

# **Quantum Machine Learning and Optimisation**

Carla Rieger (CERN, TU Munich)

CERN Openlab Summer Student Lectures, 30.7.2024

**Recap: the key features of Quantum Computing** 





**Quantum Superposition State** 

Quantum Entanglement (here: Bell state)



30.07.24

# **Recap: the key features of Quantum Computing**





JITIATIVE

### Interlude: Bell at CERN



John Stewart Bell commenting on the famous Bell's inequalities at CERN in 1982.

Source: https://physicsworld.com/a/saved-by-bell/



30.07.24

#### **CHSH-Game**



- 1. Alice and Bob may agree on a strategy before the game starts but cannot communicate once the game has started. They act cooperatively.
- 2. The referee prepares (binary) bits *x* and *y* independently and at rand . om
- 3. Alice and Bob win if their return answers  $a \in \{0,1\}$  and  $b \in \{0,1\}$  satisfy:  $xy = a \bigoplus b$ .



John Stewart Bell commenting on the famous Bell's inequalities at CERN in 1982.

Image source: https://physicsworld.com/a/saved-by-bell/ and Wikipedia.

Scarani, Valerio. Bell nonlocality. Oxford University Press, 2019.



#### **CHSH-Game**



- 1. Alice and Bob may agree on a strategy before the game starts but cannot communicate once the game has started. They act cooperatively.
- 2. The referee prepares (binary) bits *x* and *y* independently and at rand . om
- 3. Alice and Bob win if their return answers  $a \in \{0,1\}$  and  $b \in \{0,1\}$  satisfy:  $xy = a \bigoplus b$ .
- Upper bound on winning probability:



theory



John Stewart Bell commenting on the famous Bell's inequalities at CERN in 1982.

Violation of Bell's inequality



Shared entanglement and no communication

Image source: https://physicsworld.com/a/saved-by-bell/ and Wikipedia.

Scarani, Valerio. Bell nonlocality. Oxford University Press, 2019.

OUANTUM

TECHNOLOGY



30.07.24



John Stewart Bell commenting on the famous Bell's inequalities at CERN in 1982.

# This implies that the **predictions of quantum theory cannot be accounted for by any local theory.**

Image source: https://physicsworld.com/a/saved-by-bell/

Scarani, Valerio. *Bell nonlocality*. Oxford University Press, 2019.



#### The Nobel Prize in Physics 2022

#### Alain Aspect

"for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science"



© Nobel Prize Outreach. Photo: Stefan Bladh

#### John F. Clauser

"for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science"



© Nobel Prize Outreach. Photo: Stefan Bladh

"for experiments with entangled photons,

establishing the violation of Bell inequalities

and pioneering quantum information science"

Anton Zeilinger

© Nobel Prize Outreach. Photo: Stefan Bladh



John Stewart Bell commenting on the famous Bell's inequalities at CERN in 1982.

Image sources: https://www.nobelprize.org/all-nobel-prizes-2022/ and https://physicsworld.com/a/saved-by-bell/



30.07.24

# **Recap: Basic one qubit gates**

Quantum Theory is unitary — gates are represented by unitary matrices U –



- Pauli matrices (together with identity matrix) form basis of 2x2 matrices
- any 1-qubit rotation can be written as a linear combination of Pauli gates



 $U^{\dagger}U = \mathbb{1}$ 

Nielsen, Michael A., and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.



9

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

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

Apply H on computational basis state

# **Recap: Basic two qubit gate**



Nielsen, Michael A., and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

10



30.07.24

# **Aim of Quantum Computing**

# Do classically intractable computations efficiently on a Quantum Computer leveraging Quantum Effects



# **Applications of Quantum Computing**

#### One may successfully leverage quantum effects for:

- Efficient sampling, search and optimization (e.g., Grover's search algorithm)
- Linear algebra, matrix computations and machine learning (e.g., HHL-algorithm)
- Algorithms and protocols for Cryptography and Communication (e.g., Shor's algorithm, Quantum Key distribution)

