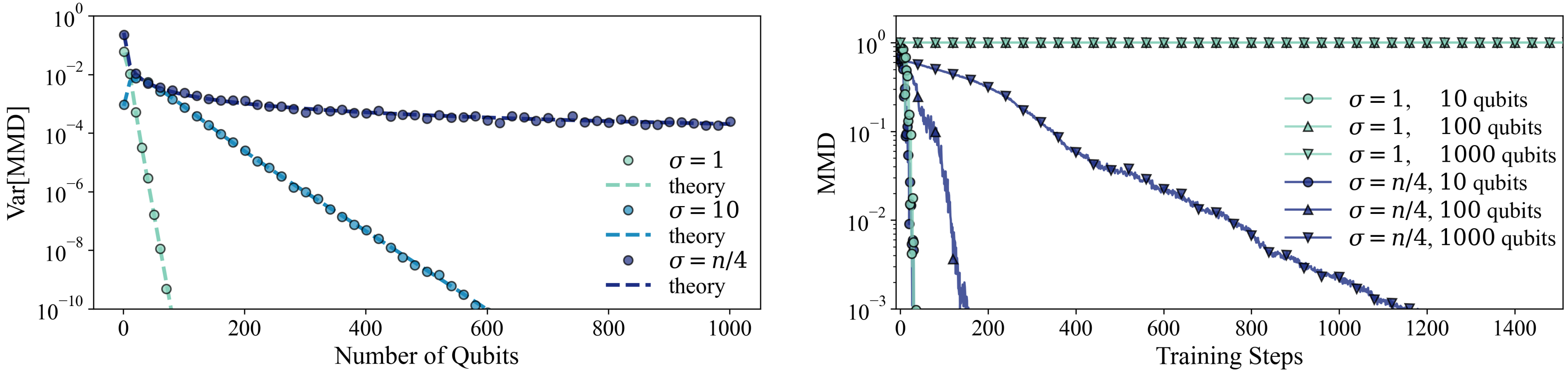
$$O^{(\sigma)}_{\mathrm{MMD}} = \sum_{l=0}^{n}\binom{n}{l}(1-p_\sigma)^{n-l}p_\sigma^l D_{2l}$$

