ML4JETS 2024, LPNHE, PARIS

META-LEARNING QUANTUM JET PROPERTIES WITH QUANTUM GENERATIVE MODELS

Vincent Croft [1], Michele Grossi [2], **Yacine Haddad [3]**, Cenk Tüysüz [4]

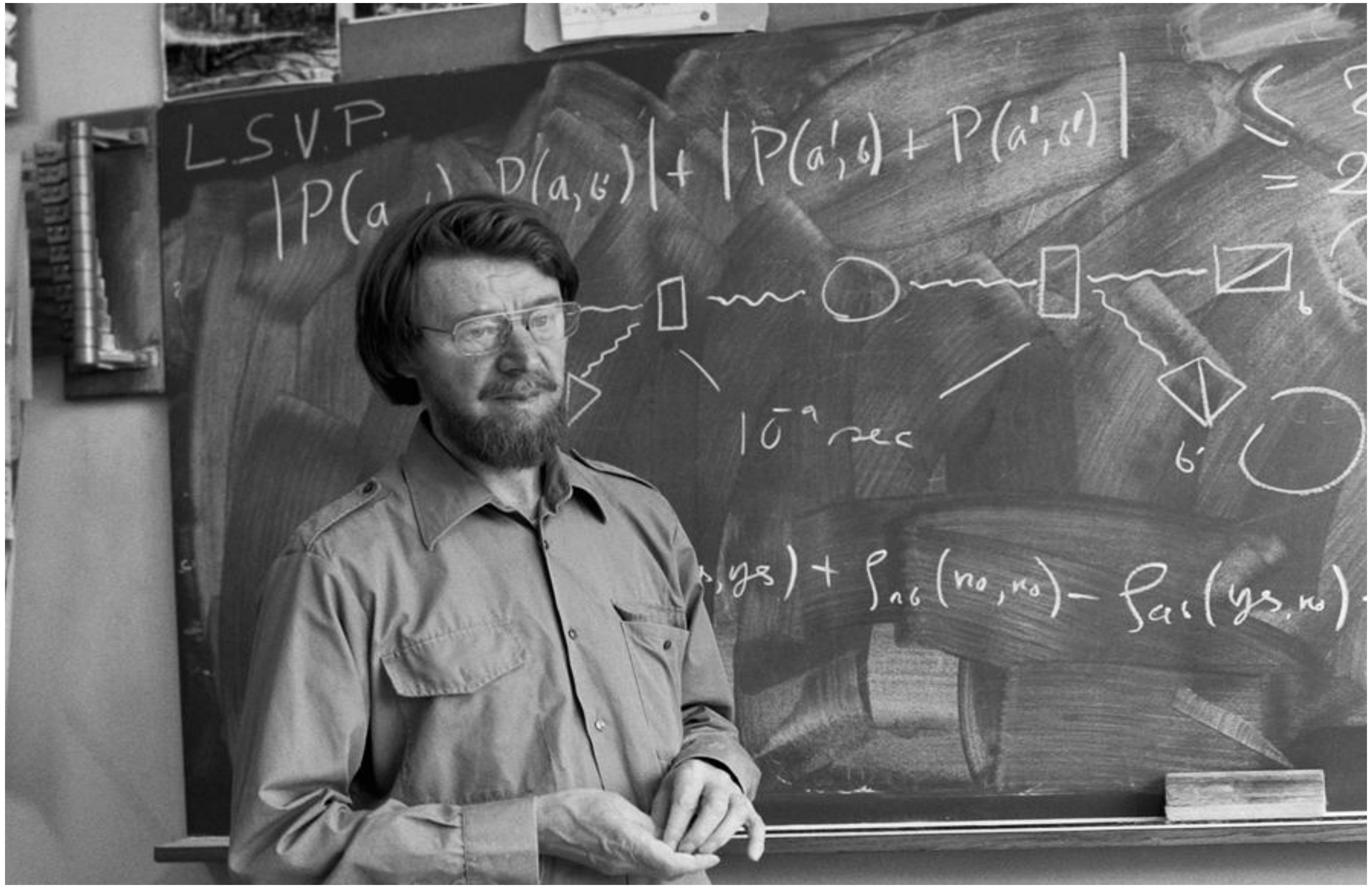
[1] Leiden Institute of Advanced Computer Science (LIACS), [2] QTI-CERN,
[3] Northeastern University, [4] CQTA DESY, Humboldt-Universität zu Berlin

Background Image: Allegory of Science
Jean Baptiste Simeon Chardin

INTRODUCTION

- **Objective:** To explore quantum generative models in jet substructure modeling. Using two different approaches:
  - **Quantum Boltmann machines**
  - **Quantum Generative Adverserial Networks**

THE SECOND QUANTUM REVOLUTION
& QUANTUM COMPUTING



- **Quantum Mechanic principles** are exploited to develop new technology
- Create **"artificial" quantum states** for a range of applications (single photons, trapped ions, superconductors, etc.)
- 1964: **Bell inequalities** prove that no theory based on **local hidden variables** (realism) can reproduce QM results
- Major step confirming the possibility of using distant entangled photons as a quantum information resource

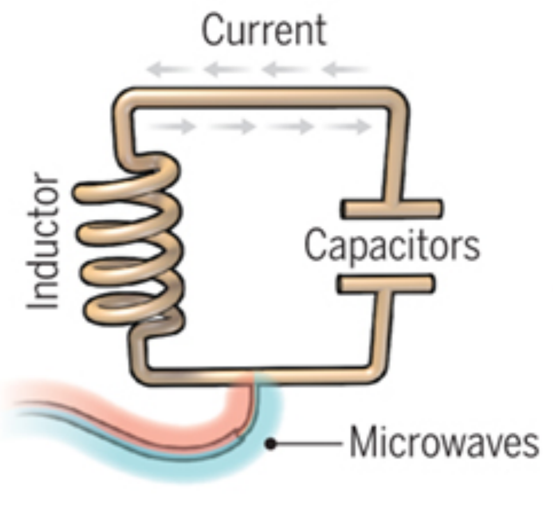
a paradigm shift in our ability to harness the unique properties of quantum mechanics for practical purposes

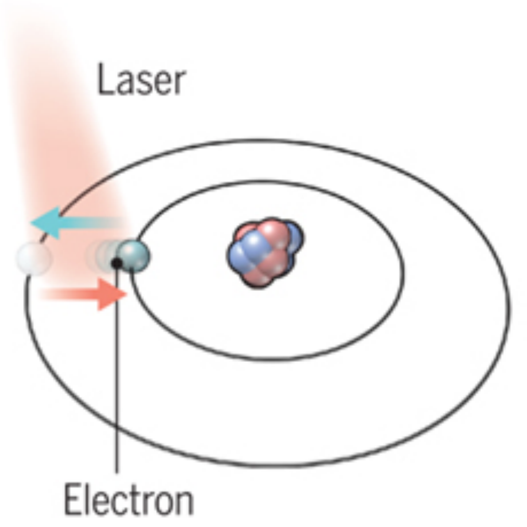
QUANTUM COMPUTING TECHNOLOGY OVERVIEW

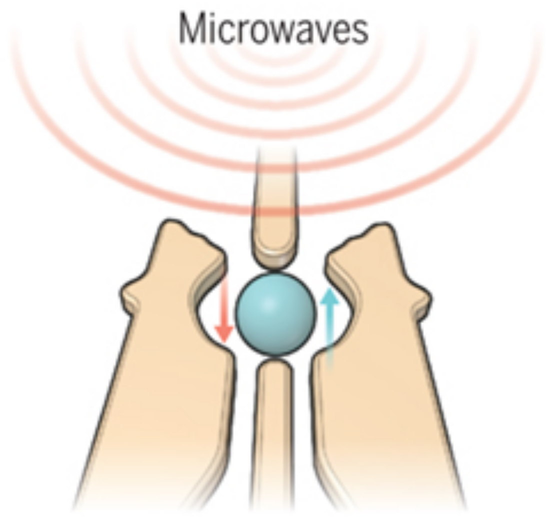
A bit of the action

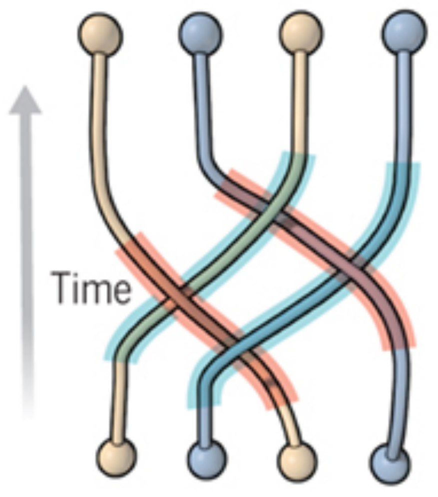
In the race to build a quantum computer, companies are pursuing many types of quantum bits, or qubits, each with its own strengths and weaknesses.

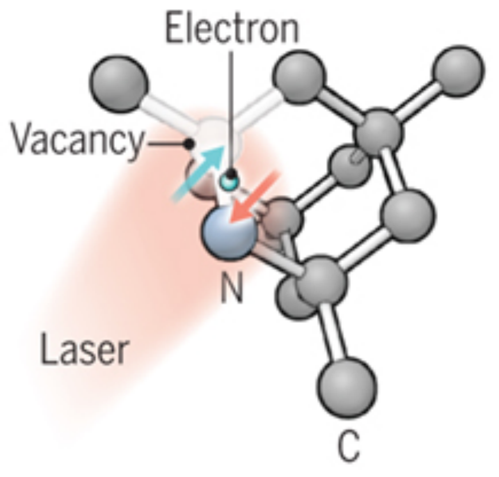
Adapted from Science, Dec 2016

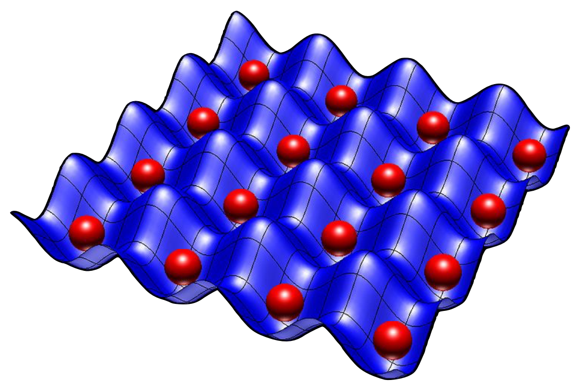


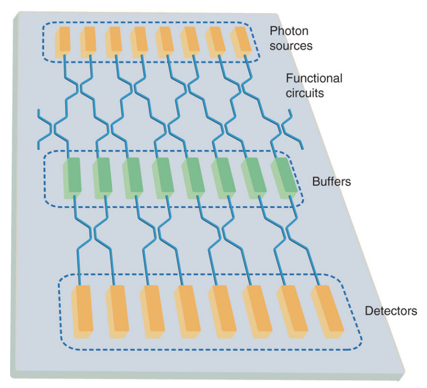












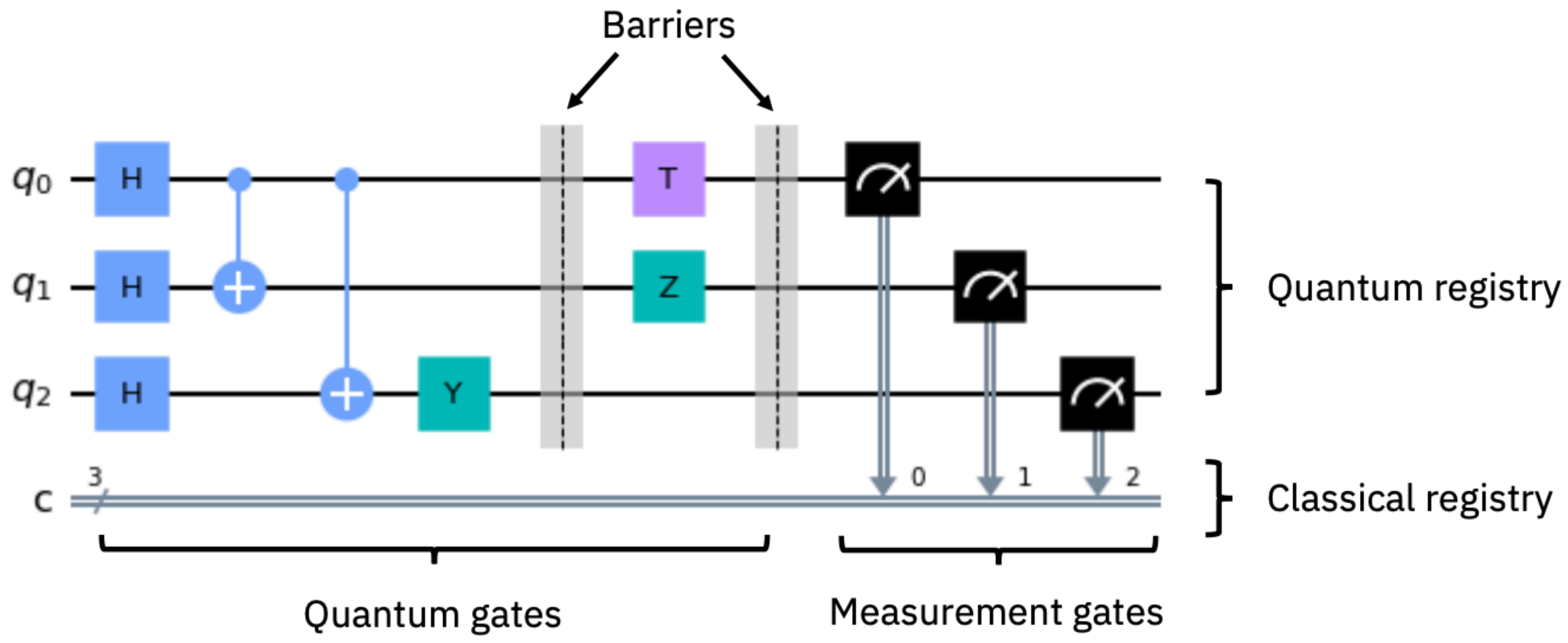
	Superconducting loops	Trapped ions	Silicon quantum dots	Topological qubits	Diamond vacancies	Silicon quantum dots	Phtonics
	A resistance-free current oscillates back and forth around a circuit loop. An injected microwave signal excites the current into super-position states.	Electrically charged atoms, or ions, have quantum energies that depend on the location of electrons. Tuned lasers cool and trap the ions, and put them in superposition states.	These "artificial atoms" are made by adding an electron to a small piece of pure silicon. Microwaves control the electron's quantum state.	Quasiparticles can be seen in the behavior of electrons channeled through semi-conductor structures. Their braided paths can encode quantum information.	A nitrogen atom and a vacancy add an electron to a diamond lattice. Its quantum spin state, along with those of nearby carbon nuclei, can be controlled with light.	Neutral atoms, like ions, store qubits within electronic states. Interaction through excitation to Raydberg states	Photonic qubits interact via linear elements
Longevity (seconds)	0.00005	>1000	0.03	N/A	10	1	N/A
Logic success rate	99.4%	99.9%	~99%	N/A	99.2%	99.6%	N/A
Number entangled	9	14	2	N/A	6	99.6%	N/A
Company support	Google, IBM, Quantum Circuits	ionQ	Intel	Microsoft, Bell Labs	Quantum Diamond Technologies	QuEra, Atom Computing	Xanadu, Psi Corp
Pros	Fast working. Build on existing semiconductor industry.	Very stable. Highest achieved gate fidelities.	Stable. Build on existing semiconductor industry.	Greatly reduce errors.	Can operate at room temperature.	Many qubits, 2D and maybe 3D.	Linear optical gate integrated on-chip
Cons	Collapse easily and must be kept cold.	Slow operation. Many lasers are needed.	Only a few entangled. Must be kept cold.	Existence not yet confirmed.	Difficult to entangle.	Lasers needed, spagetti physics, and atoms escaps.	No memory, not clear how to scale

Note: Longevity is the record coherence time for a single qubit superposition state, logic success rate is the highest reported gate fidelity for logic operations on two qubits, and number entangled is the maximum number of qubits entangled and capable of performing two-qubit operations.