Based on previous year's talk



# What is Quantum Machine Learning?



# **Fields in Quantum Computing**



Source: Qiskit Textbook



30.07.24

# **Fields in Quantum Computing**



Source: Qiskit Textbook



# **Fields in Quantum Computing**



Source: Qiskit Textbook





CERN QUANTUM TECHNOLOGY INITIATIVE

30.07.24

Source: Qiskit Textbook



Find underlying structure in labeled data

**Supervised** Learning

- Risk minimization task: minimize the discrepancy between the model's prediction and the target output;
- the best model has minimal expected loss over all data
- model should generalize • well

TIATIVE



- Train an agent to make decisions that maximize rewards
- Rewards are given through ۲ interaction with the environment

Unsupervised Learning

No loss function based on labels but formulate distance between true probability distribution and the model distribution

Find underlying

unlabeled data

structure in

Model should generalize in order to produce samples from true probability distribution

Train agent to optimize its environment-based decisions

Source: Qiskit Textbook





# **Further differentiating Quantum Machine Learning Models**

#### Variational algorithms (e.g., QNN)

- Utilize gradient-based or gradient-free optimization procedure
- May learn data embedding
- The design of the Ansatz circuit can leverage inherent symmetries of data



#### Kernel methods (e.g., QSVM)

- Based on similarity measures between data points
- Choose kernel function based on inherent data structure
- Quantum kernel functions correspond to feature maps
- Encode data in high-dimensional Hilbert space and use inner products evaluated in the feature space to model data properties



#### Energy-based models (e.g., QBM)

- Networks of stochastic binary units are optimized wrt. to their energy
- Inspired conceptually by statistical mechanics

Jerbi, Sofiene, et al. "Quantum machine learning beyond kernel methods." *Nature Communications* 14.1 (2023): 1-8.

Schuld, Maria, Ilya Sinayskiy, and Francesco Petruccione. "An introduction to quantum machine learning." *Contemporary Physics* 56.2 (2015): 172-185.



# **Quantum Machine Learning life cycle**





Source: S. Vallecorsa

## **Quantum Circuits and the Born rule**



Initialization:

→ initialize qubits in computational basis state

An arbitrary quantum circuit generating the state  $|\Psi\rangle$ 



## **Quantum Circuits and the Born rule**



**Evolve initial state:** 

→ Apply set of **unitary** gates that may **encode classical input data** *x* and include **parametrized gates** 

An arbitrary quantum circuit generating the state  $|\Psi\rangle$ 



## **Quantum Circuits and the Born rule**



An arbitrary quantum circuit generating the state  $|\Psi\rangle$ 

#### **Quantum Measurement**

→ retrieve a classical output distribution  $|\langle x|\Psi\rangle|^2$ of classical output states

(with  $x \in \{0,1\}^n$ ) according to Born rule



### Parametrized Quantum Circuits – the data processing pipeline





### Supervised Learning in Quantum Computing: Quantum Classifiers

Goal: learn the *input-output relation* of *labeled* data



Parametrized Quantum Circuit

![](_page_26_Picture_4.jpeg)

QUANTUM

TECHNOLOGY

# **Embedding classical information in a Quantum Circuit**

A tradeoff between depth of input encoding quantum circuit and exponential compression of classical input data

### Angle encoding:

- Classical input encoded using rotational gates (e.g.,  $R_x(\theta)$ )
- Constant depth wrt. to number of encoded features
- Number of qubits n scales linearly with the number of features N'

$$h \propto O(N)$$
,  $n_{gates} \propto n$ 

### Amplitude encoding:

- Classical input is encoded as amplitudes of the quantum state
- *N*-dimensional data point *x* is encoded by a *n*-qubit quantum state with  $N = 2^n$
- Much deeper circuit depth for encoding, see scaling:

$$n \propto O(\log(N)), n_{gales} \propto O(poly(N)) = O(poly(2^{n}))$$

 $(0) \neq 2$  (0) + (1) (0) + (1) (1)

$$|\phi(x)\rangle = \bigotimes_{i=1}^{n} \left( R_x(x_i) |0\rangle \right) = \bigotimes_{i=1}^{n} \left( \cos(x_i/2) |0\rangle - i \sin(x_i/2) |1\rangle \right)$$

![](_page_27_Picture_14.jpeg)

Schuld, Maria, and Francesco Petruccione. *Supervised learning with quantum computers*. Vol. 17. Berlin: Springer, 2018. Image source: Nielsen, Michael A., and Isaac Chuang. "Quantum computation and quantum information." (2002).