CAN BE **GLOBAL** OR **LOW-BODY/LOCAL**

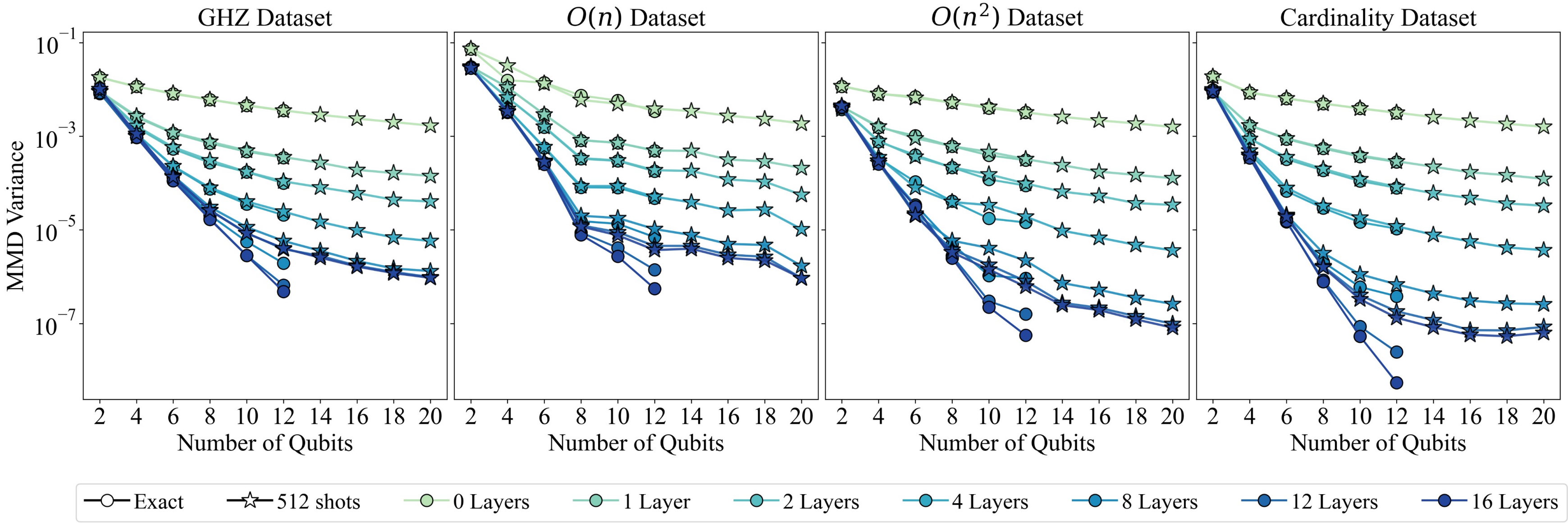


MMD Trainability

Product States

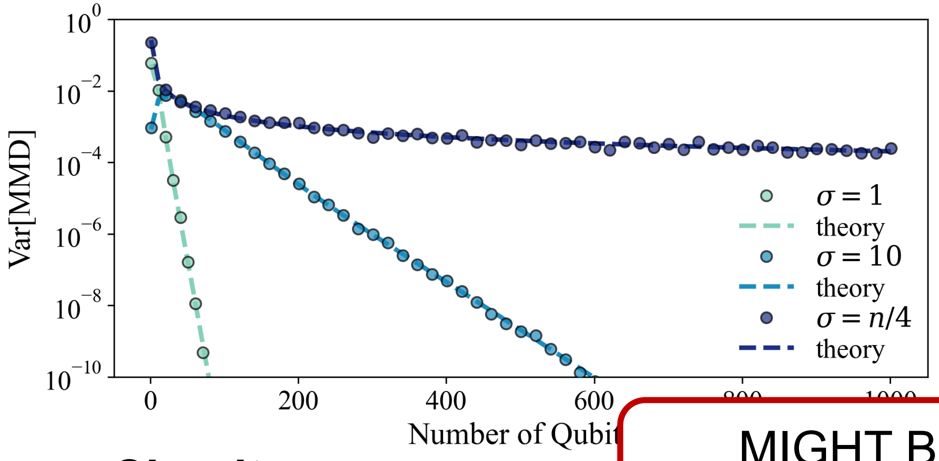


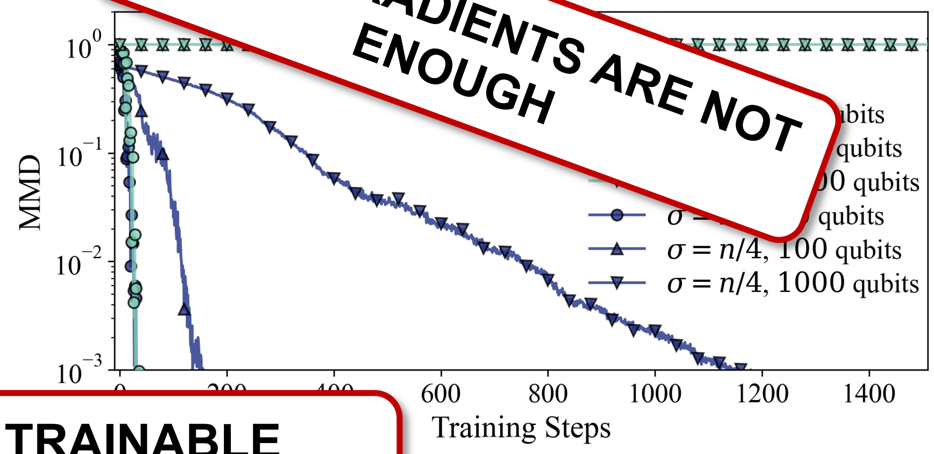
Deeper Circuits



MMD Trainability

Product States

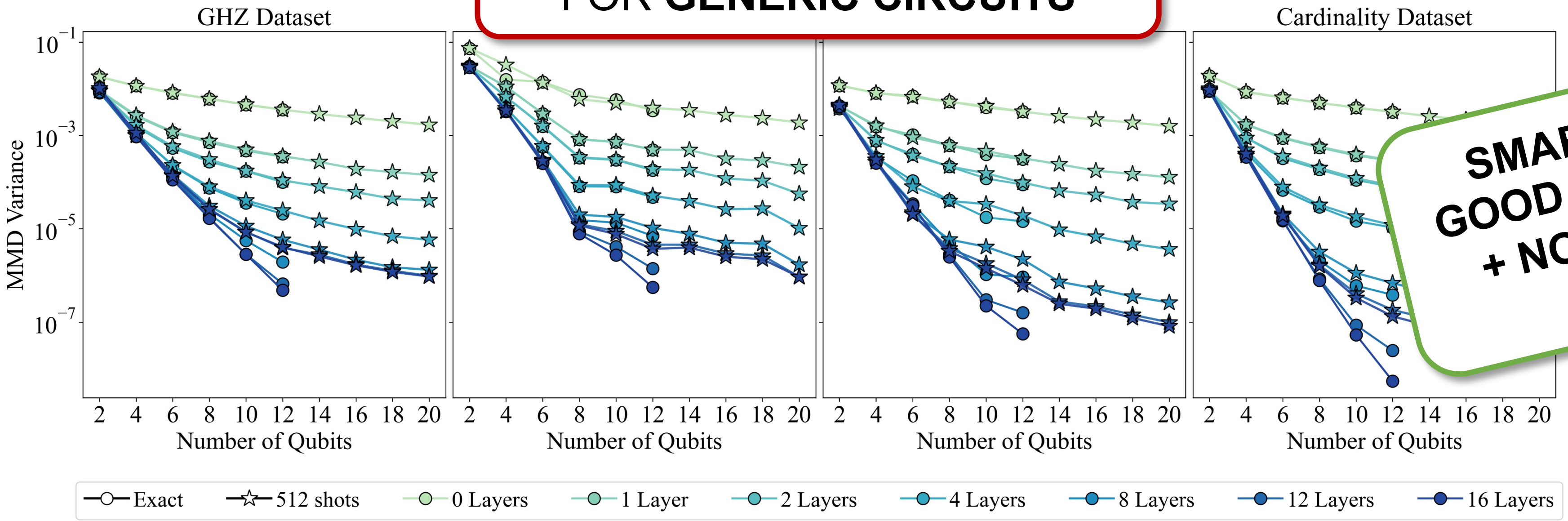




LARGE GRADIENTS ARE NOT ENOUGH

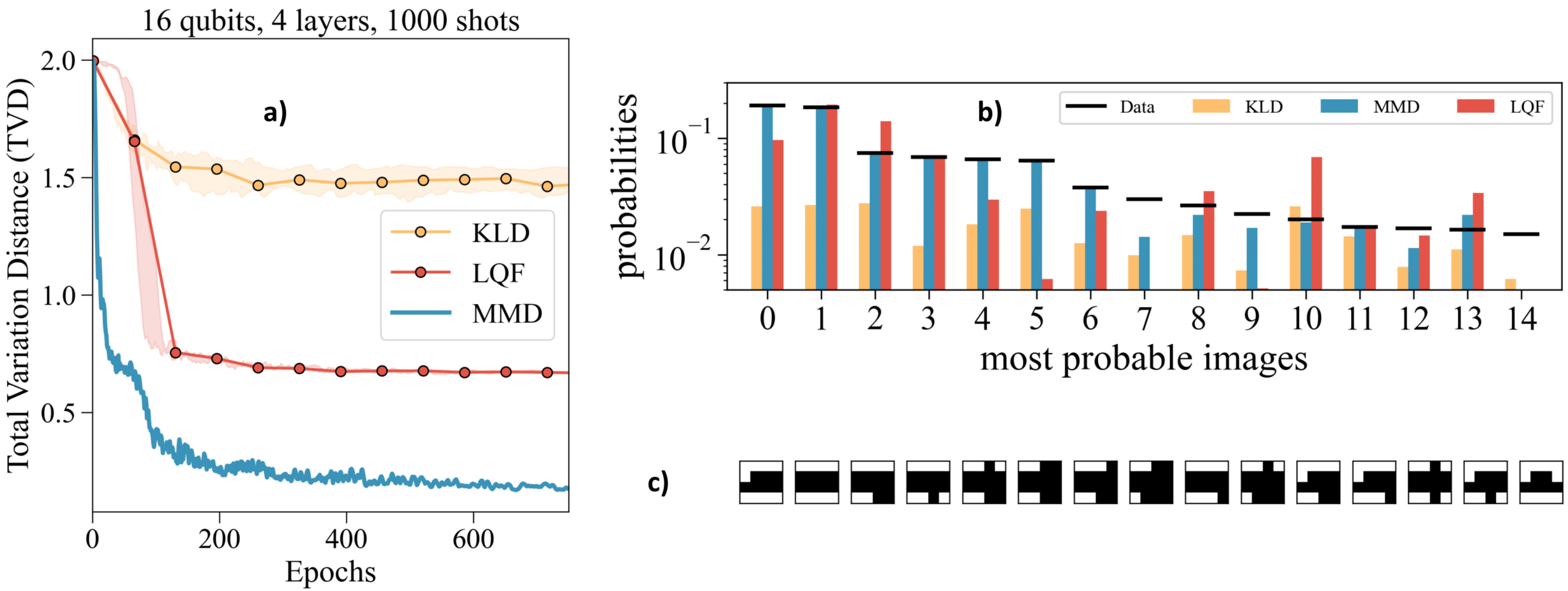
MIGHT BE **TRAINABLE**
FOR **GENERIC CIRCUITS**

Deeper Circuits



SMART CIRCUIT + GOOD INITIALIZATION + NOT TOO LOCAL LOSSES

Final Benchmarks



Summary

Explicit losses are a no-go

Implicit losses can work

Circuit depth	**Explicit loss (pairwise)**		**Implicit loss (MMD)**
	Conventional strategy	Quantum strategy	
Product	**No** (Corollary 2)	**Yes** (Local Quantum Fidelity [31])	**Yes** ($\sigma \in \Theta(n)$, Theorem 2)