QUANTUM INFORMATION IN NUTSHELL

Composing quantum gates: quantum circuits

- Set of actions to be performed to the selected qubits
  - qubits initialization
  - single-qubit gates, multi-qubit gates
  - measurements



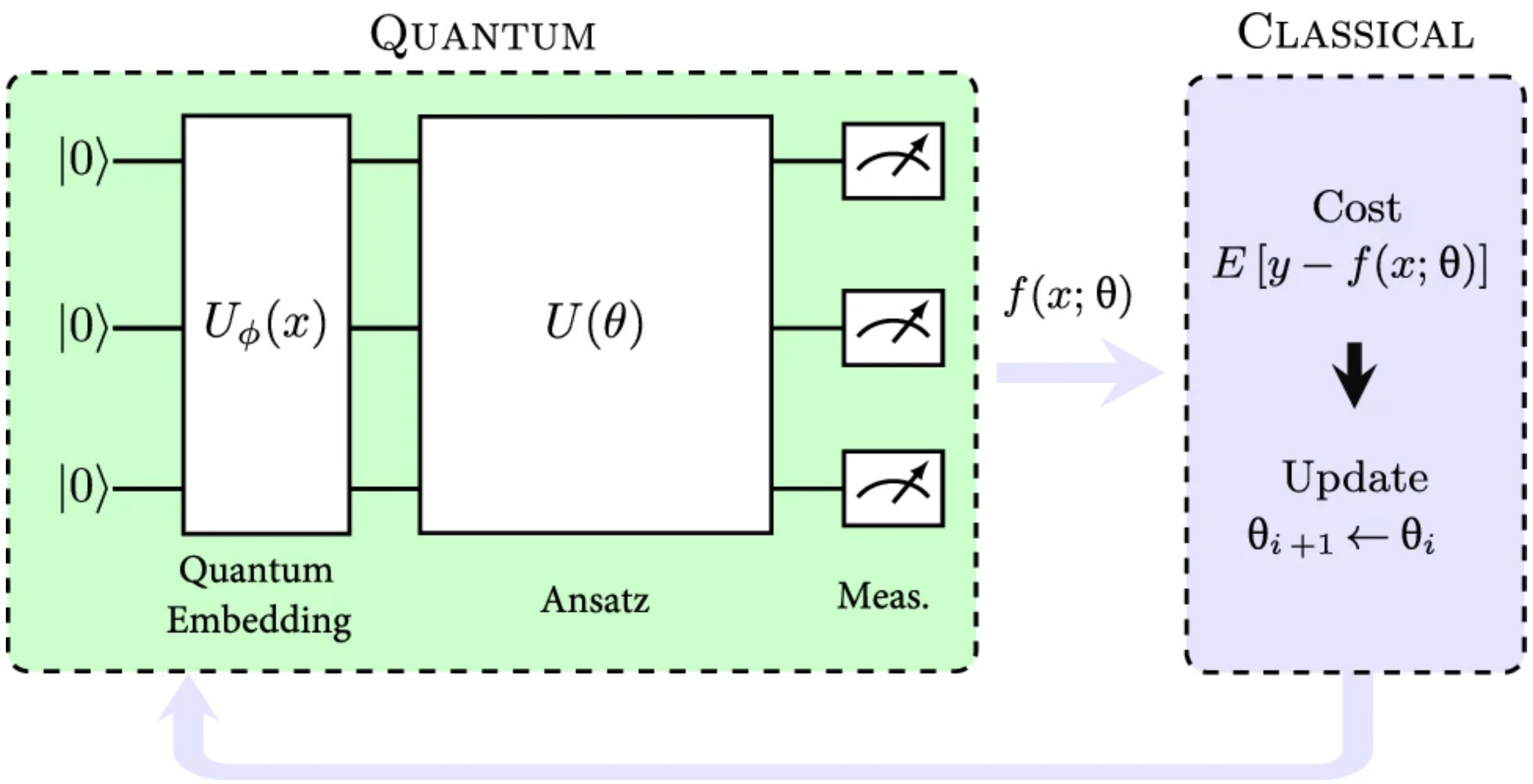
Principles of quantum computation

- Quantum algorithm: set of quantum circuits performing certain task
  - Purely quantum, e.g. *Shor*
  - Hybrid classical-quantum, e.g. *VQE*

- Quantum Simulation: simulation of time evolution of quantum system
  - Analog Simulator
  - Digital Simulator: quantum logic gates, more flexible

QUANTUM MACHINE LEARNING 101

- Can be used in Noisy Intermediate Scale Quantum Devices
  - **Circuit width:** limited number of qubits
  - **Circuit depth:** limited number of operations per qubit (small decoherence times)
  - **Hardware noise**



Macaluso, A. **Quantum Supervised Learning**. Künstl Intell (2024). https://doi.org/10.1007/s13218-024-00856-7

Variational algorithms - EXPLICIT

Gradient-free or gradient-based optimisation

Data-embedding can be learned

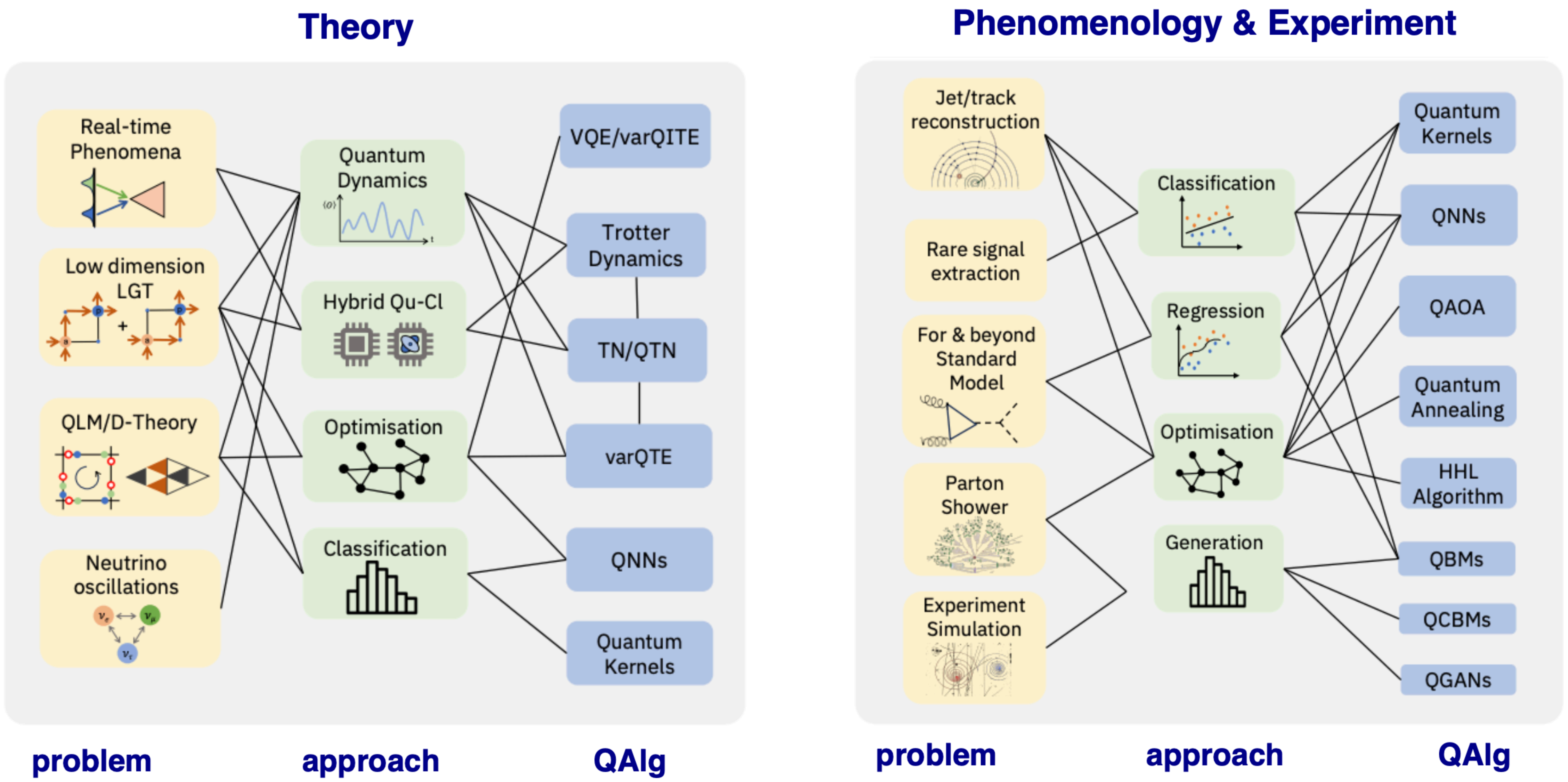
Ansatz design can laverage data symmetries

Current Hardware limitations:

Feature reduction required for realistic datasets

QUANTUM COMPUTING IN HEP

Alberto Di Meglio et al. PRX Quantum 5, 037001



Relevance in HEP requires (eventually) the quantum algorithms to **outperform classical algorithms (including ML/AI/HPC)** for the same task

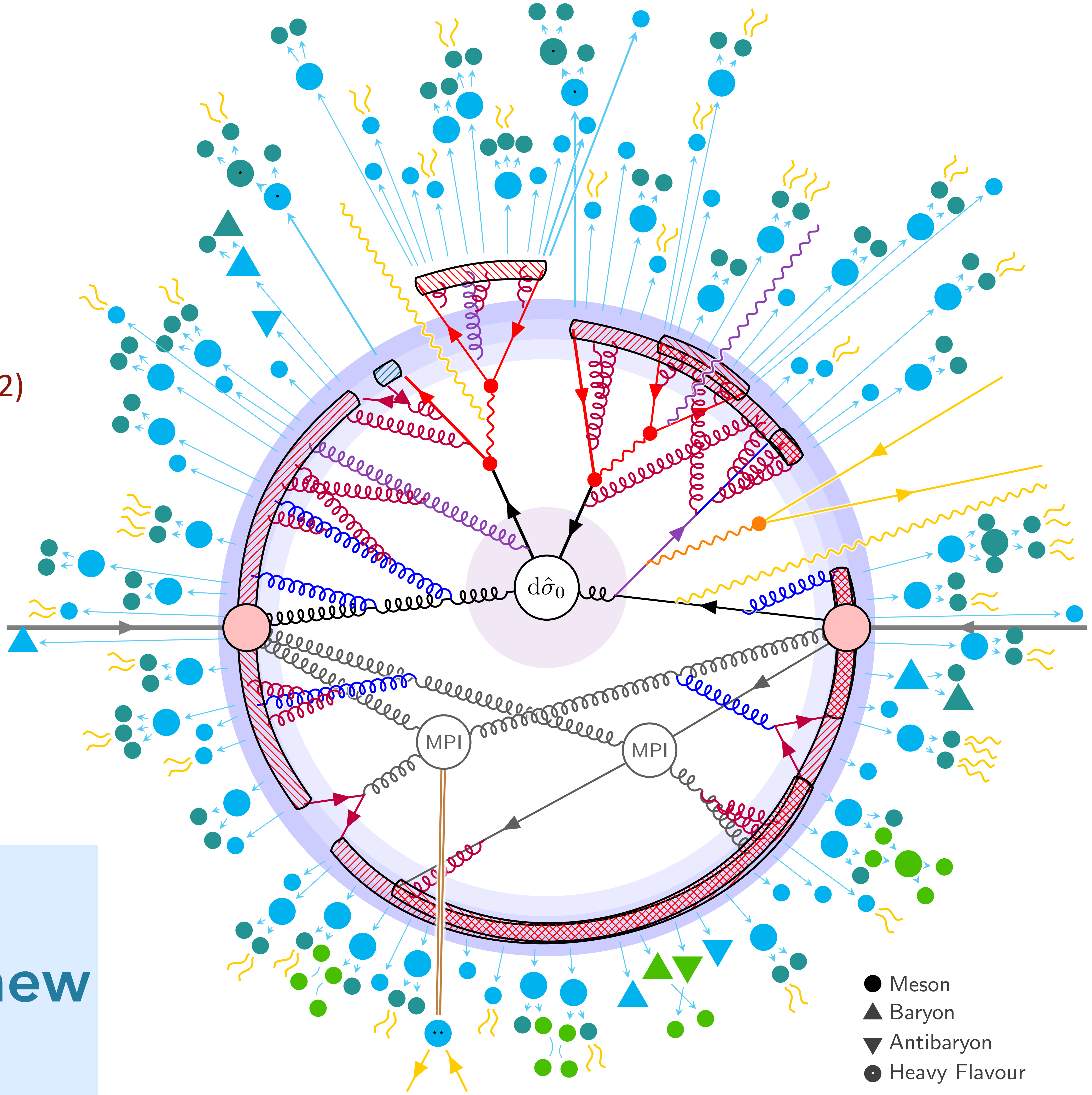
WHY QML FOR JETS?

- QC can efficiently compute scattering amplitudes, at all possible field configurations, a task exponentially challenging for classical computers.

Jordan, Lee, Preskill, Science 336,1130-1133(2012)

- QML, particularly Quantum Generative Models, can effectively approximate Hamiltonians, enabling efficient simulation of quantum systems.

Aim: train a quantum circuit to reproduce jets observables and inner correlations → provide a new avenue for understanding QCD and jets



Christian Bierlich et al, arXiv:2203.11601

QUANTUM BOLTZMANN MACHINES (QBM)

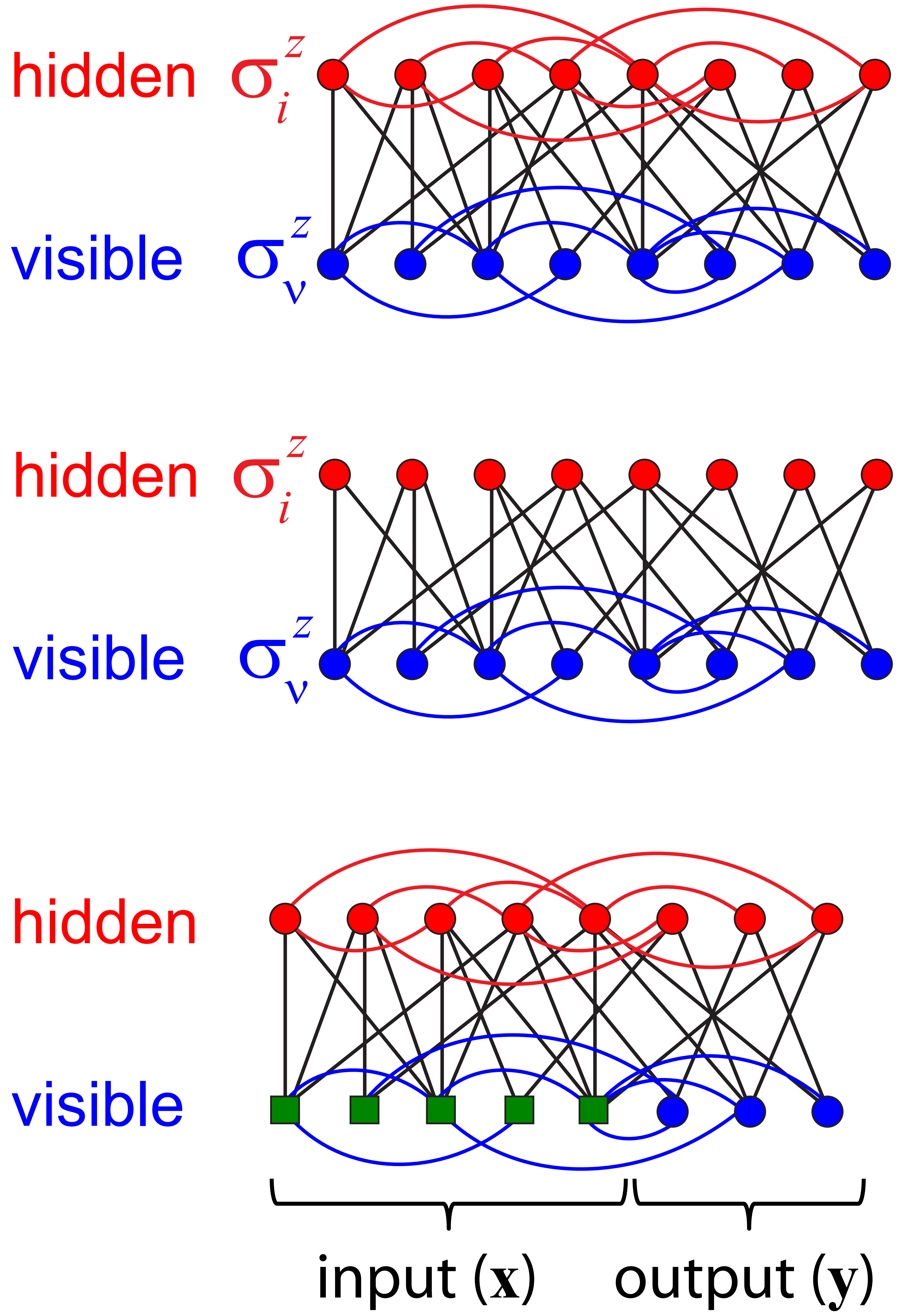
Amin, Mohammad H., et al. "**Quantum boltzmann machine.**" *Physical Review X* 8.2 (2018): 021050.

A **Quantum Boltzmann Machine** (QBM) is a quantum machine learning model that can be used for generatively modelling classical and quantum data.

Quantum Botzmann Machine Replaces the energy function with the Hamiltonian of qubit graph. Two types: Fully-visible models (only visible units). Restricted models (visible and hidden units).

Advantages over Classical BMs:

- Potential to model complex, high-dimensional distributions.