![](_page_27_Picture_16.jpeg)

### **Quantum Classifier example: Hierarchical quantum classifiers**

![](_page_28_Figure_1.jpeg)

# **Parameter optimization**

#### The parameter-shift rule (gradient-based)

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

![](_page_29_Figure_3.jpeg)

$$\begin{array}{l} \theta \rightarrow \theta - \eta \nabla_{\theta} f \\ \\ \text{meter } \theta, \\ \\ \theta \mid \text{finite} \end{array}$$

$$\Rightarrow \nabla_{\Theta} \langle \hat{A} \rangle = u \left[ \langle \hat{A} (\Theta + \frac{\pi}{\mu_{u}}) \rangle - \langle \hat{A} (\Theta - \frac{\pi}{\mu_{u}} \rangle \right]$$

 Evaluate Quantum Circuit twice at shifted parameters to compute gradient

Source: https://pennylane.ai/qml/demos/tutorial\_stochastic\_parameter\_shift https://pennylane.ai/qml/demos/tutorial\_spsa

![](_page_29_Picture_8.jpeg)

# **Parameter optimization**

# Simultaneous perturbation stochastic approximation (SPSA) (gradient-free)

→ If gradient computation is not possible, too resource-intensive,
 or noise-robustness required (slower convergence but fewer function evaluations)
 → The gradient is approximated by two sampling steps, and parameters are
 perturbed in all directions simultaneously

$$\hat{g}(\hat{\theta}) = \frac{f(\theta) + \epsilon^{''}}{c} \quad \text{random} \\ \text{output perturbation} \\ \hat{g}(\hat{\theta}_{k}) = \frac{y(\hat{\theta}_{k} + c_{k}\Delta_{k}) - y(\hat{\theta}_{k} - c_{k}\Delta_{k})}{2c_{k}\Delta_{k}i}$$

 $C_k \ge 0$ ,  $\Delta_k = (\Delta_{k_1}, \Delta_{k_2}, \dots, \Delta_{k_p})^T$  perturbation vector

Iterative update rule comparable to classical stochastic gradient descent

- https://pennylane.ai/qml/demos/tutorial\_stochastic\_parameter\_shift

https://pennylane.ai/qml/demos/tutorial\_spsa

![](_page_30_Picture_7.jpeg)

(~ randomly sampled from zero-mean distr.)

stochastic

estimat

# **Challenges when using Parametrized Quantum Circuits**

- Efficient data handling and data embedding
- Find balance: Generalization and representational power vs. Convergence
  - Problem of barren plateaus and vanishing gradients in optimization landscape
  - How well can we survey the Hilbert space (expressibility)?
- Current hardware limitations
  - Limited number of qubits and connectivity
  - Quantum Noise Effects (decoherence, measurement errors or gate-level errors)
  - Efficient interplay between a classical and a quantum computer

![](_page_31_Picture_9.jpeg)

# What is Quantum Advantage in QML?

#### **Multiple considerations:**

- 1. Runtime speed-up
- 2. Sample complexity
- 3. Representational power

![](_page_32_Figure_5.jpeg)

Bloch sphere: only the marked points are produced by the Clifford operators acting on a computational basis state

#### This includes considerations regarding **classical intractability**:

### Focus on Quantum Circuits that are not efficiently simulable classically

Nielsen, Michael A., and Isaac Chuang. "Quantum computation and quantum information." (2002). Gottesman, Daniel. "The Heisenberg representation of quantum computers." *arXiv preprint quant-ph/9807006* (1998). See also: - Kübler, Jonas, Simon Buchholz, and Bernhard Schölkopf. "The inductive bias of quantum kernels." *Advances in Neural Information Processing Systems* 34 (2021): 12661-12673. - Huang, HY., Broughton, M., Mohseni, M. *et al.* Power of data in quantum machine learning. *Nat Commun* **12**, 2631 (2021). https://doi.org/10.1038/s41467-021-22539-9

![](_page_32_Picture_10.jpeg)

# Interlude: Efficient classical simulation of Clifford circuits

The Gottesman-Knill theorem

A quantum circuit build up of Clifford gates can be efficiently simulated on a classical computer. (Qubit preparation and measurement in

computational basis.)