Shallow			**Yes** ($\sigma \in \Theta(n)$, Conjecture 1)
Deep		**No** [22, 30]	**No** [22, 30]

Expressivity BP

Quantum strategies?

Paper link

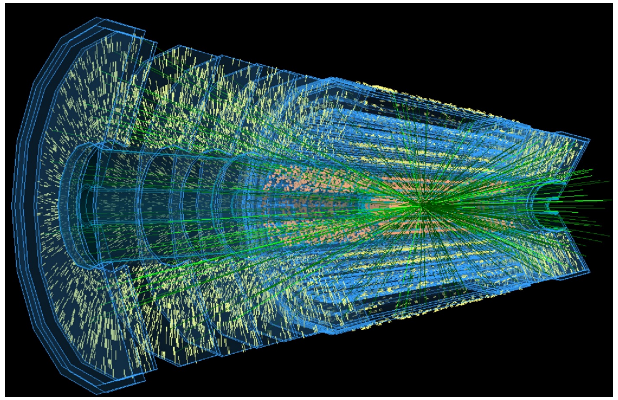
CERN QTI & Motivation

Trainability issues with GM

Summary of QC4HEP applications

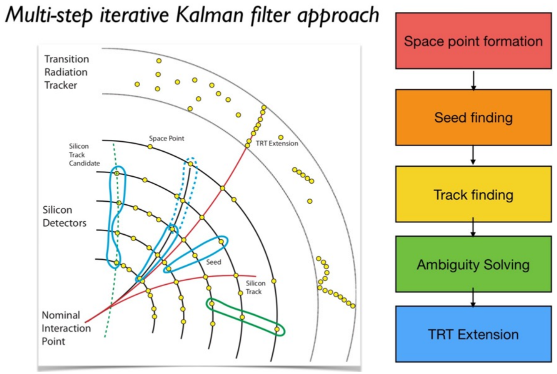
$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + h.c. + \psi_i y_{ij}\psi_j\phi + h.c. + |D_\mu\phi|^2 - V(\phi)$$

Theory

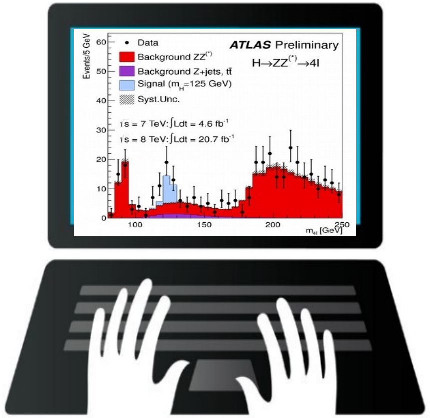


Data acquisition

Data Analysis



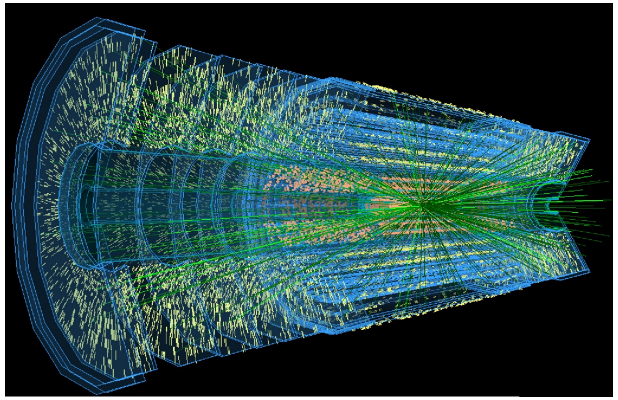
Feature extraction



$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\Psi}\not{D}\Psi + h.c. + \Psi_i y_{ij}\Psi_j\Phi + h.c. + |D_\mu\Phi|^2 - V(\Phi)$$