Relevance to Jet Substructure Modelling

- Jets involve complex quantum correlations and interference effects
- QBMs can potentially capture these effects more effectively than classical models

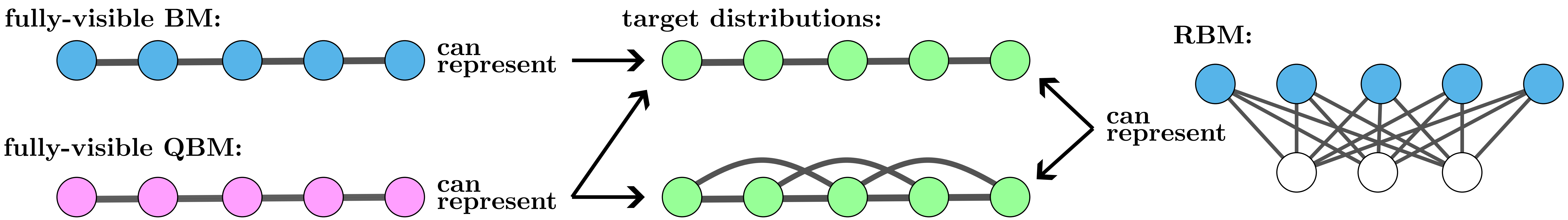
APPLICATION OF QBMS TO JET MODELLING

Learning to generate high-dimensional distributions with low-dimensional quantum Boltzmann machines

Cenk Tüysüz,[1,2,*] Maria Demidik,[1,3] Luuk Coopmans,[4] Enrico Rinaldi,[5,6]
Vincent Croft,[7] Yacine Haddad,[8] Matthias Rosenkranz,[4] and Karl Jansen[3,1]

[1] *CQTA, Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, 15738 Zeuthen, Germany*
[2] *Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany*
[3] *Computation-Based Science and Technology Research Center, The Cyprus Institute, 20 Kavafi Street, 2121 Nicosia, Cyprus*
[4] *Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom*
[5] *Quantinuum K.K., Otemachi Financial City Grand Cube 3F, 1-9-2 Otemachi, Chiyoda-ku, Tokyo, Japan*
[6] *Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS) Program, RIKEN, Wako, Saitama 351-0198, Japan*
[7] *Applied Quantum Algorithms Leiden $\langle aQa^L \rangle$ and Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Niels Bohrweg 2, 2333 CA Leiden, Netherlands*
[8] *Department of Physics, Northeastern University, Boston, MA 02115 USA*

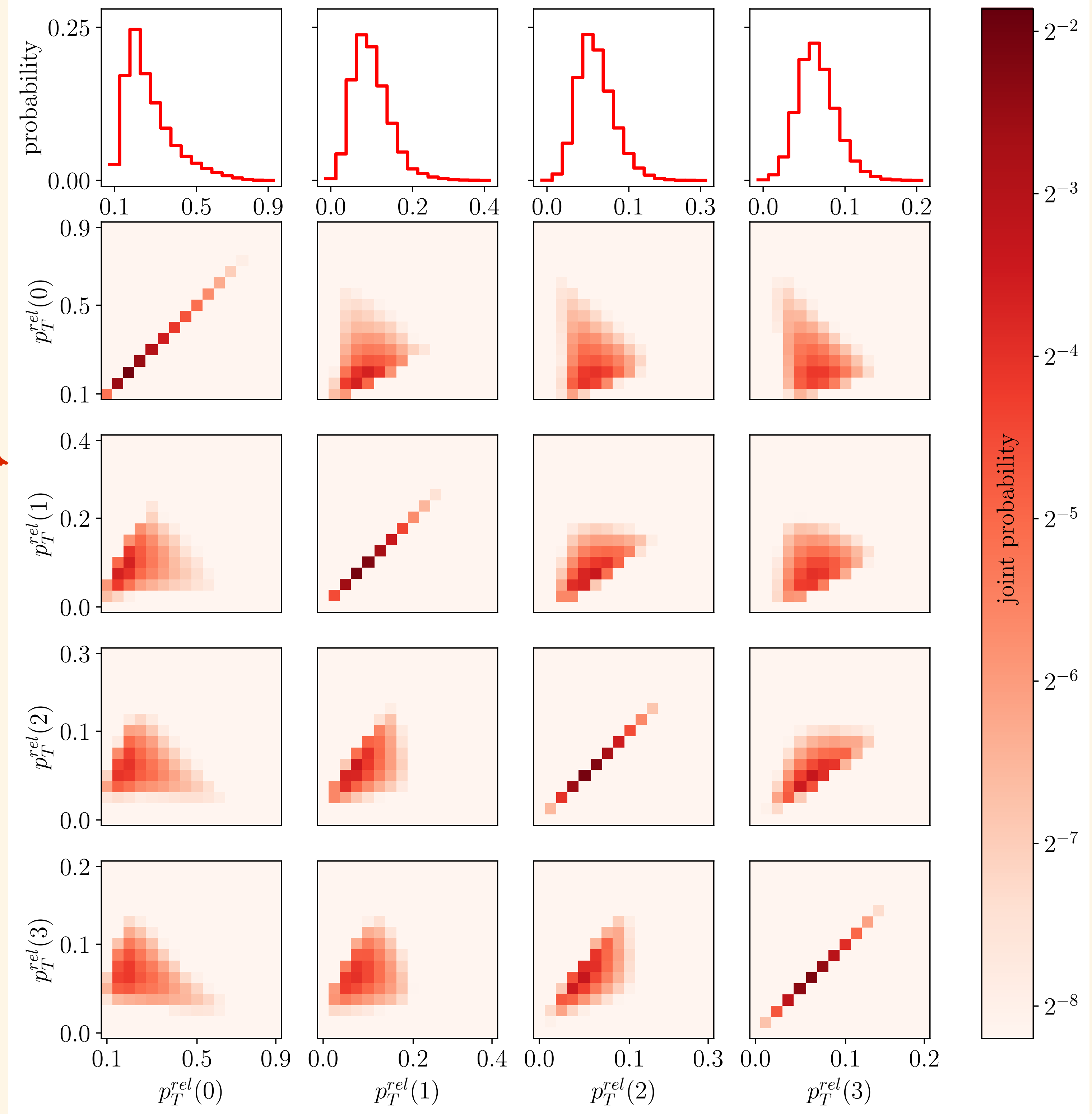
arXiv:2410.16363



- QBMs effectively generate reduced-size particle jet events.
- Capable of capturing complex correlations that classical models miss.
- First application of QBMs to particle jet event generation.

FRAMEWORK AND TRAINING

- **Quantum Relative Entropy as Loss Function**
  $S(\eta\,||\,\rho_\theta) = \mathrm{Tr}(\eta \log \eta) - \mathrm{Tr}(\eta \log \rho_\theta)$
  - $\eta \leftarrow$ Target density matrix (data)*
  - $\rho_\theta = \dfrac{e^{-\beta H_\theta}}{\mathrm{Tr}(e^{-\beta H_\theta})} \leftarrow$ Model density matrix (QBM)
  - Model Hamiltonian† $H_\theta = \sum_i \theta_i H_i$
- **Objective is to minimise $S(\eta\,||\,\rho_\theta)$** to train the QBM
- For classical BMs, minimising quantum relative entropy is equivalent to minimising KL divergence



Taregt distributions for 4 particles and $n_{bins} = 16$

Distributions of m highest p_T^{rel} particle constitutents of W-bosons jets from the JetNet datasets → binned in multi-dimentional historgam

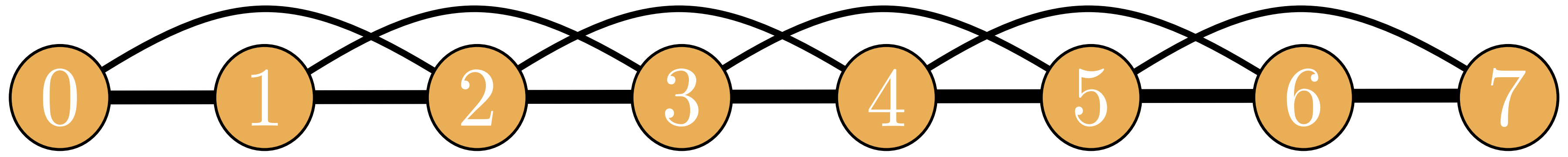
(*) encoding in the computational basis: $\eta = |\psi\rangle\langle\psi|, \quad |\psi\rangle = \sqrt{p(s)}e^{i\alpha(s)}|s\rangle$

(†) $H_\theta = \sum_{k\in\mathscr{P}_1}\sum_{i\in\mathscr{V}} \theta_i^k \sigma_i^k + \sum_{(k,l)\in\mathscr{P}_2}\sum_{(i,j)\in\mathscr{E}} \theta_{i,j}^{k,l}\sigma_i^k\sigma_j^l$

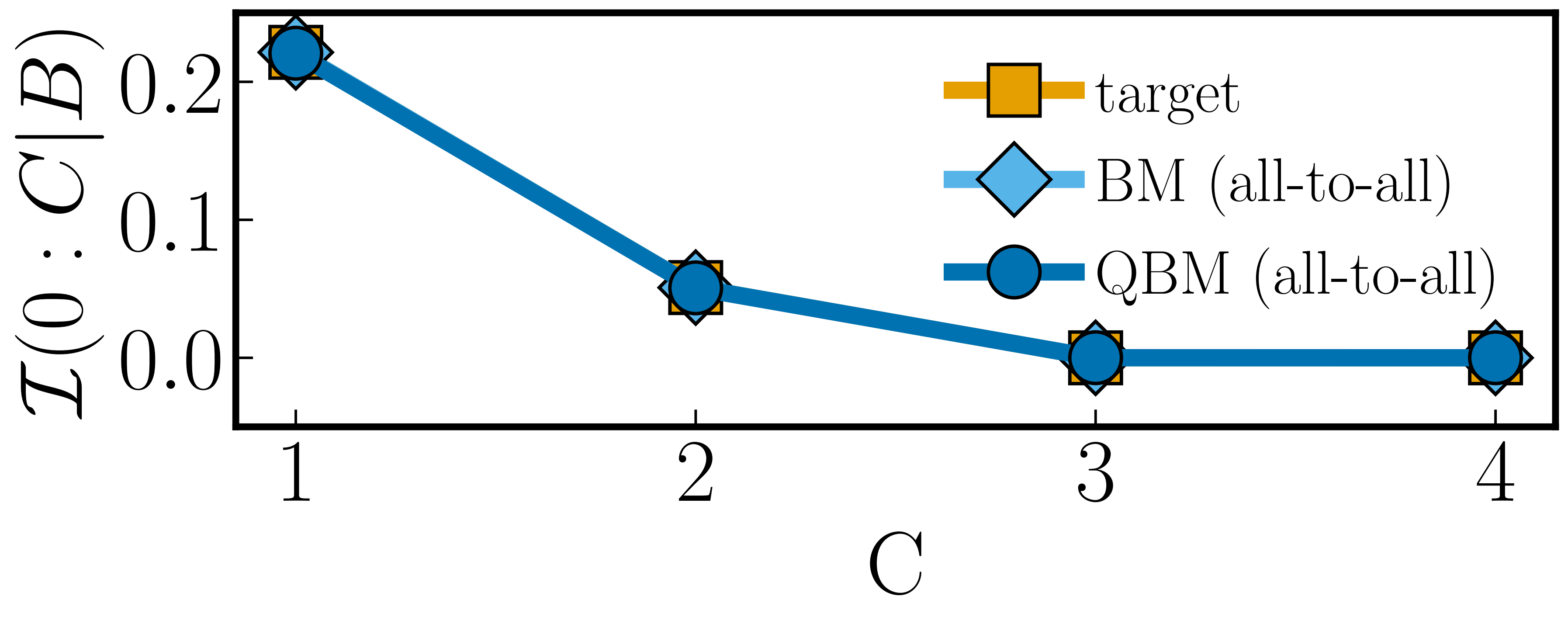
σ_i^k denotes the Pauli matrix applied on the i-th qubit with $k \in \mathscr{W} = \{X, Y, Z\}$ and $\mathscr{P}_1 \subseteq \mathscr{W}, \mathscr{P}_2 \subseteq \mathscr{W} \otimes \mathscr{W}$

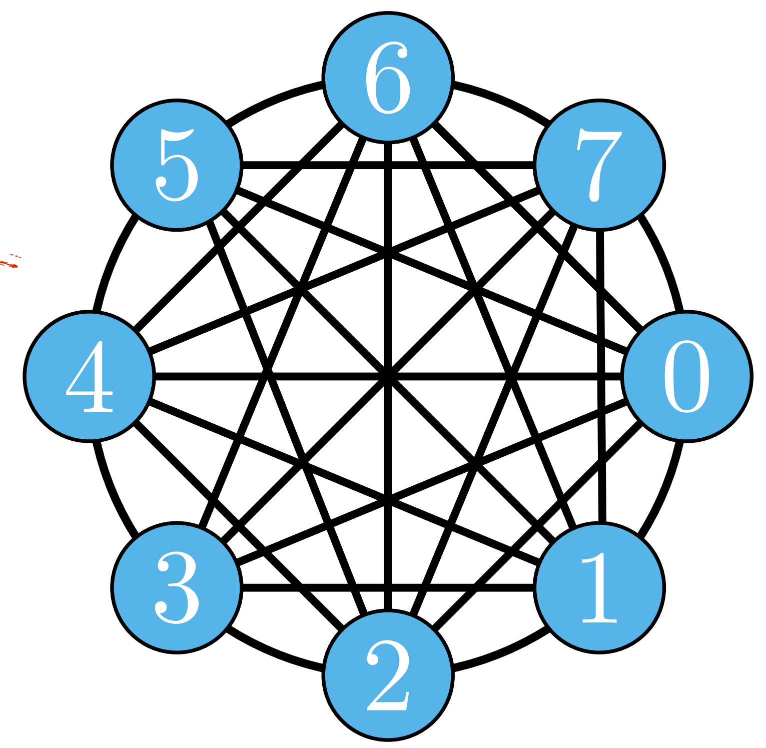
CONDITIONAL MUTUAL INFORMATION

- Models trained on next-nearest-neighbor (NNN) distribution.