There are more detailed considerations of cases with different computational complexities.

→ Even highly entangled states can be simulated efficiently classically.

30.07.24

Generating set of the Clifford group: (H, S, CNOT)

Nielsen, Michael A., and Isaac Chuang. "Quantum computation and quantum information." (2002). Gottesman, Daniel. "The Heisenberg representation of quantum computers." *arXiv preprint quant-ph/9807006* (1998).

![](_page_33_Picture_8.jpeg)

# What is Quantum Advantage in QML?

#### **Multiple considerations:**

- Runtime speed-up
- Sample complexity
- Representational power

### **Practical advantage** → Practical implementations on NISQ devices

→ Need for performance metrics and fair comparisons to classical models

Nielsen, Michael A., and Isaac Chuang. "Quantum computation and quantum information." (2002). Gottesman, Daniel. "The Heisenberg representation of quantum computers." *arXiv preprint quant-ph/9807006* (1998). See also: - Kübler, Jonas, Simon Buchholz, and Bernhard Schölkopf. "The inductive bias of quantum kernels." *Advances in Neural Information Processing Systems* 34 (2021): 12661-12673.

- Huang, HY., Broughton, M., Mohseni, M. *et al.* Power of data in quantum machine learning. *Nat Commun* **12**, 2631 (2021). https://doi.org/10.1038/s41467-021-22539-9

![](_page_34_Picture_9.jpeg)

![](_page_35_Picture_0.jpeg)

![](_page_35_Picture_1.jpeg)

30.07.24

# **Quantum Circuit Born Machine for Event Generation**

![](_page_36_Figure_1.jpeg)

Muon fixed target scattering experiment

- MFCs are bosons which appear in beyondthe-standard-model theoretical frameworks and are candidates for dark matter
- Monte Carlo calculations are expensive in time and CPU consumption

QUANTUM

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

![](_page_36_Picture_6.jpeg)

![](_page_37_Figure_1.jpeg)

![](_page_37_Figure_2.jpeg)

Muon fixed target scattering experiment

![](_page_37_Figure_4.jpeg)

### **Quantum Circuit Born Machine for Event Generation**

#### **Born machine:**

using parametrized quantum circuit  $|\psi(\theta)\rangle$ 

Produces statistics according to Born's measurement rule

# **Quantum Circuit Born Machine for Event Generation**

- Generate samples of discrete PDFs with Born machine
- Train using Maximum Mean Discrepancy loss function:

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

 efficient way to generate multivariate (and conditional) distributions for NISQ devices (suggested by numerical evidence)

![](_page_38_Figure_5.jpeg)

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

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

![](_page_38_Picture_8.jpeg)

![](_page_39_Figure_1.jpeg)

 $w_{i,j} = 1$  for  $\forall i, j$ 

#### The MaxCut problem (NP-complete)

**Goal:** partition the graph into two groups and maximize the number of edges connecting both partitions

→ assign binary variables to nodes

Zhou, Leo, et al. "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices." *Physical Review X* 10.2 (2020): 021067.

![](_page_39_Picture_6.jpeg)

Aiming to solve a **QUBO problem** of the form:  $f_Q(x) = x^T Q x$ ,  $x \in \{0,1\}^N$ Map to an **Ising Hamiltonian** of the general form  $H = \sum_{i,j} J_{i,j} Z_i Z_j$ 

![](_page_40_Figure_2.jpeg)

Zhou, Leo, et al. "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices." *Physical Review X* 10.2 (2020): 021067.