2013

Theory

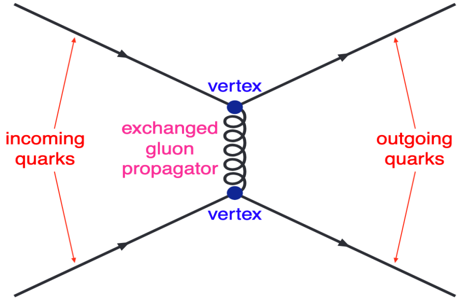


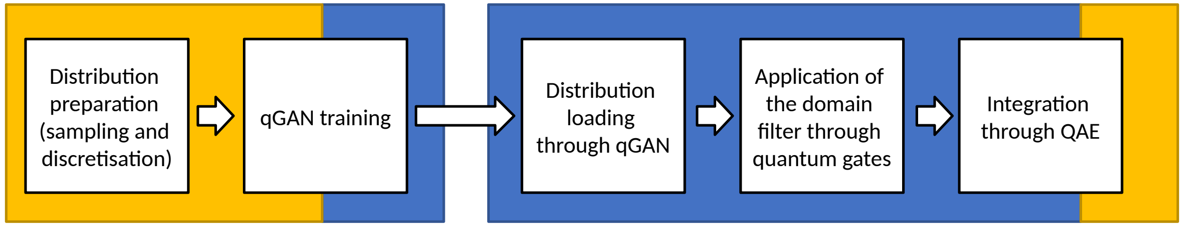
$$\sigma = \frac{1}{F}\int d\Phi |M|^2 \Theta(\Phi - \Phi_c)$$