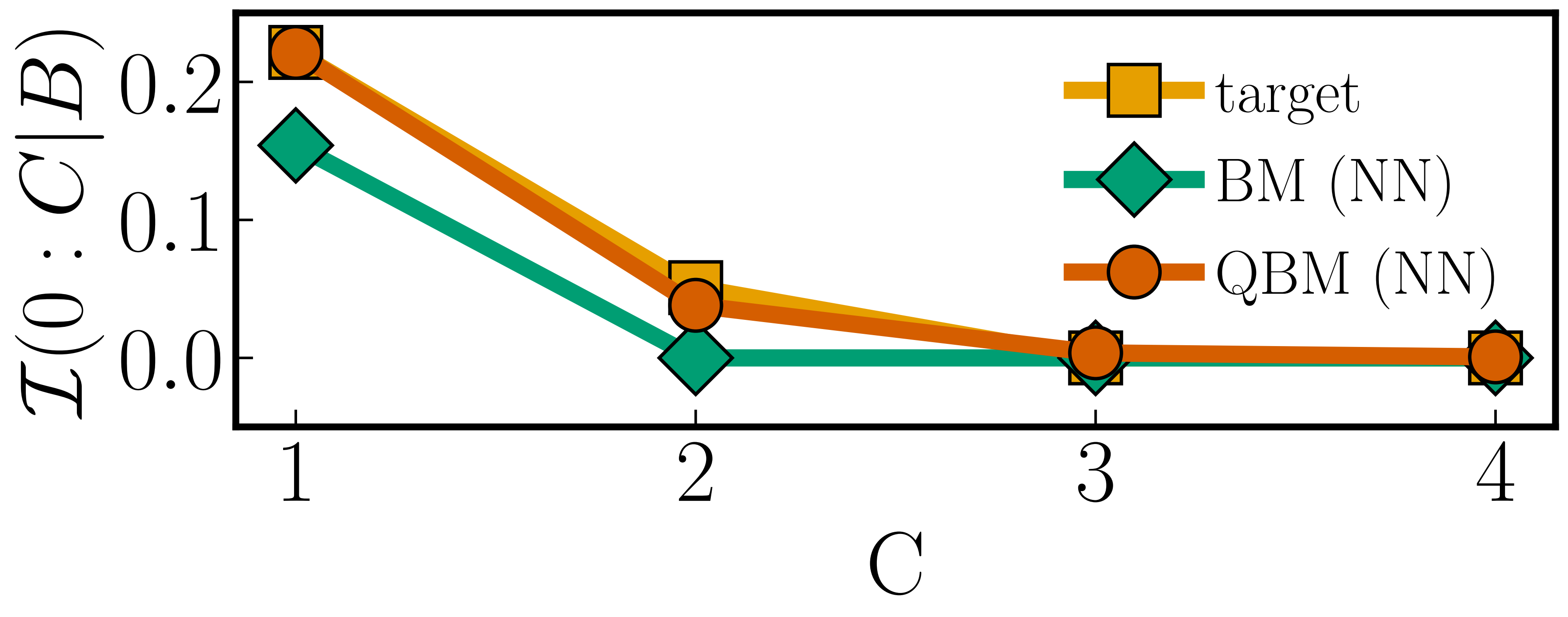


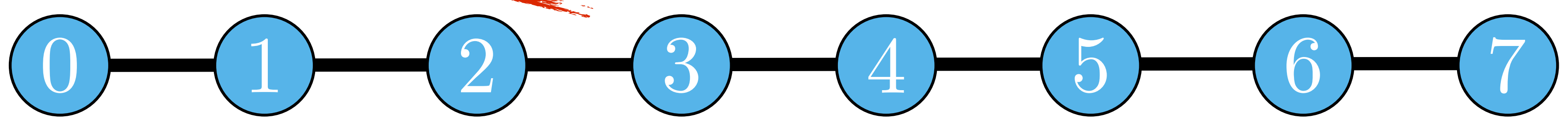
target connectivity





all-to-all connectivity

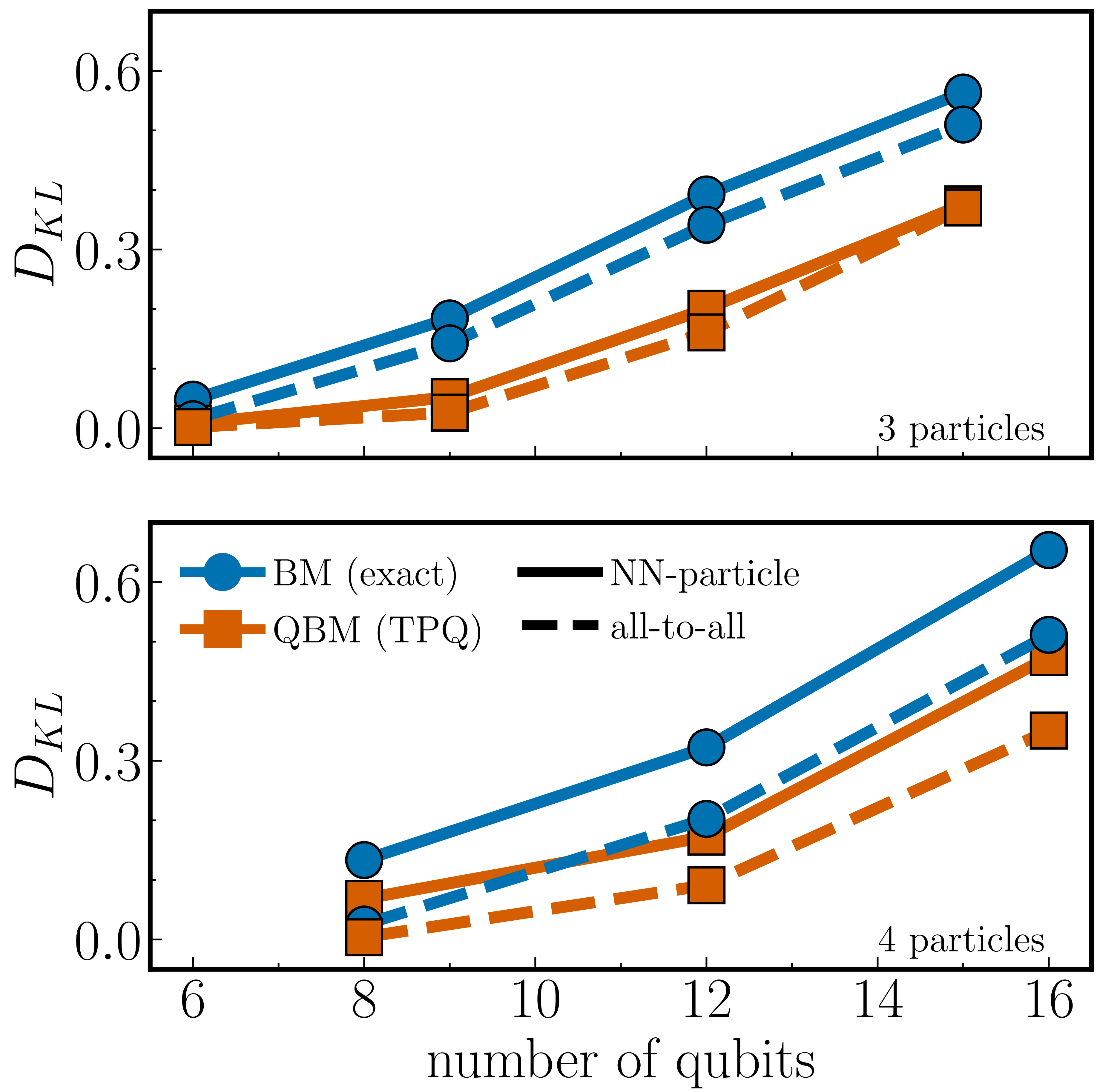


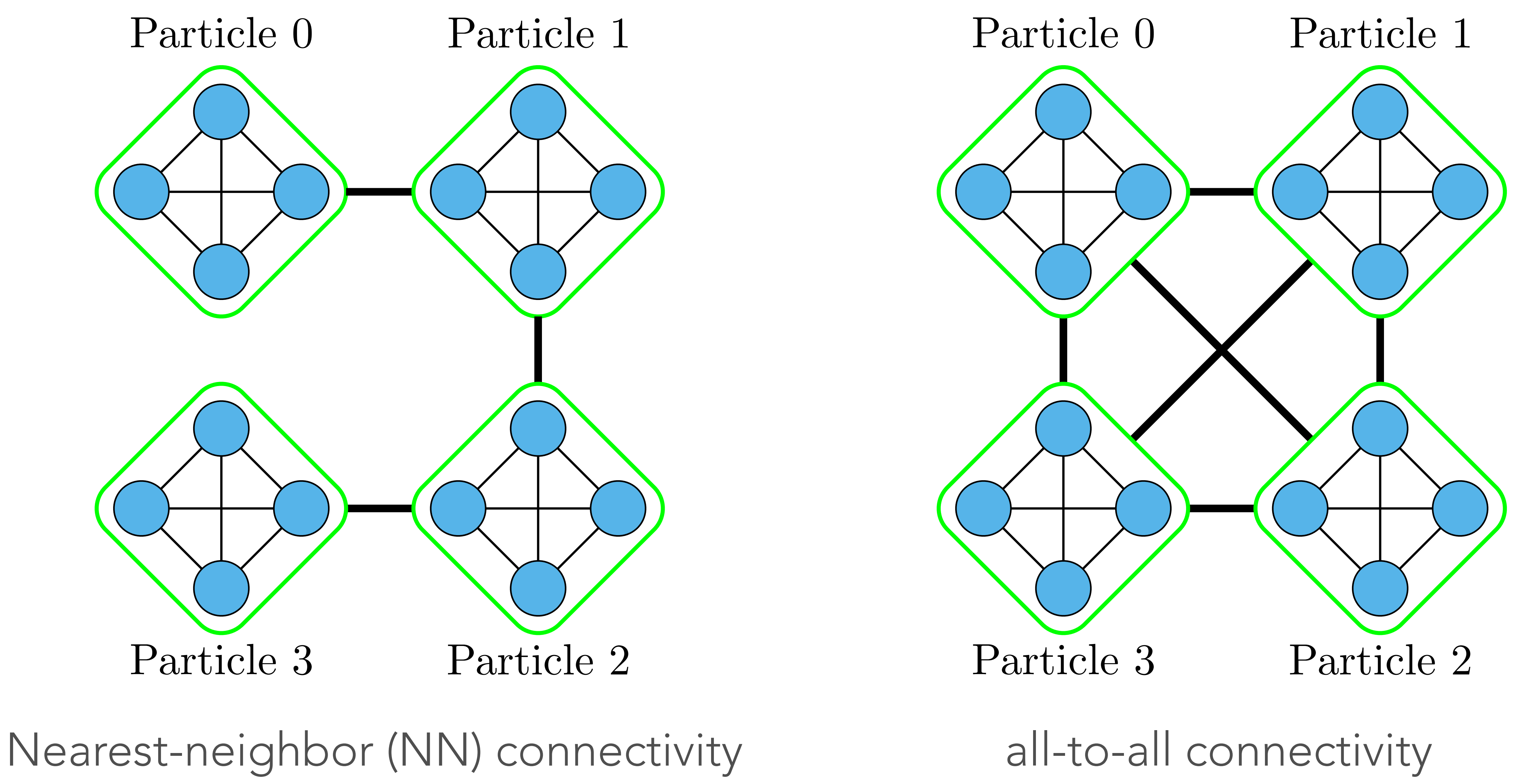


Nearest-neighbor connectivity

- QBMs better capture Conditional Mutual Information compared to classical BMs
- Significant improvement in modeling higher-order correlations.
- Demonstrates QBMs' enhanced expressivity even with limited connectivity.

PARTICLE JET EVENT GENERATION

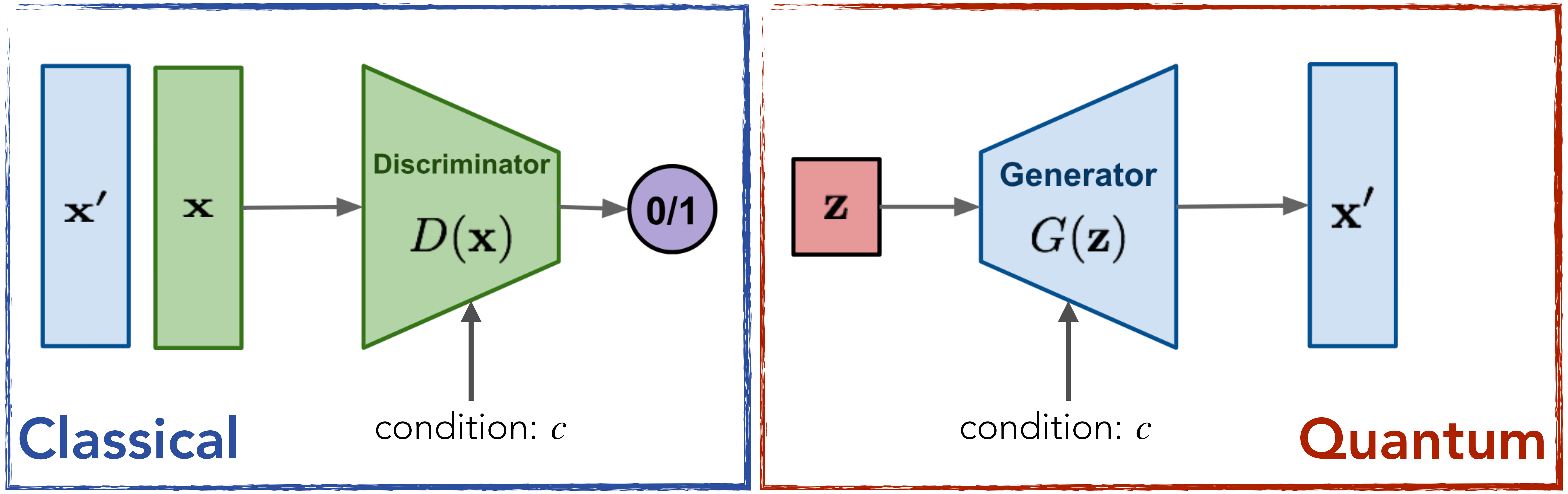




- Difficulty in accurately learning particle correlations for the Classical BM
- Significant improvement in modeling higher-order correlations.
- **Demonstrates QBMs' enhanced expressivity even with limited connectivity.**

QUANTUM GENERATIVE ADVERSARIAL NETWORKS

- **Generative Adversarial Networks :** two networks competing, generator produces fake data, and a discriminator distinguishes between real and fake data
- **Quantum GAN (QGAN) replaces the generator network by a parameterised quantum circuit**

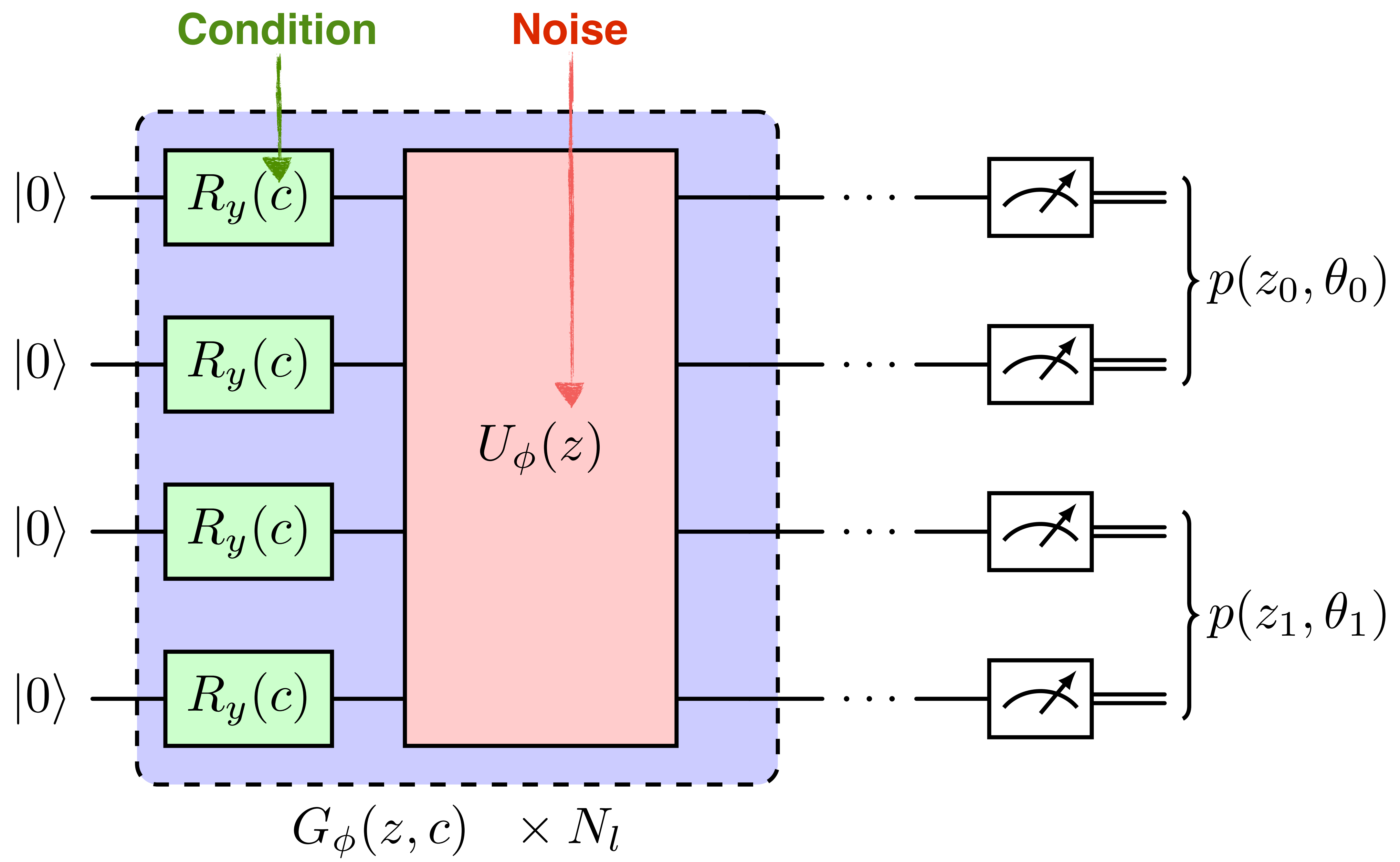


- **Can we use quantum technologies to create an enhanced GAN?**
- **Can a QGAN trained on jet substructure give insights into the correlation structure in a way that a classical GAN cannot?**
- **Is it possible to associate the complexity of the shower with the complexity of the circuit**
- **Would a QGAN have less parameters?**

Modified from Lilian Weng: https://lilianweng.github.io/posts/2018-10-13-flow-models/

QUANTUM GENERATOR DESIGN

Quantum Generator model ← A parametrised quantum cricuit $U_\phi(\mathbf{z})$



- Data: anti-k_T ($R = 0.4$) jets generated with Pythia8 ($p_T > 30$ GeV)
- Each jet constituent represented by two features:
  - Momentum fraction $z_i = p_T^i / p_T^{jet}$
  - Angle with reference to the jet axis $\theta_i = \Delta R_i / R$
- 1 qubit = 1 feature: $\hat{x} = \{\langle\sigma_Z^0\rangle, \langle\sigma_Z^1\rangle, \ldots, \langle\sigma_Z^n\rangle\}$
- Style-based approach (*): the noise is inserted in every gate: $\phi_i \rightarrow \phi_i(z, w, b) = w \cdot z + b$
  - The noise z, is sampled from $\mathcal{U}(0,1)$
- Jet p_T included as condition via R_y rotations in styled-base approach: $c \rightarrow c = w \cdot c + b$

Training: estimates the earth mover distance between generated & real data: Wasserstein GAN(**)

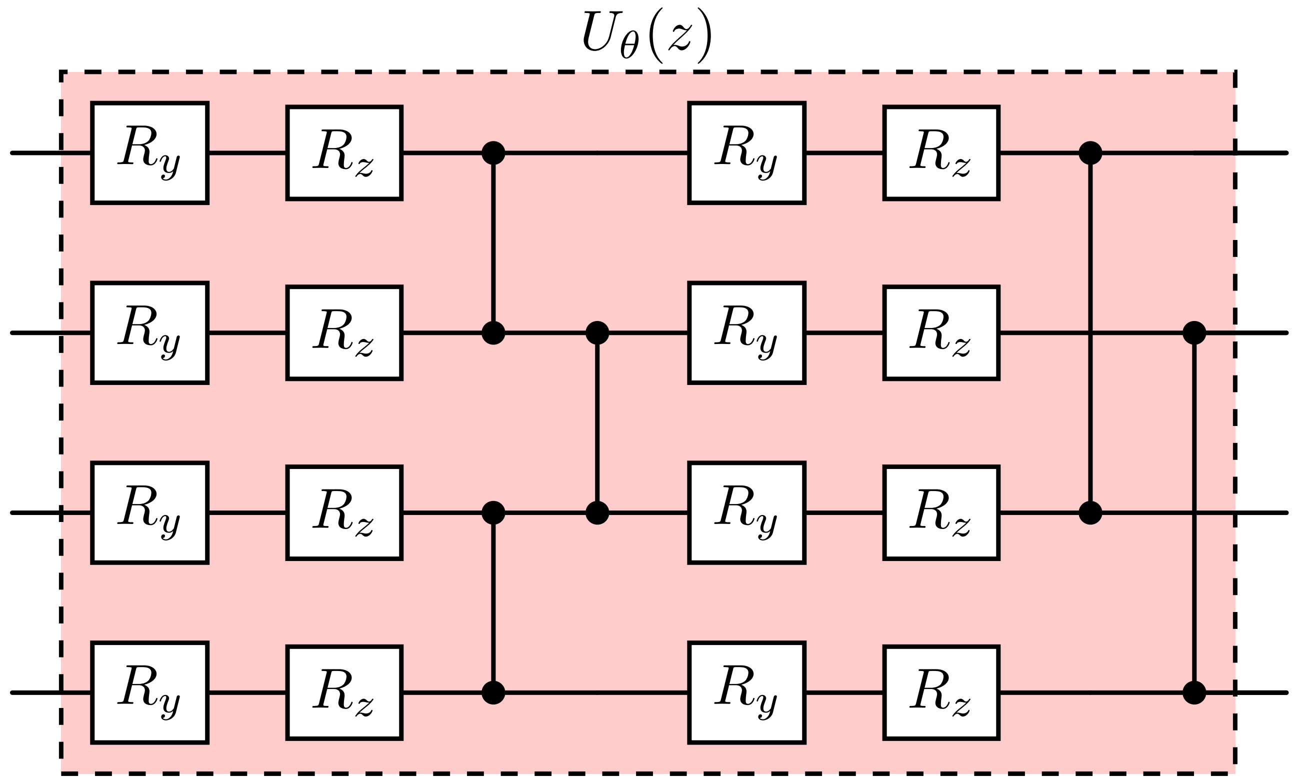
$$L = \underset{\tilde{\boldsymbol{x}} \sim \mathbb{P}_g}{\mathbb{E}}\left[D(\tilde{\boldsymbol{x}})\right] - \underset{\boldsymbol{x} \sim \mathbb{P}_r}{\mathbb{E}}\left[D(\boldsymbol{x})\right] + \lambda \underset{\hat{\boldsymbol{x}} \sim \mathbb{P}_{\hat{\boldsymbol{x}}}}{\mathbb{E}}\left[\left(\|\nabla_{\hat{\boldsymbol{x}}} D(\hat{\boldsymbol{x}})\|_2 - 1\right)^2\right].$$