![](_page_40_Picture_4.jpeg)

$$f_Q(x) = x^T Qx = \sum_{i,j=1}^{N} Q_{i,j} x_i x_j \qquad x^* = \underset{x \in \{0,1\}^N}{\operatorname{arg min}} f_Q(x) \quad \text{QUBO problem}$$

$$\int_{z_i \in \{1,-1\}}^{x_i \in \{0,1\}} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

$$\int_{Q_i = 1}^{N-1} \sum_{j>i}^{N-1} Q_{i,j} (1-z_i) (1-z_j) \quad \text{Ising-type Hamiltonian}$$

30.07.24

QUANTUM

TECHNOLOGY INITIATIVE  $w_{i,j} = 1 \text{ for } \forall i, j$   $w_{12}$   $w_{12}$   $w_{25}$   $w_{45}$   $w_{15}$   $W_{15}$   $W_{35}$   $W_{5}$   $W_{45}$   $W_{45}$   $W_{15}$   $W_{15}$ 

MaxCut problem

Zhou, Leo, et al. "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices." *Physical Review X* 10.2 (2020): 021067.

QAOA Ansatz and hybrid optimization procedure

![](_page_42_Figure_2.jpeg)

![](_page_42_Figure_3.jpeg)

MaxCut Graph

Classically optimize angles: Hybrid procedure

> Zhou, Leo, et al. "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices." Physical Review X 10.2 (2020): 021067.

![](_page_42_Picture_7.jpeg)

## **Quantum Databases in a General Context**

![](_page_43_Figure_1.jpeg)

Rieger, Carla, et al. **"Operational Framework for a Quantum Database."** *arXiv preprint arXiv:2405.14947* (2024).

![](_page_43_Picture_3.jpeg)

# **Quantum Databases (QDB)**

How do we operate on superposition states containing a quantum index register correlated to data registers?

- QDB's are relevant for quantum algorithms to operate on quantum database states and dynamically manipulate them.
- Make use of exponential compression due to the usage of the superposition principle.
  - Manipulation operations are defined to mimic classical database operations.

![](_page_44_Picture_5.jpeg)

 $\left| \text{QDB}^{(k)} \right\rangle = \sum^{k-1} \psi_j \left| j \right\rangle_I \left| d_j \right\rangle_D \in \mathcal{H}_I \otimes \mathcal{H}_D$ 

Rieger, Carla, et al. **"Operational Framework for a Quantum Database."** *arXiv preprint arXiv:2405.14947* (2024).

![](_page_44_Picture_8.jpeg)

# **Quantum Databases (QDB)**

**Example**: QDB Prepare Algorithm with qubits for  $S_{ta}^{[0]}$ 

![](_page_45_Figure_2.jpeg)

Constant-depth quantum circuit

![](_page_45_Figure_4.jpeg)

 $\tilde{Y}(1)$ 

 $\left(\frac{b1}{b1+1}\right)$ 

 $\tilde{Y}$  (

 $\tilde{Y}(1)$ 

 $\tilde{Y}\left(\frac{b100}{b101+1}\right)$ 

 $Y\left(\frac{1}{2}\right)$ 

 $Y\left(\frac{1}{2}\right)$ 

 $Y\left(\frac{1}{2}\right)$ 

 $|0\rangle$ 

 $|0\rangle$  -

 $|0\rangle$  –

Linear-depth (in no. qubits) quantum circuit (e.g., k = 14)

![](_page_45_Figure_6.jpeg)

arXiv:2405.14947 (2024).

CERN QUANTUM TECHNOLOGY INITIATIVE

QML and Optimization

# **Outlook on QML and summary**

Research on QML applications in High Energy Physics is producing a **large number of prototypical algorithms for potential future use-cases:** 

- Currently focus on *algorithms for data processing* in a *controlled* environment for current hardware
- Preliminary hints for advantage in terms of *representational power of quantum states*
- Mostly, algorithm performance is *as good as* the classical counterpart
- Need more robust studies to relate architecture of quantum computational model and its performance to data sets
- *Identify use-cases* where quantum approach is provably *more efficient* than classical model
- Studying QML algorithms today *links Quantum computing and Learning Theory* and draw separation between classical and quantum learner

![](_page_46_Picture_9.jpeg)

![](_page_47_Picture_0.jpeg)

![](_page_47_Picture_1.jpeg)

# carla.sophie.rieger@cern.ch

![](_page_47_Picture_3.jpeg)

30.07.24