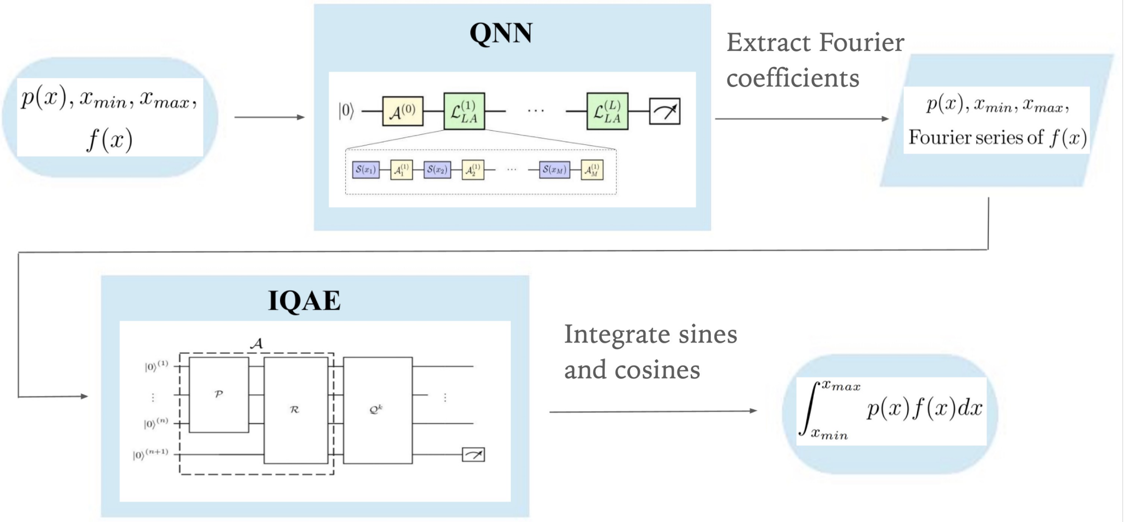
phase-space factor

matrix element

phase-space cuts



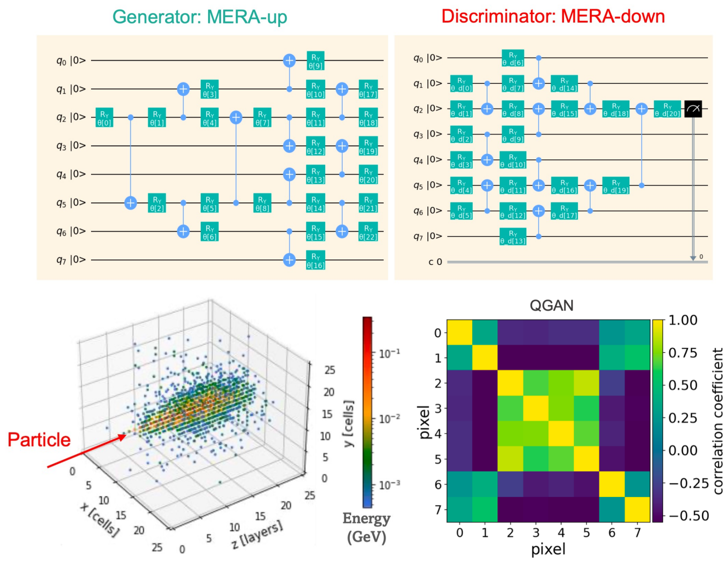




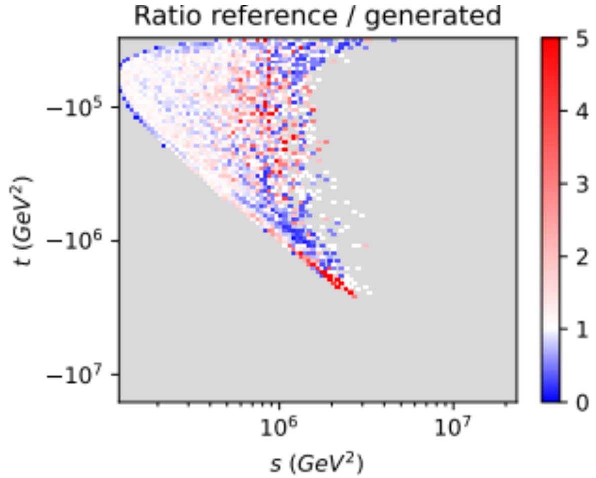
Agliardi, Grossi, Pellen, Prati "**Quantum integration of elementary particle processes**." https://doi.org/10.1016/j.physletb.2022.137228

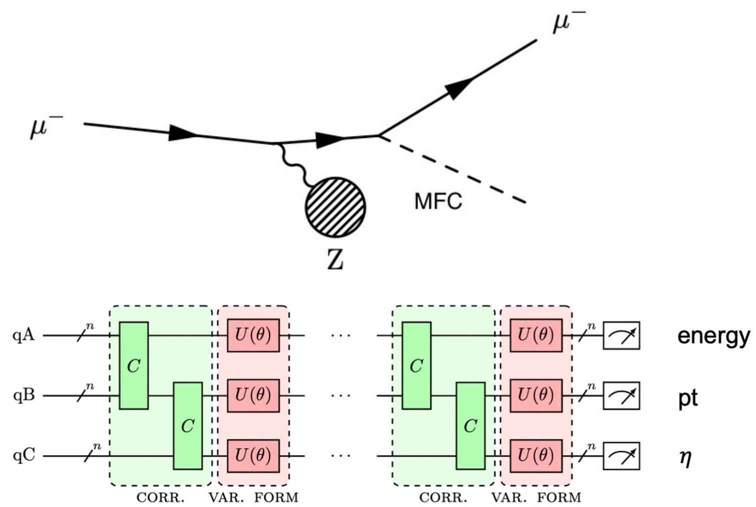
Jorge J. Martinez de Lejarza, Michele Grossi, Leandro Cieri and German Rodrigo: arXiv: 2305.01686

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

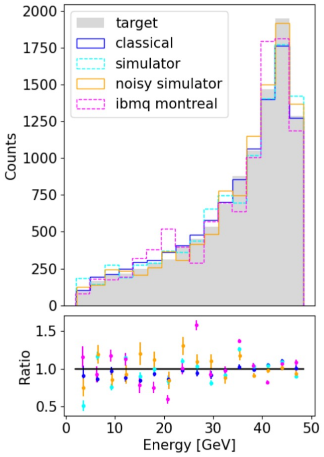


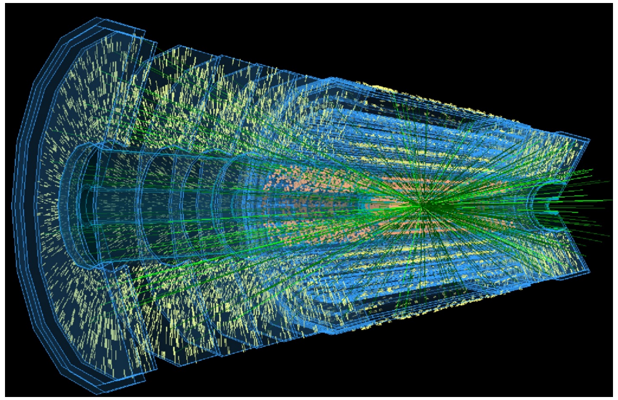
Bravo-Prieto, Carlos, et al. "**Style-based quantum generative adversarial networks for Monte Carlo events**." *Quantum 2022*



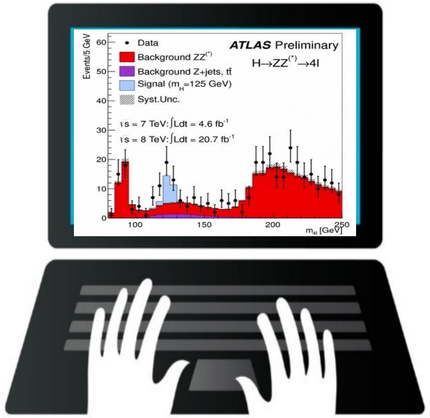


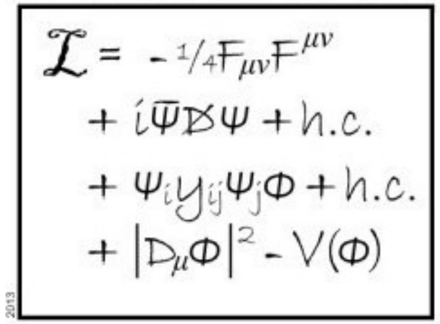
Kiss O., Grossi M. et all., **Conditional Born machine for Monte Carlo events generation**, *Phys. Rev. A* ***106***, *022612 (2022)*



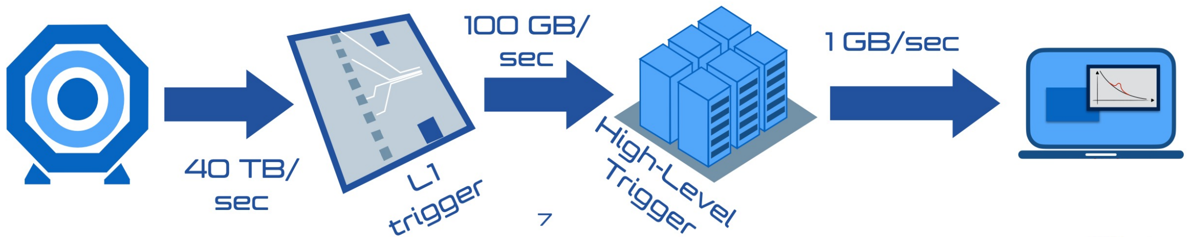


Data acquisition





Are we using the right data?