(*) Carlos Bravo-Prieto et al. Quantum 6, 777 (2022).
Su Yeon Chang et al. arXiv:2406.02668

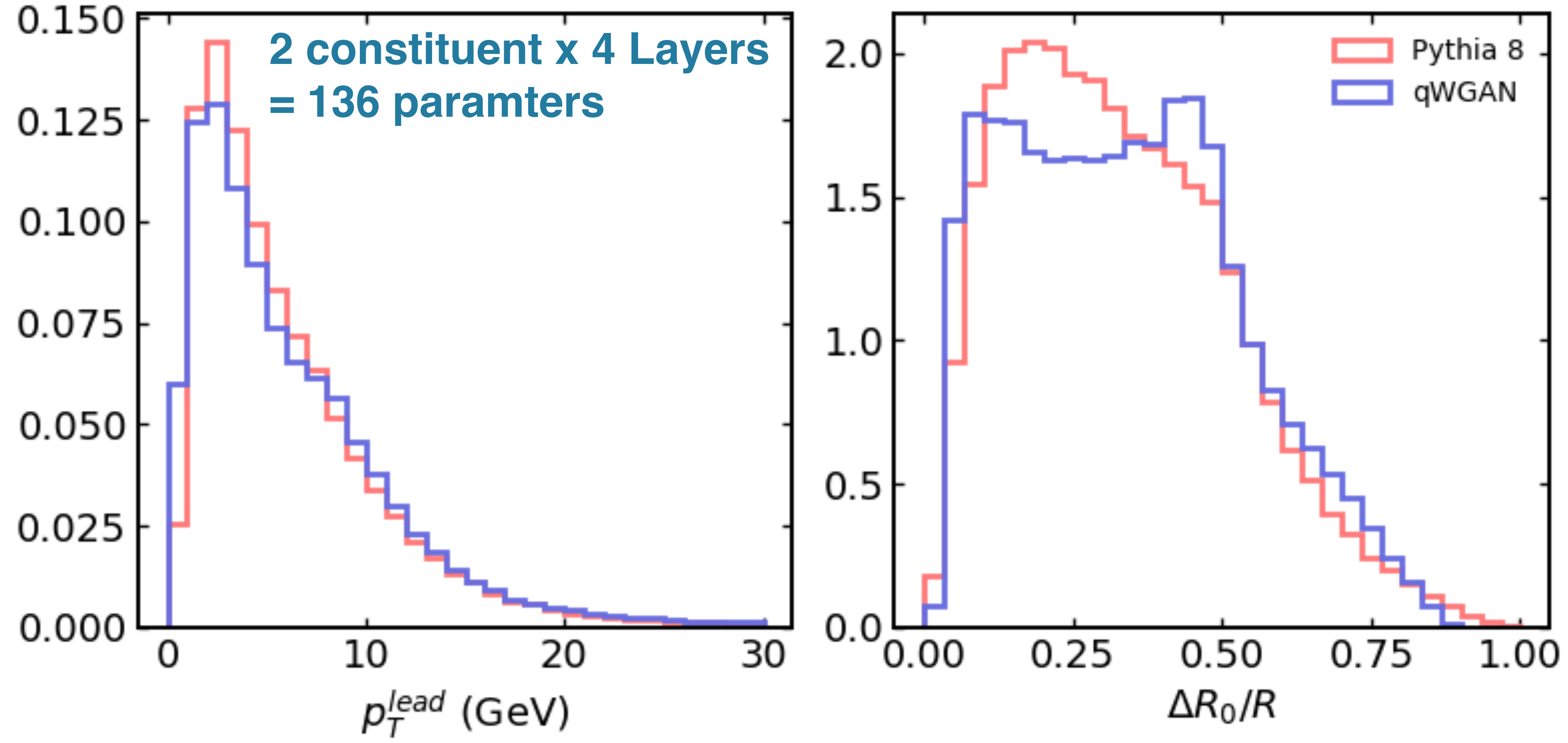
(**) Ishaan Gulrajani et al. arXiv:1704.00028v3

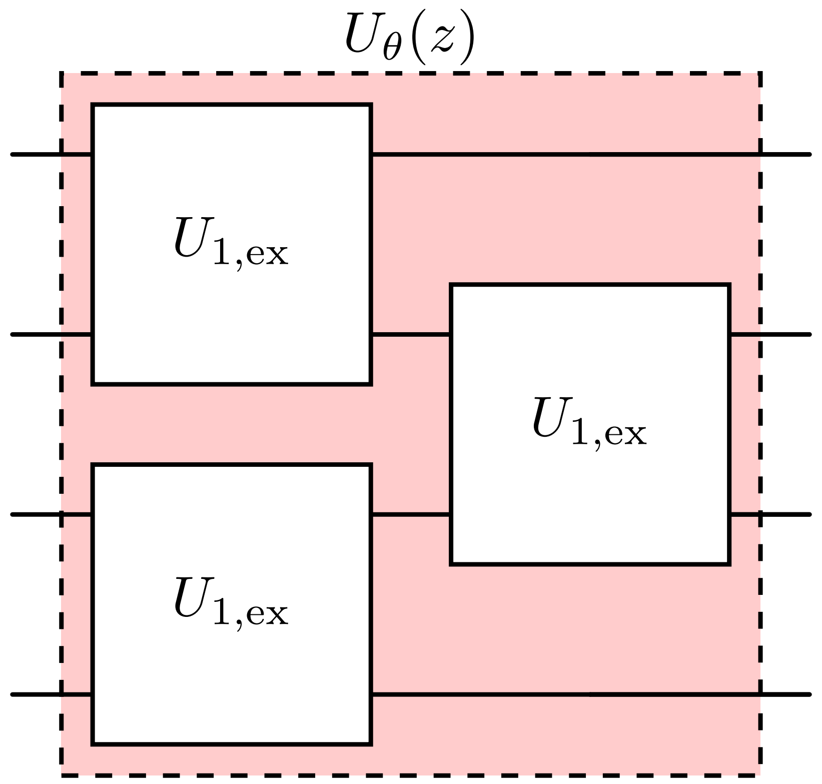
SIMULATION WITH JET DATA

we choose two main anzats for $U_\theta(z)$



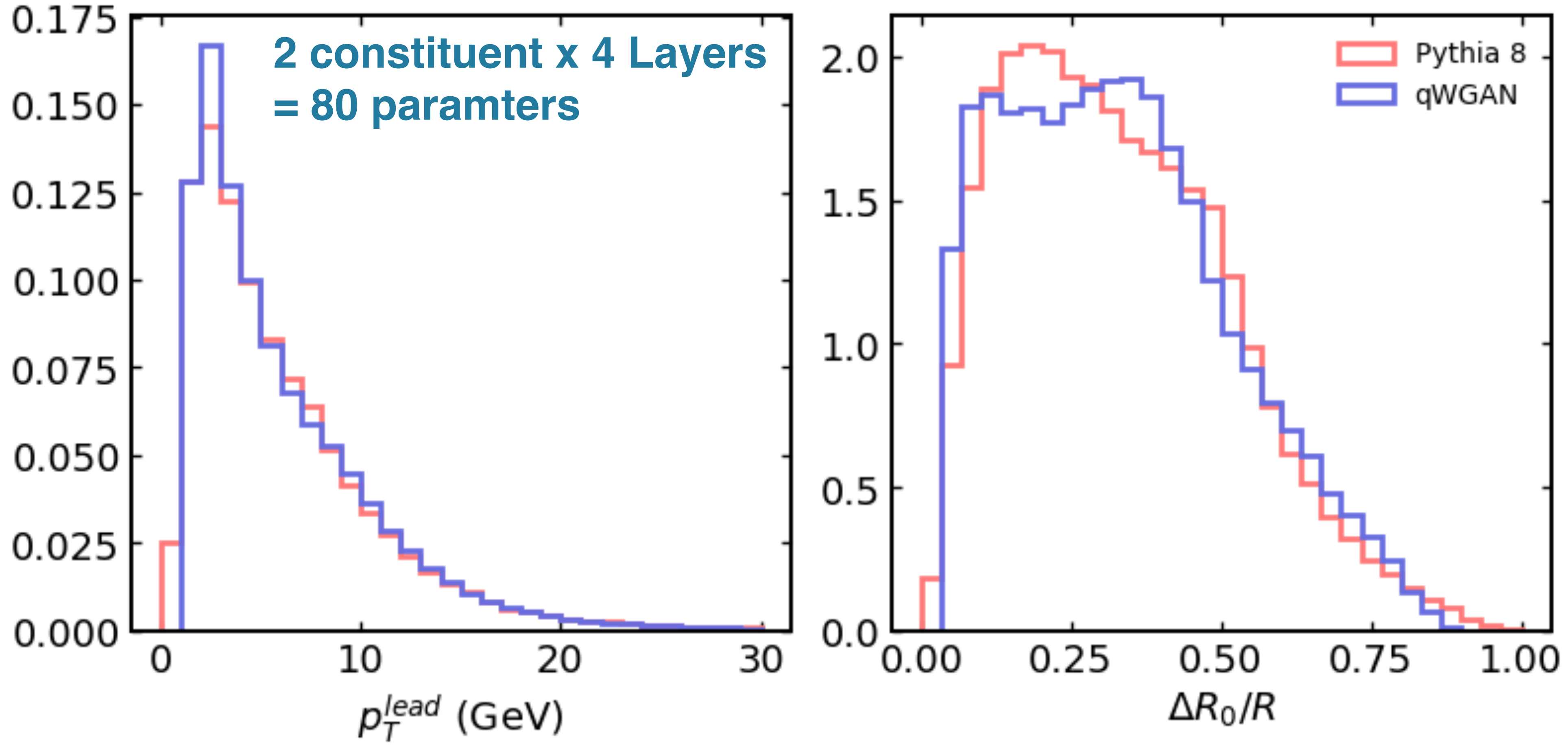
Circuit-centric design inspired from M. Shuld et al arXiv:1804.00633





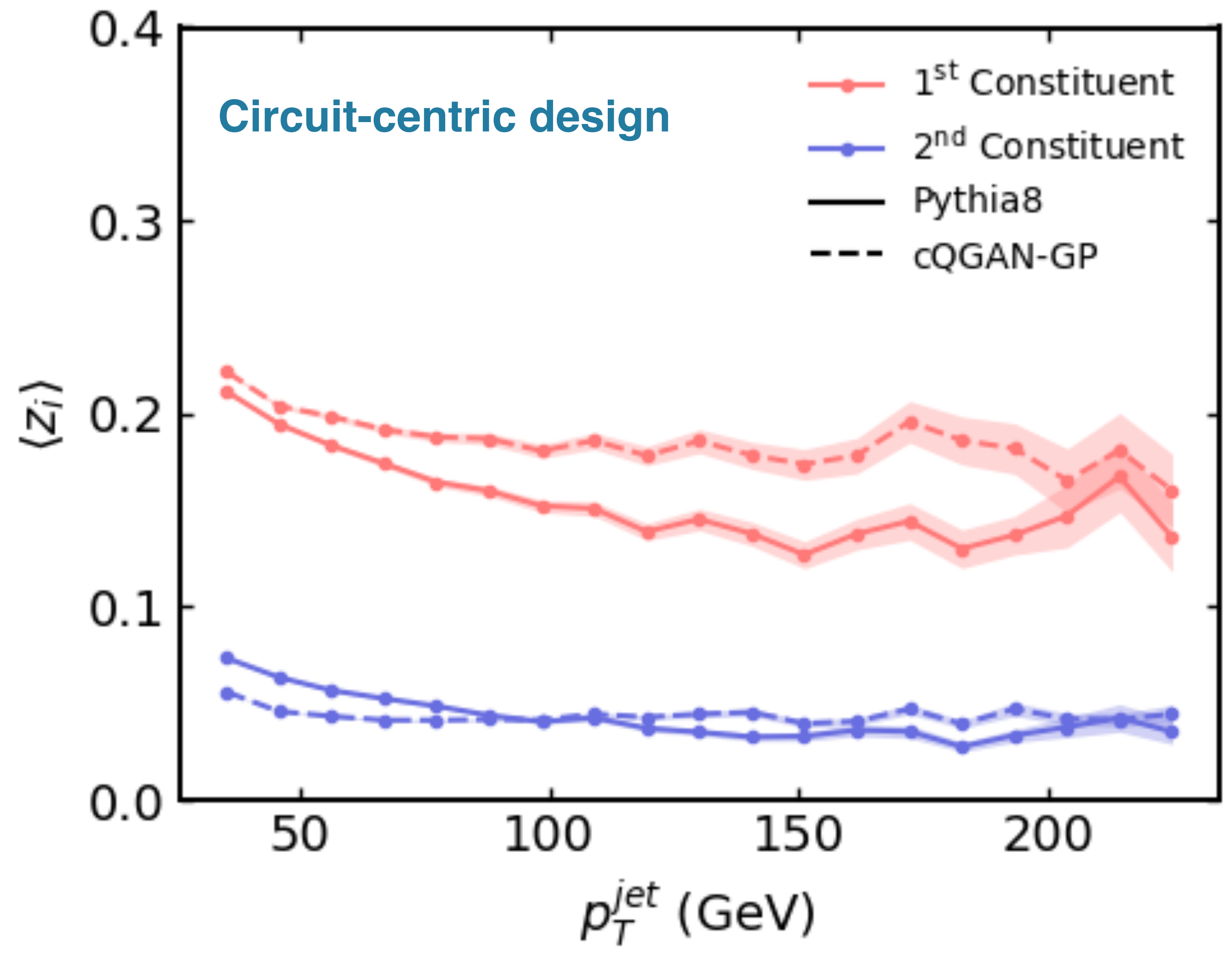
$$U_{1,\mathrm{ex}}(\phi,\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & e^{i\phi}\sin(\theta) & 0 \\ 0 & e^{-i\phi}\sin(\theta) & -\cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

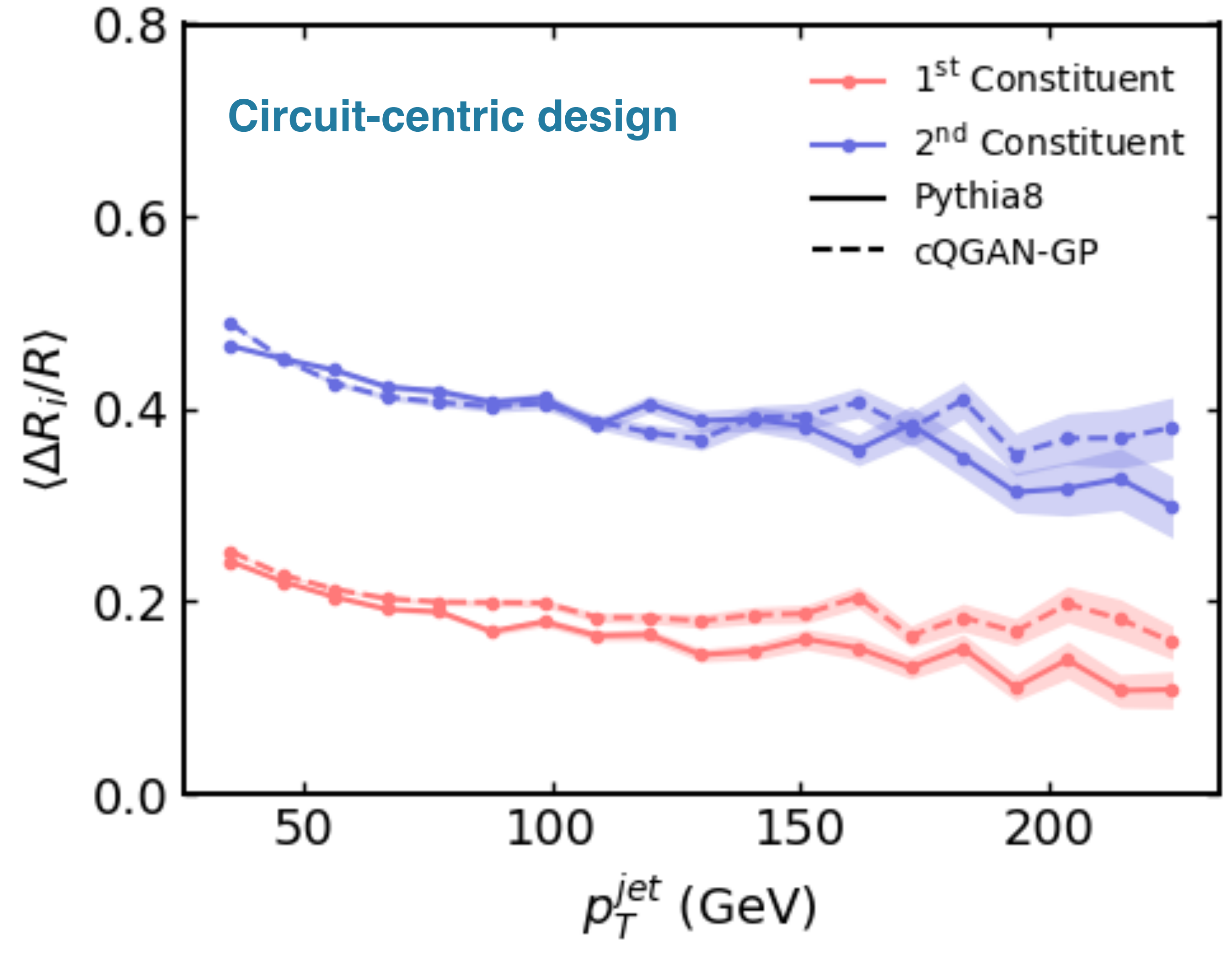
Brick-Wall design with Particle-conserving gate $U_{1,\mathrm{ex}}$ proposed by Barkoutsos et al. in arXiv:1805.04340

