香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Quantum Recurrent Embedding Neural Networks

Mingruij Jing, Erdong Huang, Xiao Shi, Shengyu Zhang, Xin Wang

IW-QBM 2025





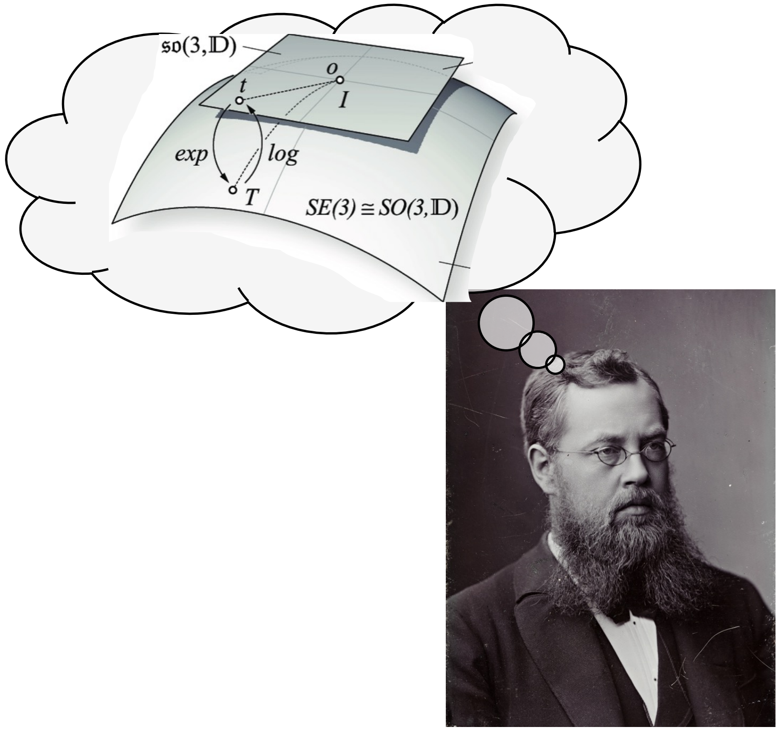
[1] Quantum Recurrent Embedding Neural Network https://arxiv.org/pdf/2506.13185 Email: shengyzhang@tencent.com; felixxinwang@hkust-gz.edu.cn

Content:

- Preliminary of Quantum Machine Learning
  - Quantum neural networks
  - Dynamical Lie algebra

- Quantum Recurrent Embedding Neural Network
  - Circuit framework
  - Quantum supervised learning
  - Trainability
  - SPT phase detection

- Concluding remarks
  - Take-Home message

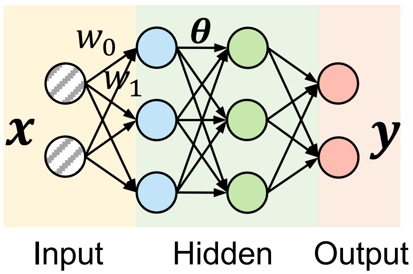


香港科大(广州)
HKUST(GZ)

Preliminary of Quantum Machine Learning

Quantum Neural Networks

Classical NNs

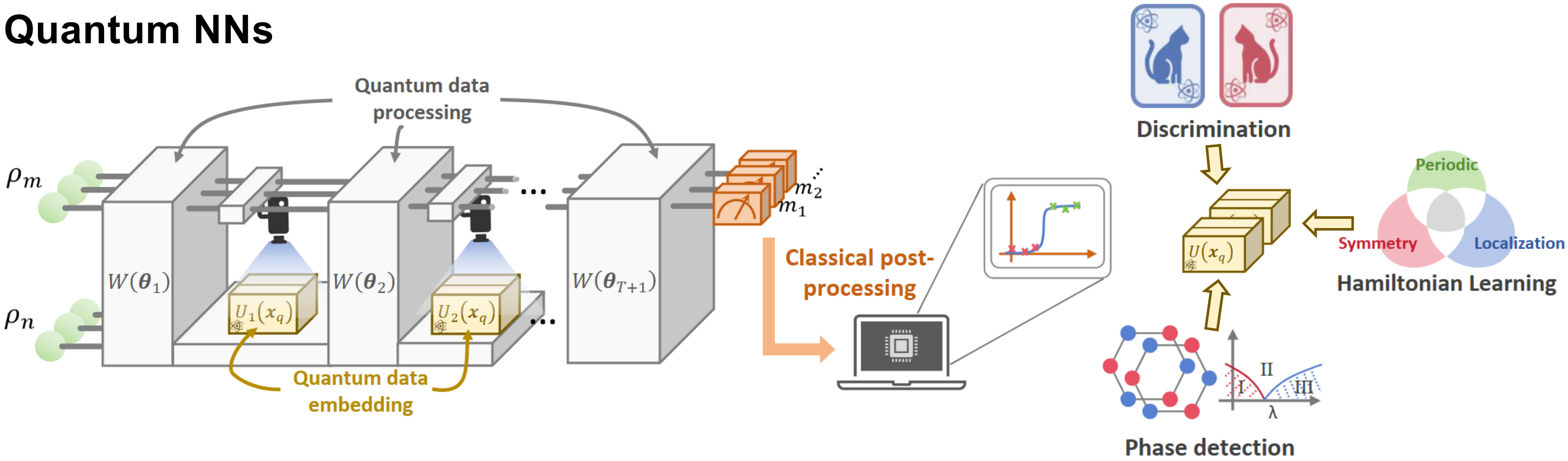


Data vector $\boldsymbol{x} = [x_0, \cdots, x_d]^T$

Process: Adjust the weights of the connections $\boldsymbol{w}, \boldsymbol{\theta} = \{w_0, w_2, \cdots, \theta_1, \theta_2, \cdots\}$