What if you do not know the signal or where to look for new-physics ?

Data Analysis

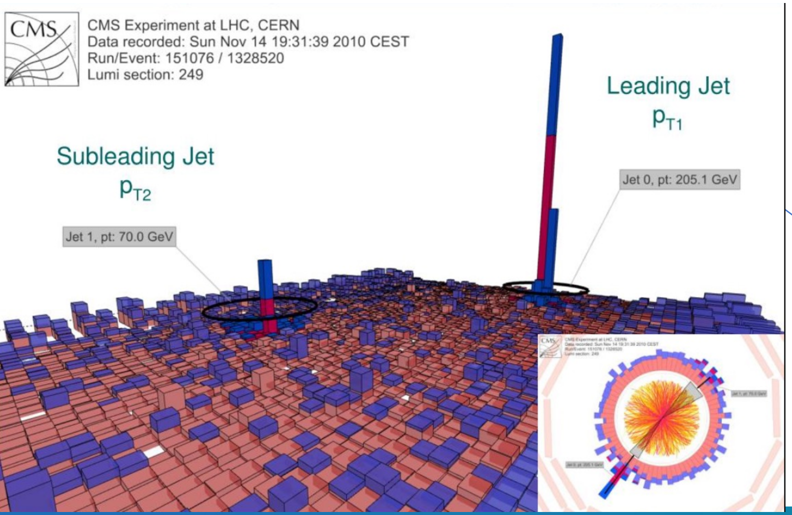
Re-embracing the scientific method: *starts gathering information about nature*

… our baseline is the SM (from 1970!) → let's change the approach

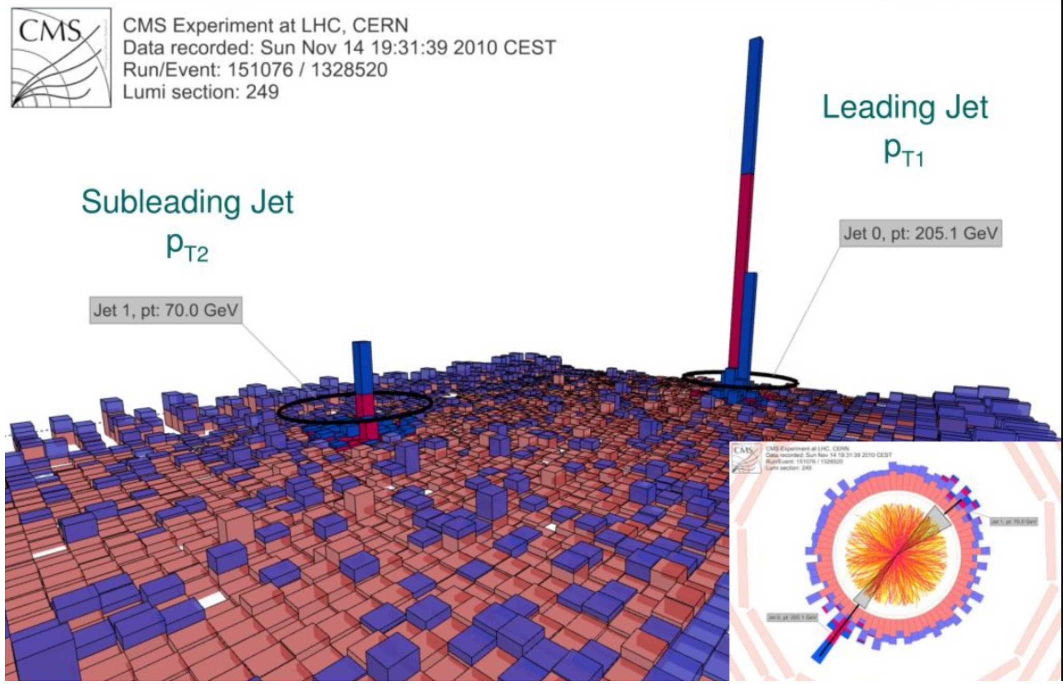
Rather than specifying a signal hypothesis upfront, we could start looking at our data

Based on what we see (e.g., clustering alike objects) we could formulate a signal hypothesis

EXAMPLE: star classification was based on observed characteristics …



Standard Model jet data



Simulate QCD multijet production at the LHC (64 fb $^{-1}$)

Jet is built of **100 highest-p_T particles** within $\Delta R < 0.8$ from its axis.

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

100 particles

Event selection:

- Two jets with $p_T > 200$ GeV and $|\eta| < 2.4$
- $m_{jj} > 1260$ GeV (emulate online selection)
- Each event is represented by its two highest-p_T jets.