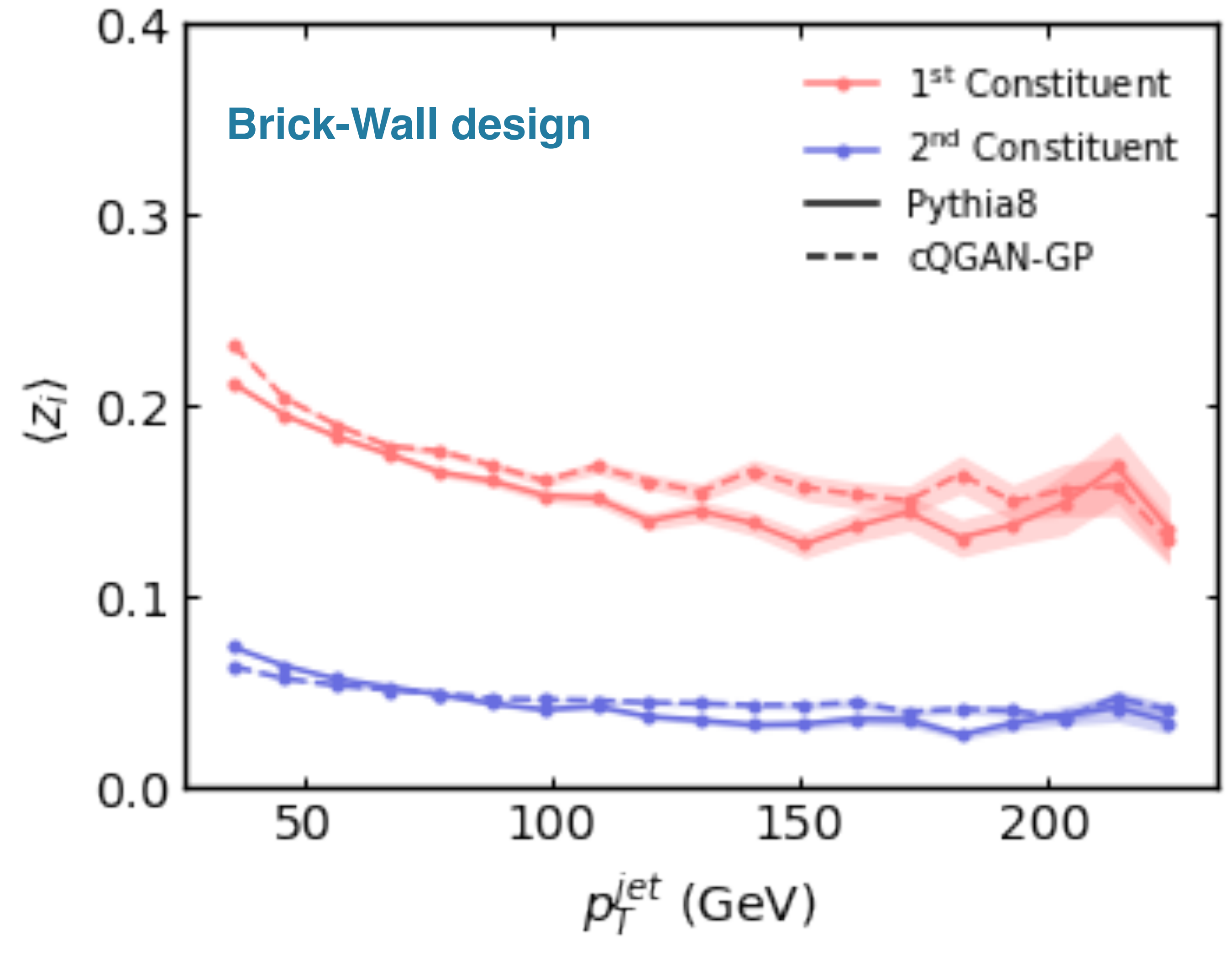


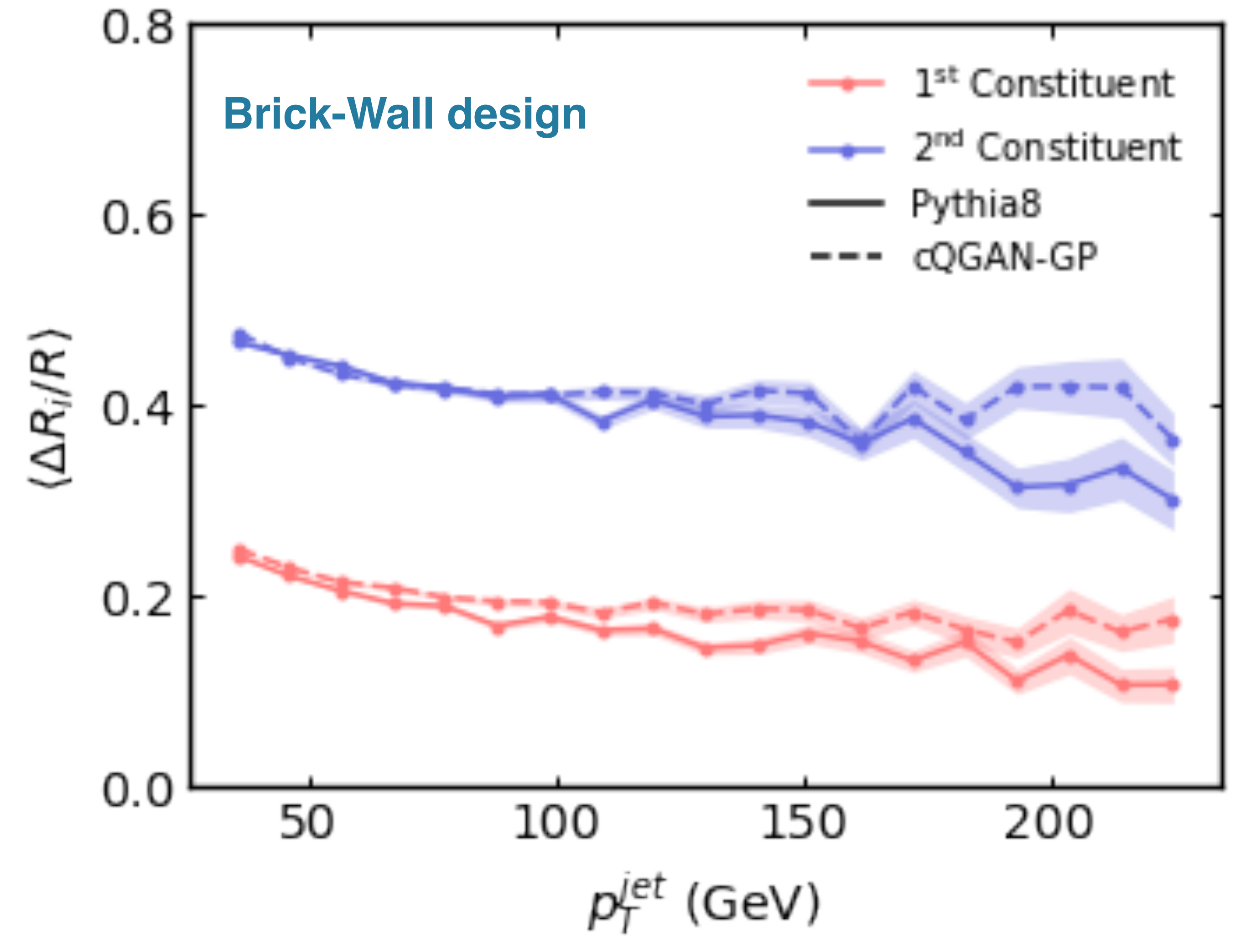
Simulation perfomed on a classical computer with a quantum simulator in Pennylane and PyTorch frameworks

SIMULATION WITH JET DATA



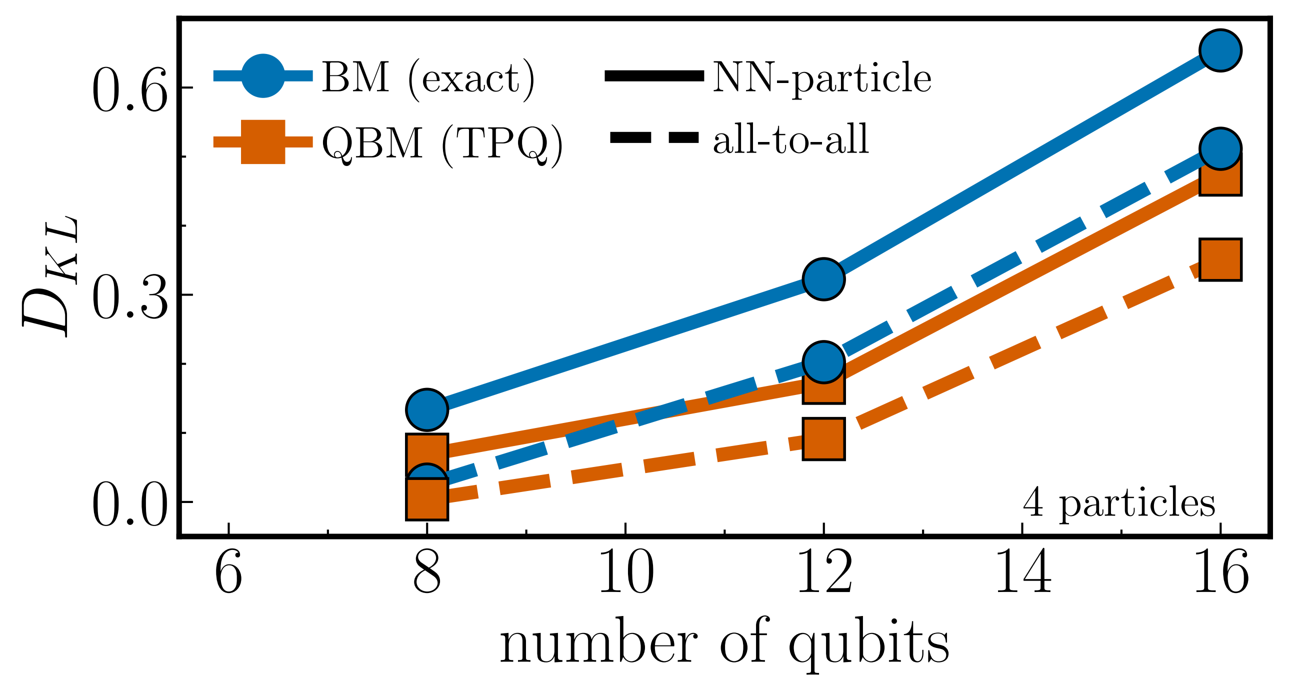


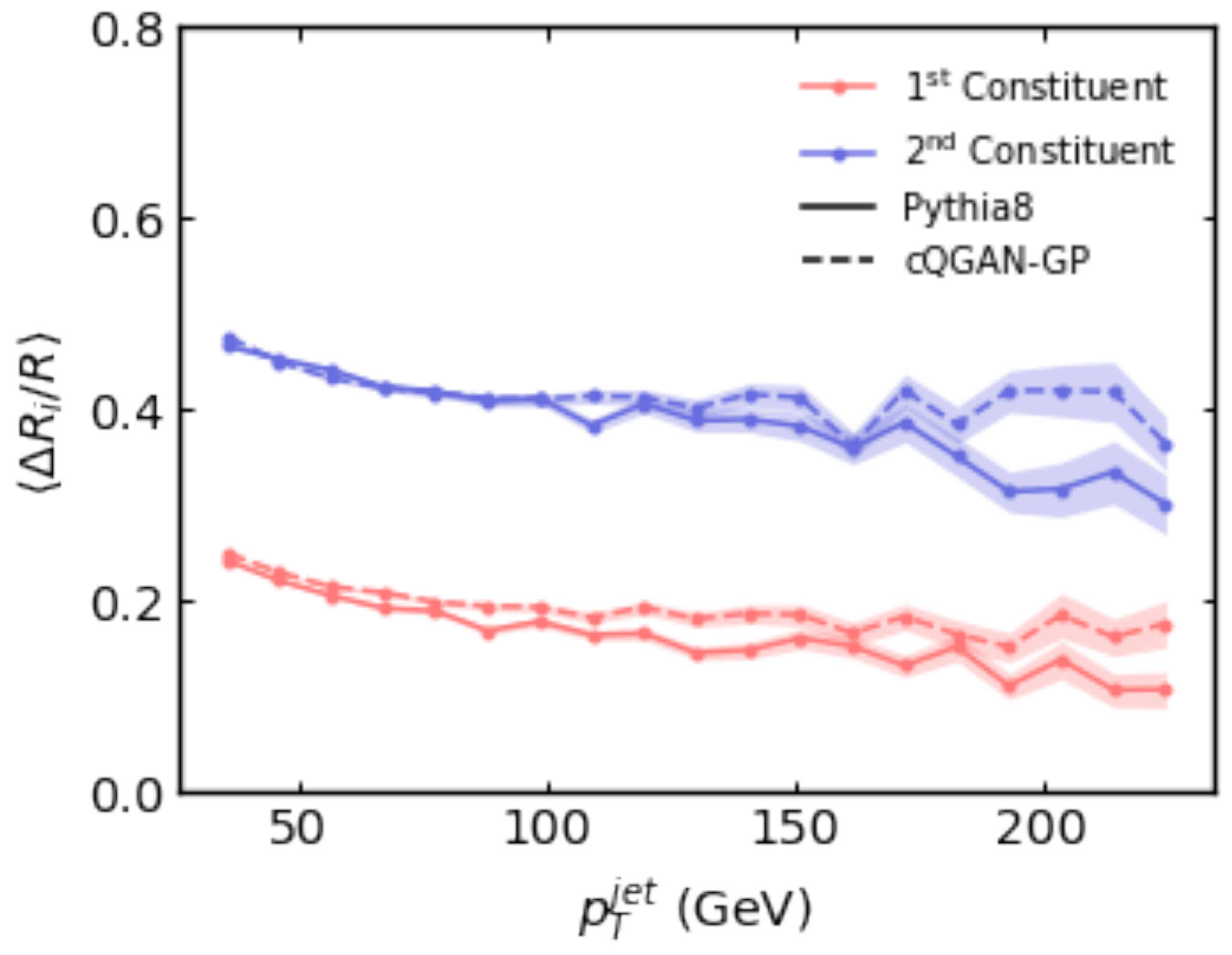




SUMMARY

- **QML Growth in HEP**: QML is rapidly advancing in high-energy physics, creating useful prototype algorithms.
- **Potential of Quantum Generative Models**: QBMs and QGANs hold promise for accurate jet substructure modeling.
- **Better High-Dimensional Modeling**: Quantum models capture complex distributions and correlations beyond classical capabilities.
- **Need for theory-centric approach:** More studies are needed to link quantum model structure with performance.
- **QML is not a magic bullet:** Need quantum centric problems to leverage quantum learners