Convolutional AutoEncoder compresses particle jet learning the **internal structure**

- Trained on background events

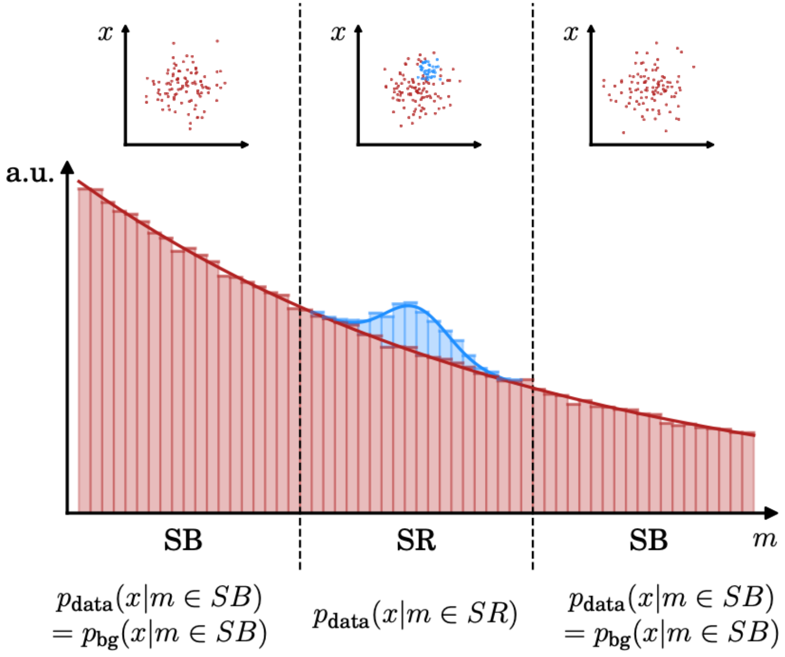
$$\mathbb{R}^{300} \rightarrow \mathbb{R}^{\ell}, \ell = 4, 8, 16$$

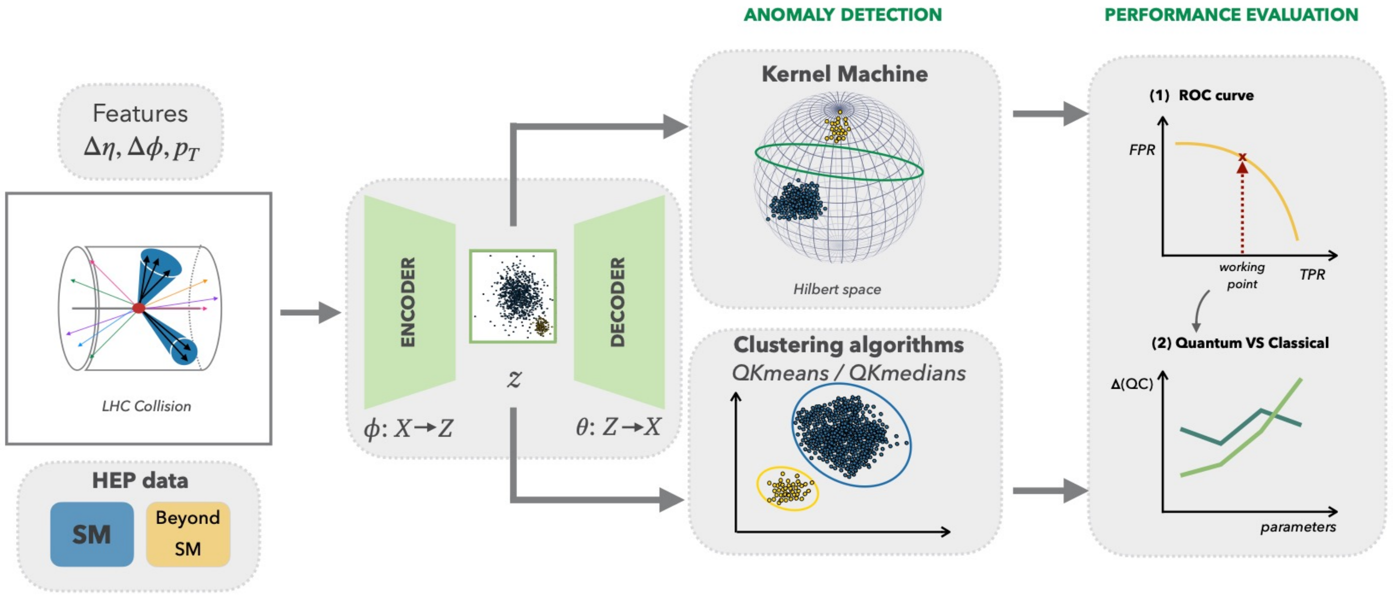
A typical hybrid QML workflow

Paper Code

Anomaly detection can point to new physics at the LHC **Model-agnostic**!

- *Narrow and Broad Graviton resonance G → W^+W^- → Multi-jet final state*
- *New scalar boson A → HZ → ZZZZ (Multi-jet final state)*





«Normal» training data | Data compression | Quantum algorithm | Output

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

Wózniak, Belis, Grossi, Tavernelli, Vallecorsa et al. - https://arxiv.org/abs/2301.10780

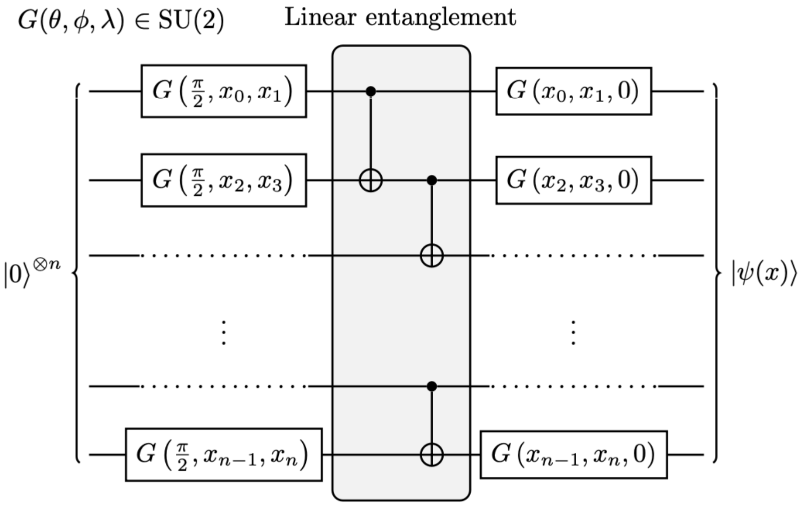
Results

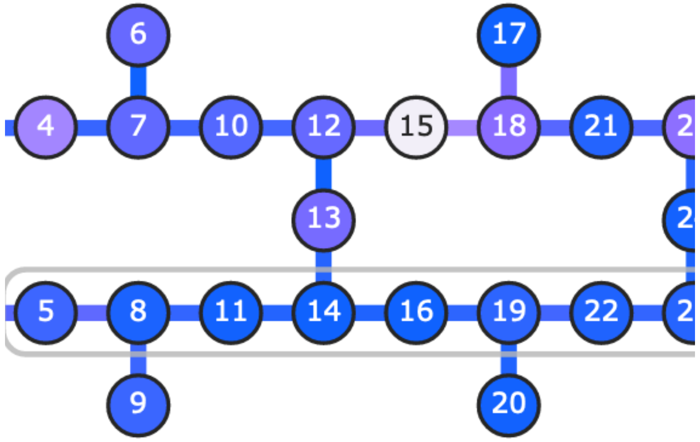
Comparison to best-performing classical algorithm with similar complexity trained and tested on the same data

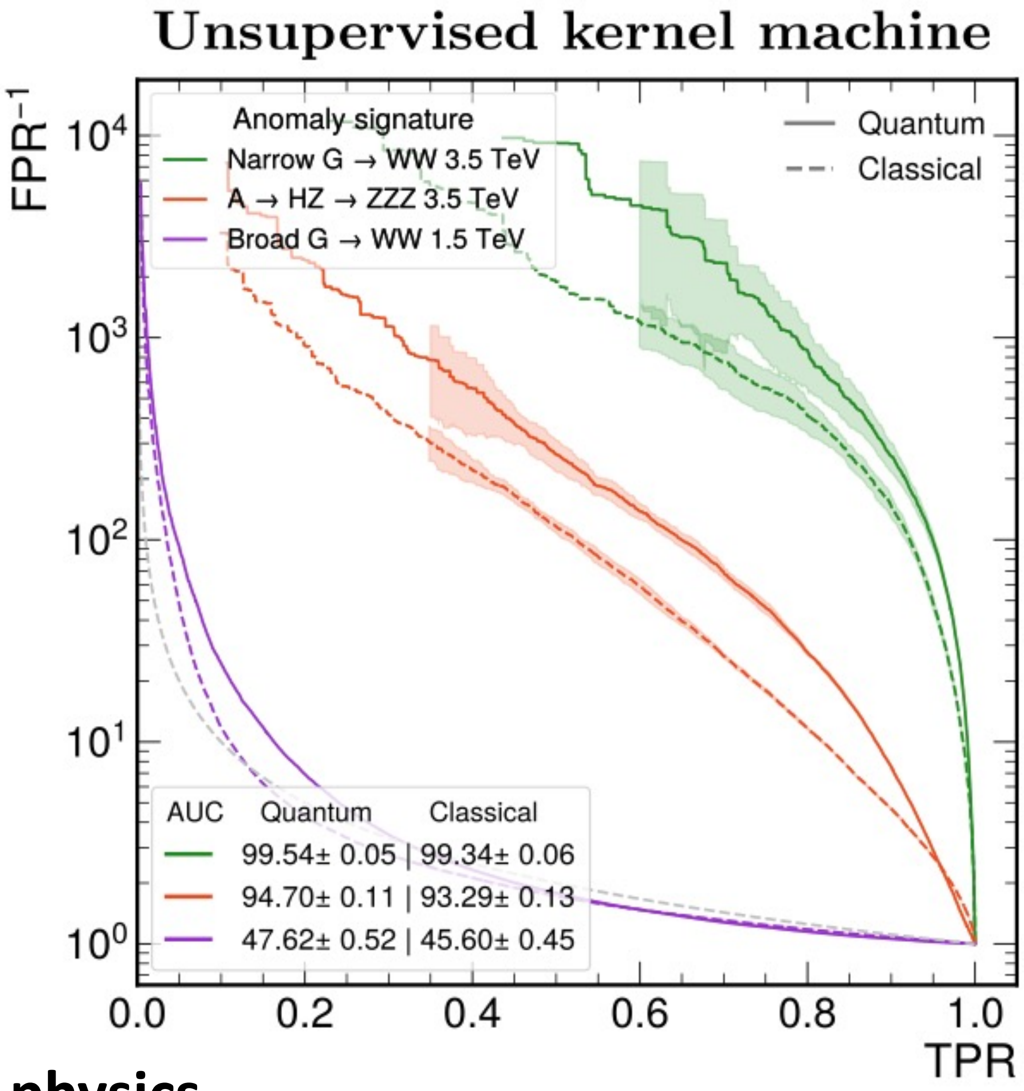
- RBF –based SVM

AUC shows **consistent advantage** for quantum algorithm

Evaluate performance at **typical working**, where ε_s = 0.6, 0.8







Quantum kernel machine works best for more complex physics

Outlook and Questions

- ***Supervised searches served HEP well so far***
- ***We need new directions to search for as an alternative workflow, where data guide us***
- **Studying the behaviour of models in the NISQ regime is useful**

- **Can we reduce the impact of data reduction techniques?**
- **Can we find the right balance of trainability vs generalization?**
- **Can Quantum Anomaly Detection being a good candidate?**
- **What is the role on Quantum Data for HEP?**

CERN 19-24 November 2023

Annual international conference focusing on the interdisciplinary field of quantum technology and machine learning

QTML

Quantum Techniques in Machine Learning

CERN QTI

https://quantum.cern/