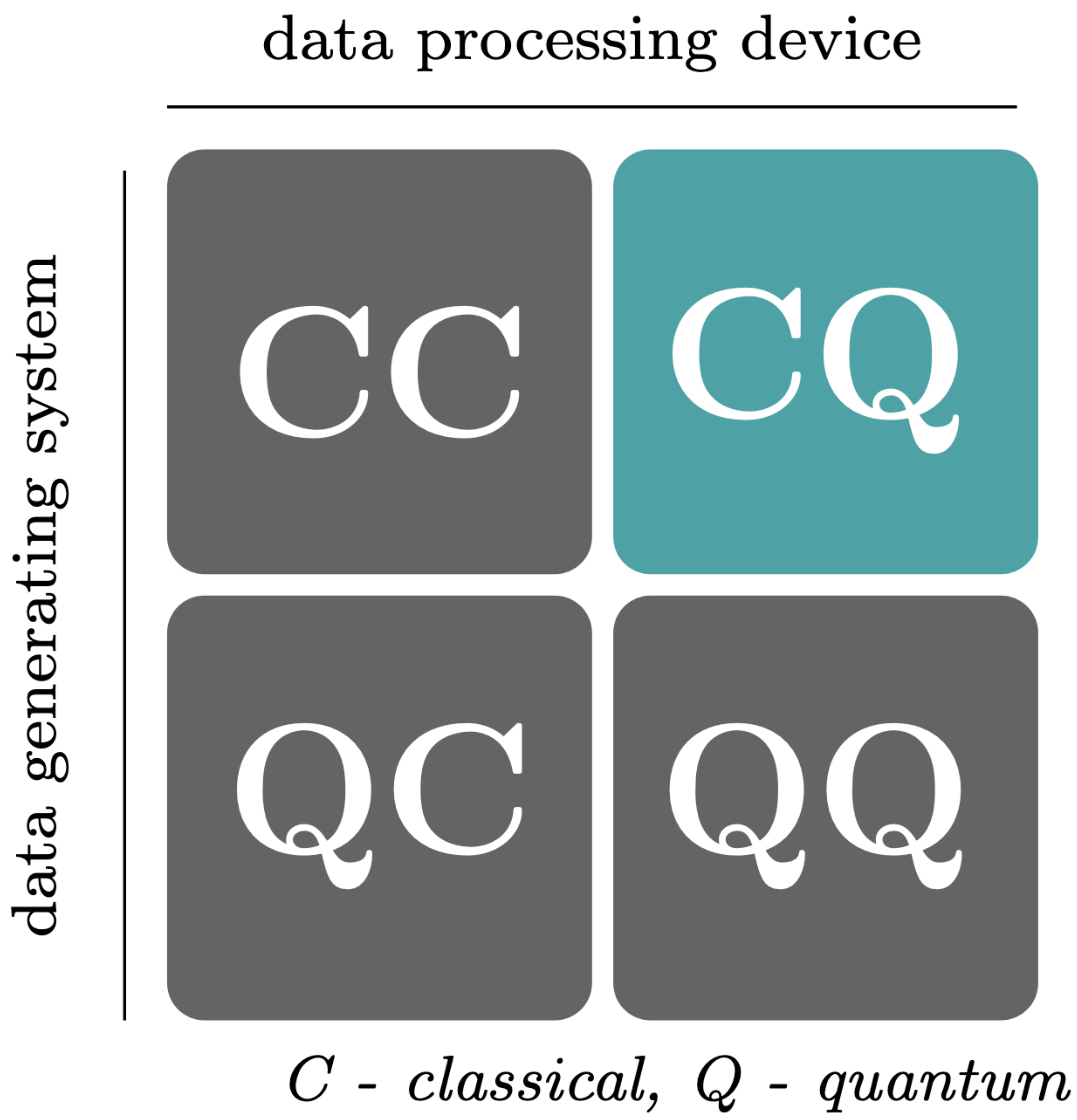




SUMMARY

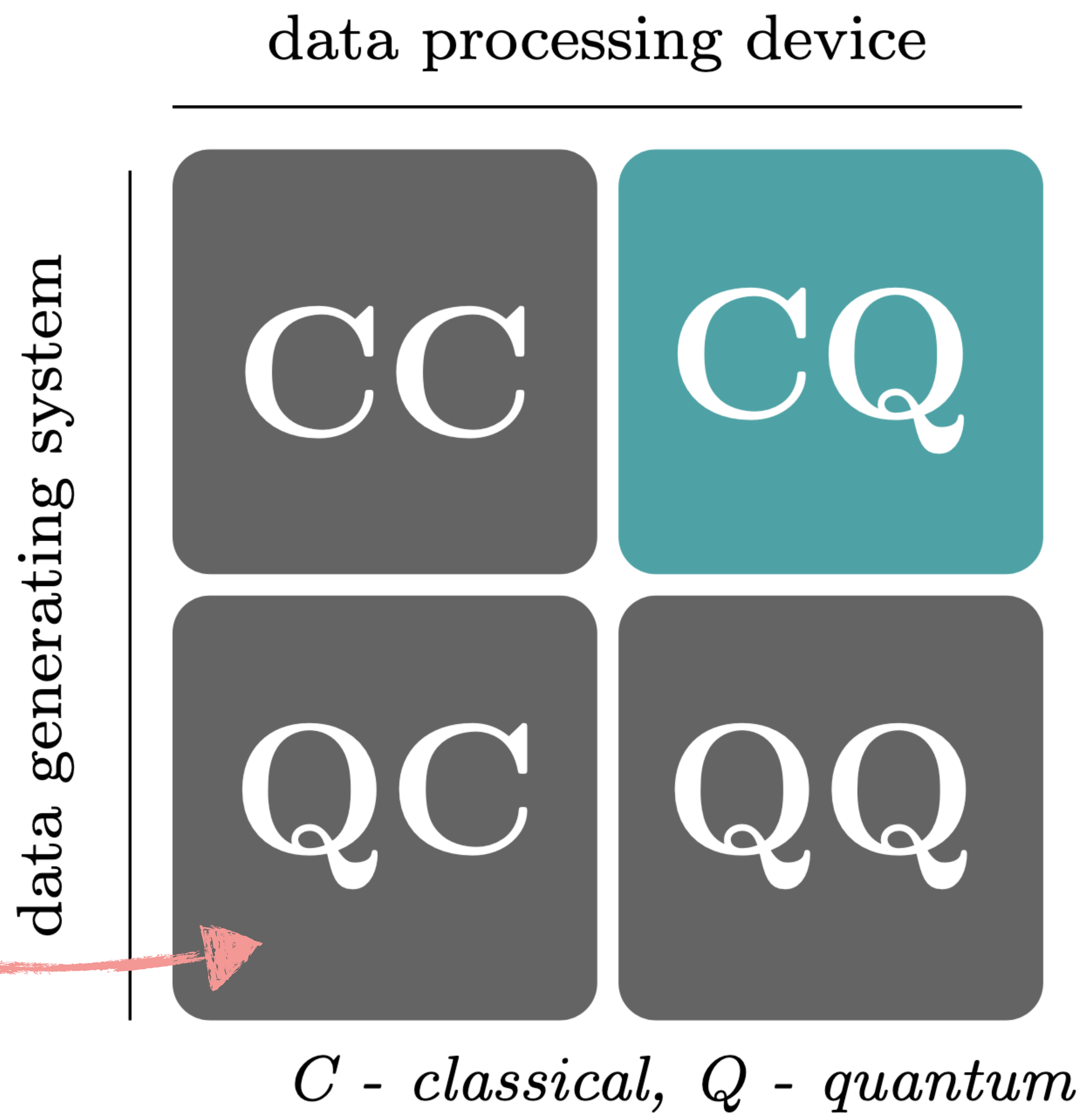
BACKUP

FIELDS IN QUANTUM COMPUTING



Schuld, M., Petruccione, F. (2021)

FIELDS IN QUANTUM COMPUTING

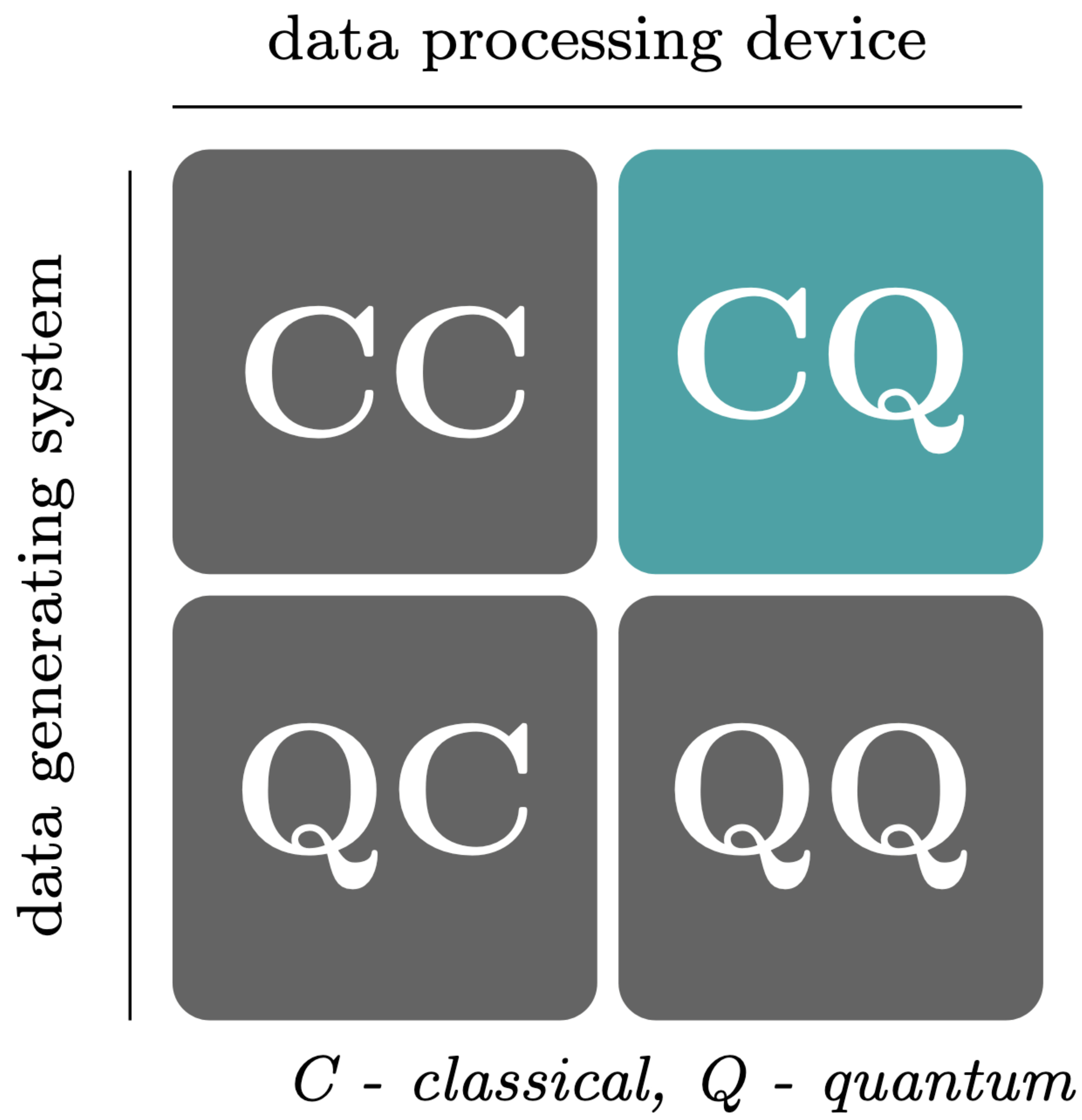


Simulation of Quantum Systems using classical ML

Simulation of Quantum Systems using a Quantum Computer

Schuld, M., Petruccione, F. (2021)

FIELDS IN QUANTUM COMPUTING

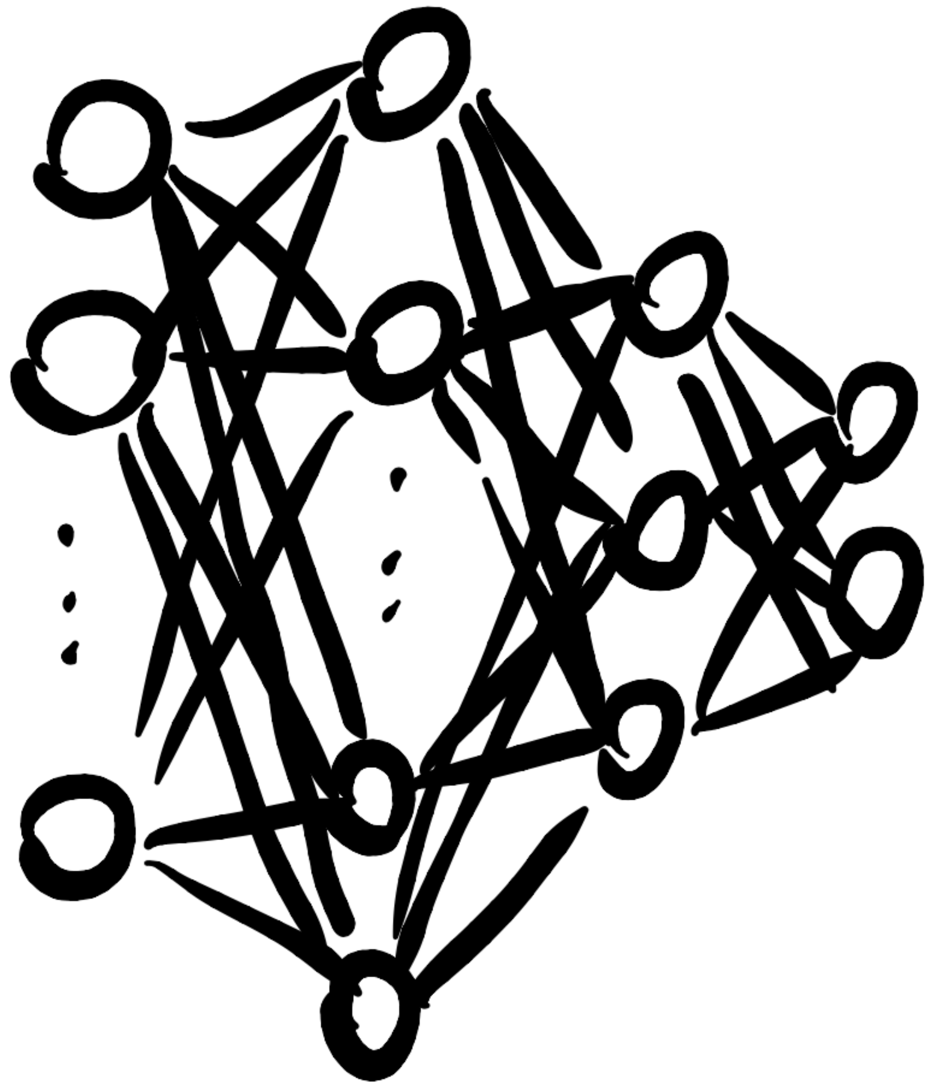


Here: **focus on quantum algorithm with classical input data**

Schuld, M., Petruccione, F. (2021)

SUPERVISED LEARNING IN QUANTUM COMPUTING: QUANTUM CLASSIFIERS

Goal: learn input-output relation of labeled data

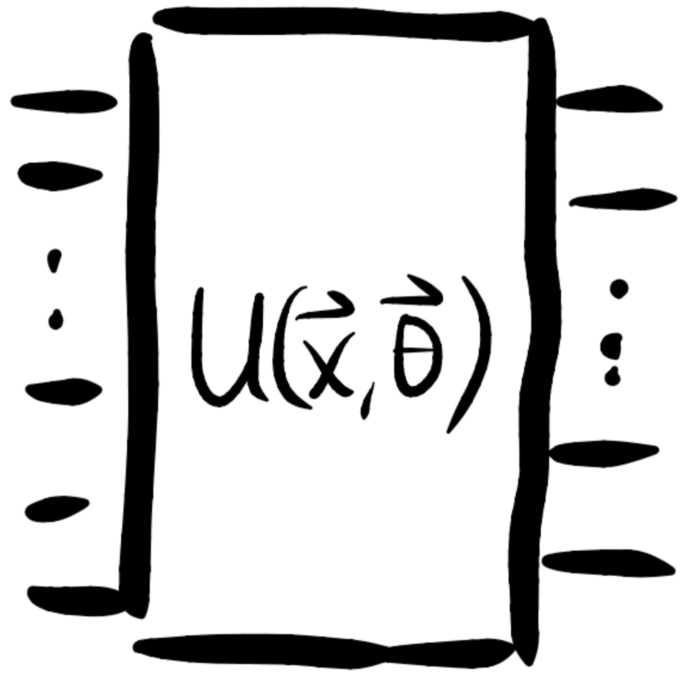


$$\Psi(\vec{x},\vec{\theta}): \mathbb{R}^{n_{in}} \to \mathbb{R}^{n_{out}}$$

input trainable weights

Classical Neural Network

$|0\rangle^{\otimes n_{in}}$



$$|\Psi_{out}(\vec{x},\vec{\theta})\rangle = U(\vec{x},\vec{\theta})|0\rangle^{\otimes n_{in}}$$

input trainable weights

$$y(\vec{x},\vec{\theta}) = \langle \Psi_{out}(\vec{x},\vec{\theta})|\,\hat{O}\,|\Psi_{out}(\vec{x},\vec{\theta})\rangle$$

Parametrized Quantum Circuit

courtesy Michele Grossi

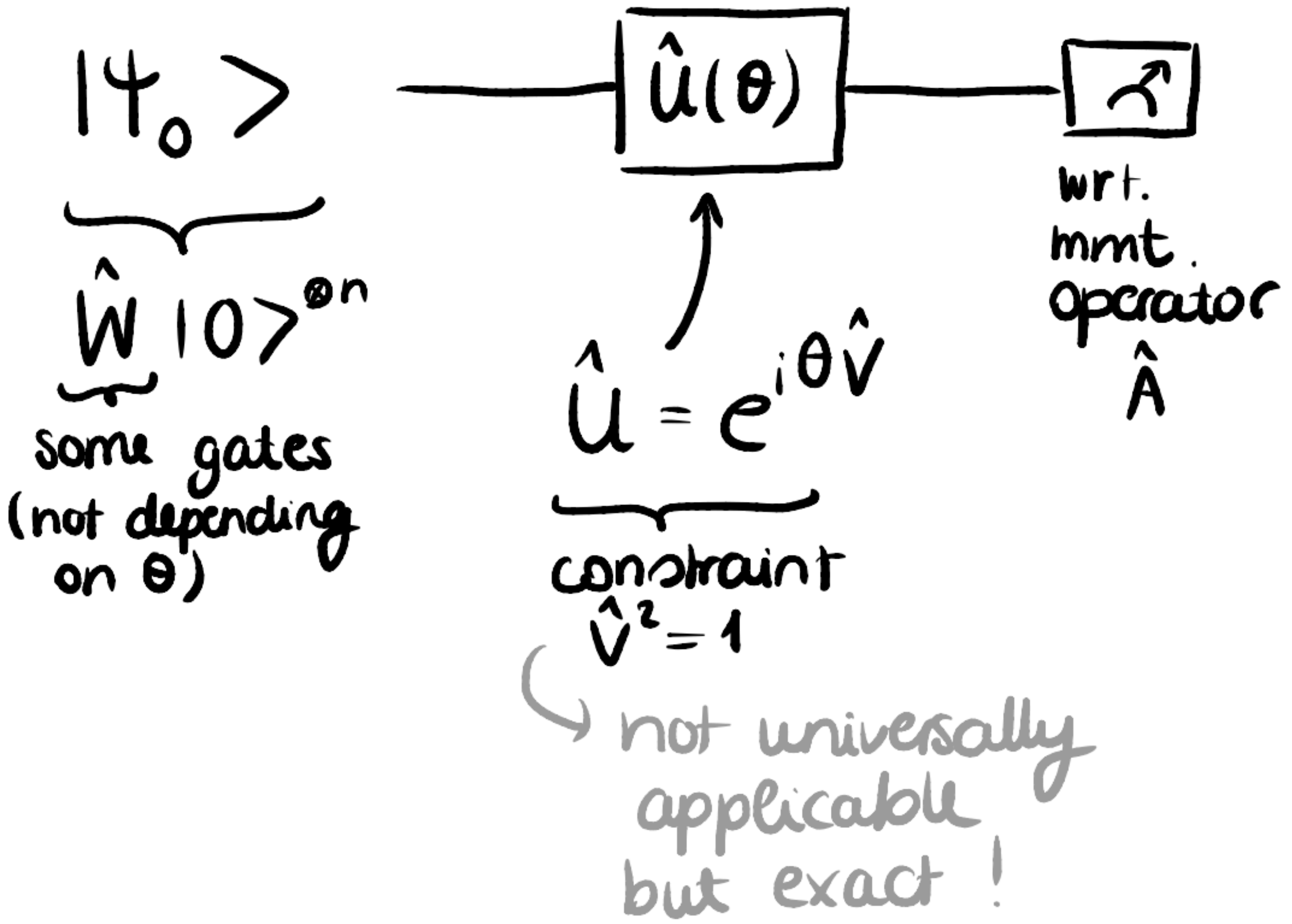
PARAMETER OPTIMIZATION

$$\theta \rightarrow \theta - \eta \nabla_\theta f$$

$$\uparrow \langle \hat{A}(\theta) \rangle$$

The parameter-shift rule (gradient-based)

⟶ Compute **partial derivative** of variational circuit parameter θ, alternative to analytical gradient computation and classical finite difference rule (numerical errors and resource cost considerations)



$$\Rightarrow \nabla_\theta \langle \hat{A} \rangle = u\left[\langle \hat{A}(\theta + \tfrac{\pi}{4u}) \rangle - \langle \hat{A}(\theta - \tfrac{\pi}{4u}) \rangle\right]$$

⟶ Evaluate Quantum Circuit twice at shifted parameters to compute gradient

Source:https://pennylane.ai/qml/demos/tutorial_stochastic_parameter_shift/

courtesy Michele Grossi

$$\theta_{k+1} \leftarrow \theta_k - a_k \underbrace{\hat{g}_k(\hat{\theta}_k)}_{\text{stochastic estimate of } \nabla_\theta f}$$

Simultaneous Perturbation Stochastic Approximation (SPSA)
(gradient-free)

⟶ If gradient computation not possible, too resource-intensive, or noise-robustness required (slower convergence but fewer function evaluations)

⟶ Gradient is approximated by two sampling steps and parameters are perturbed in all directions simultaneously

" $y(\theta) = f(\theta) + \varepsilon$ "

random output perturbation

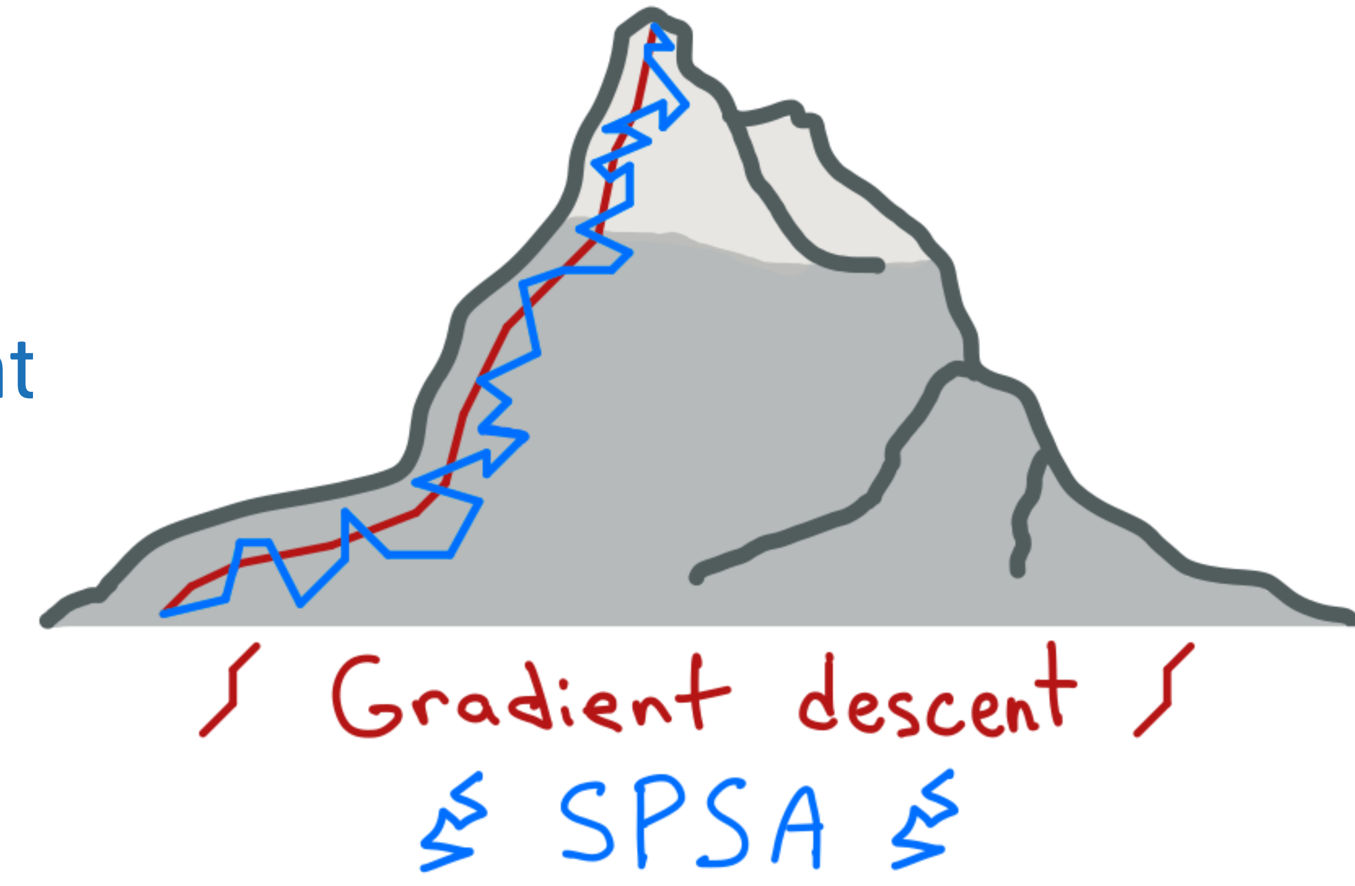
$$\hat{g}_{ki}(\hat{\theta}_k) = \frac{y(\hat{\theta}_k + c_k\Delta_k) - y(\hat{\theta}_k - c_k\Delta_k)}{2c_k\Delta_{ki}}$$

$c_k \geq 0$, $\Delta_k = (\Delta_{k_1}, \Delta_{k_2}, \ldots, \Delta_{kp})^T$ perturbation vector

(~ randomly sampled from zero-mean distr.)

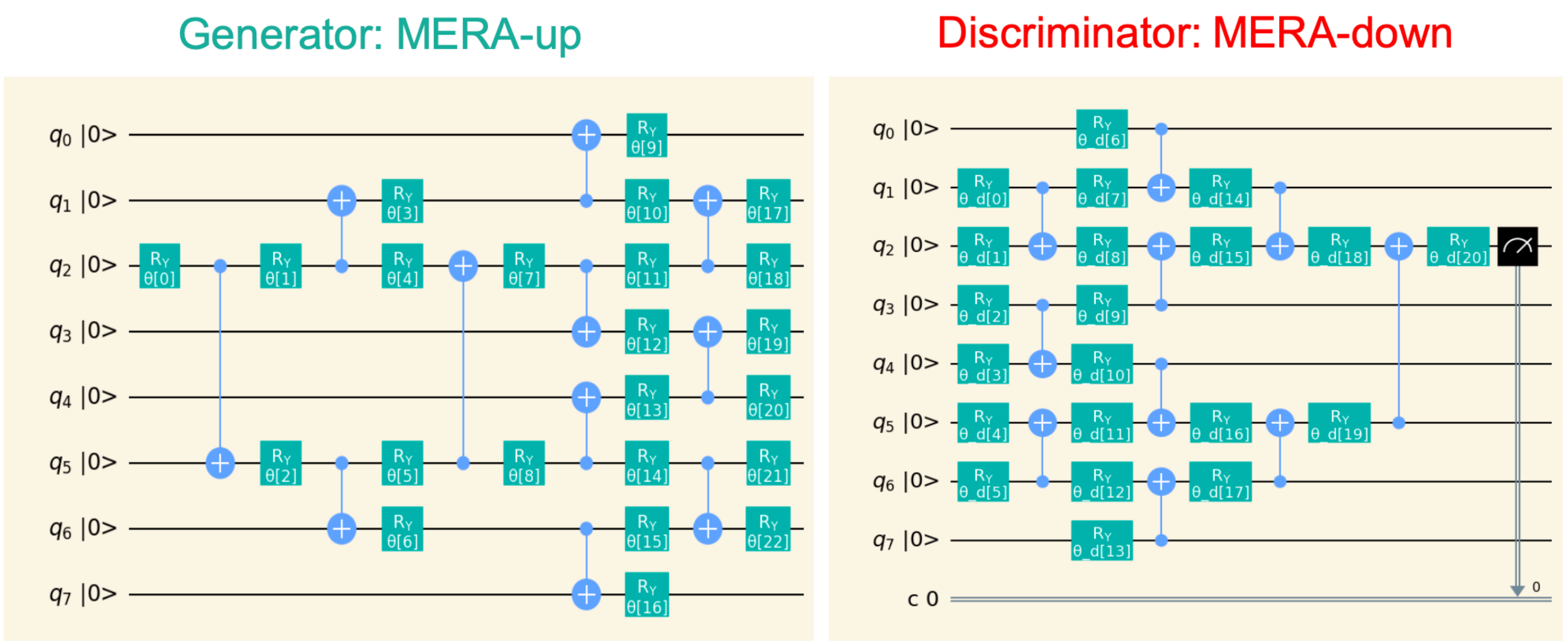
Iterative update rule comparable to classical stochastic gradient descent

https://pennylane.ai/qml/demos/tutorial_spsa

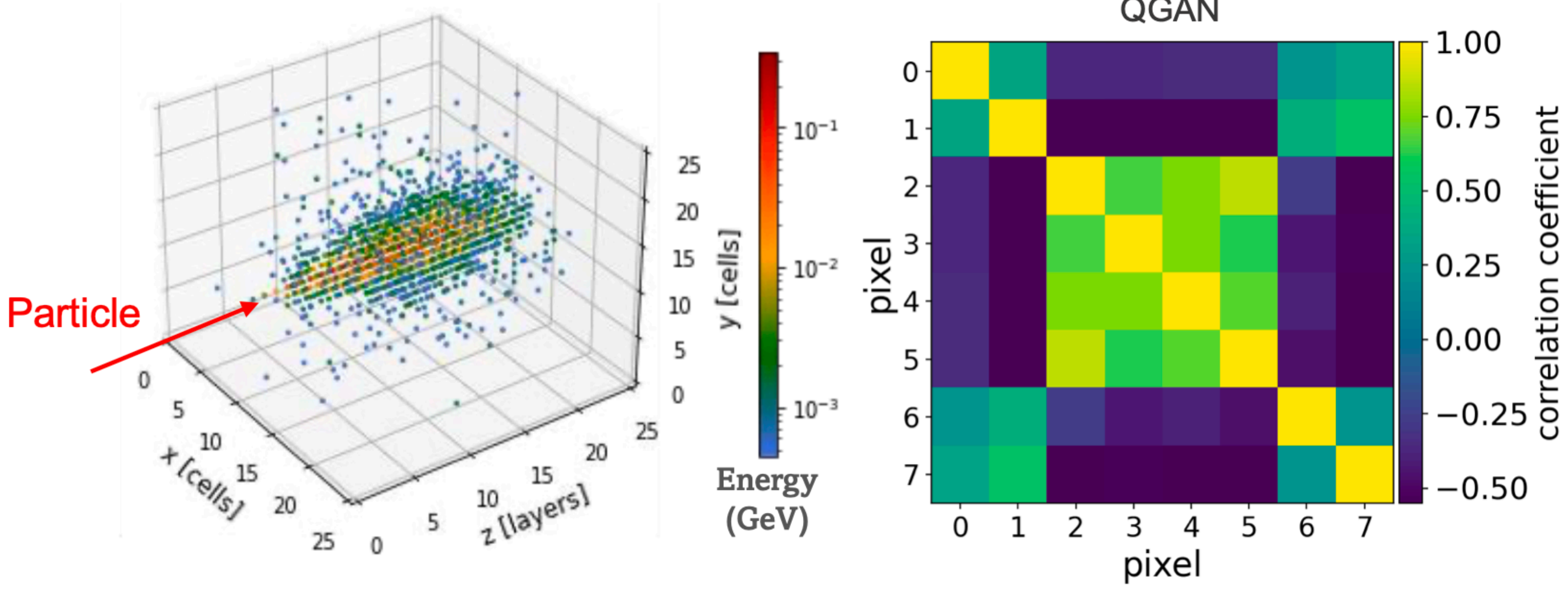


courtesy Michele Grossi

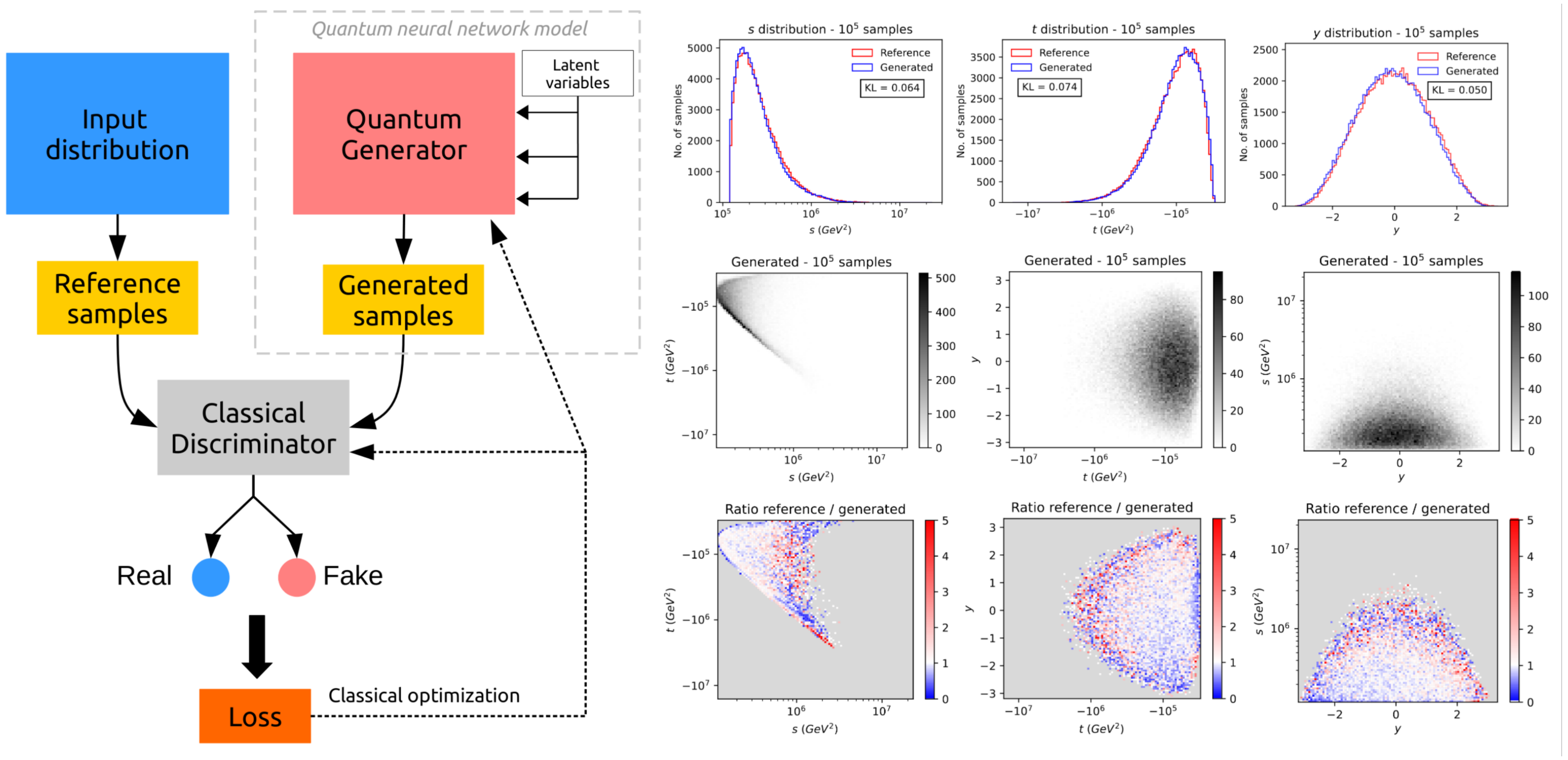
QUANTUM GENERATIVE MODELS IN HEP



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*



Rudolph M., Grossi M. et al., **Trainability Barriers and opportunity in quantum Generative Model.** [arXiv:2305.02881](https://arxiv.org/abs/2305.02881)