Output: $\boldsymbol{y} = [y_0, y_1, \cdots, y_m]^T$

Quantum NNs



Quantum Neural Networks

香港科大(广州) HKUST(GZ)

- QNNs have various applications in Quantum Machine Learning (QML) [1]

QNNs

Enhanced quantum computing

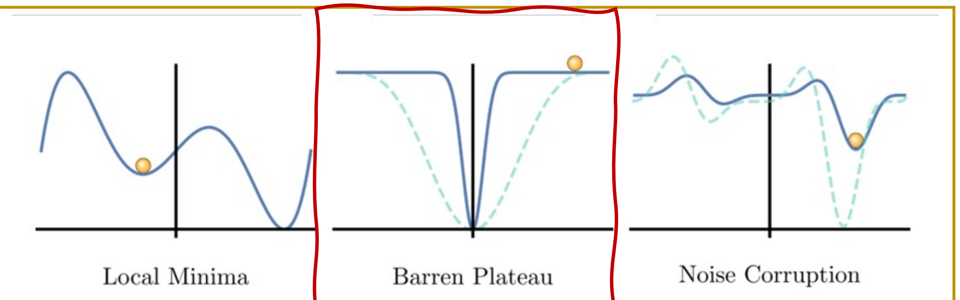
Classical data analysis

Quantum machine perception

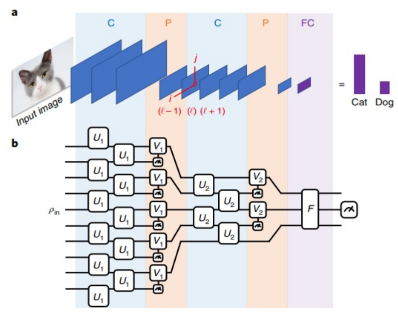
Quantum simulation

Quantum sensing, predict properties, compiling circuits, … …

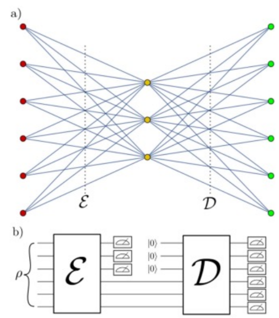
Challenges



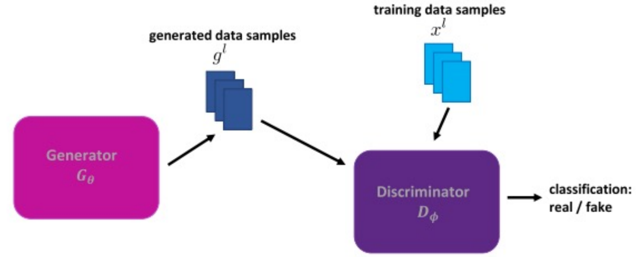
- Famous examples



Quantum convolutional neural network



Quantum autoencoder



Quantum Generative Adversarial Networks

[1] Cerezo, Marco, et al. "Challenges and opportunities in quantum machine learning." *Nature computational science* 2.9 (2022): 567-576.

Dynamical Lie algebra

- For QNN expressed as $U(\boldsymbol{\theta}) = \prod_{l=1}^{L}\left(\prod_{k=1}^{K} e^{i\theta_{l,k}H_l}\right)$, the DLA of the circuit is defined as [1],

$$\mathfrak{g} = \text{span}_{\mathbb{R}}\langle iH_1, iH_2, \cdots, iH_L\rangle_{Lie} = \text{span}_{\mathbb{R}}\langle i\mathcal{G}\rangle_{Lie}$$

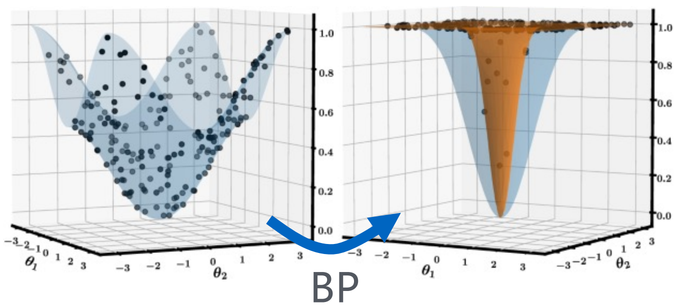
- In the finite case, $\mathfrak{g} = \mathfrak{c} \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_M$ where each $\mathfrak{g}_j$ is simple and $\mathfrak{c}$ is the center of $\mathfrak{g}$.
- DLA contains **important information** that **drive or generate** the system's evolution over time.

- If the QNN is deep enough to form a unitary 2-design on $e^{\mathfrak{g}} \subset \mathcal{U}(d)$ (compact Lie group) [2]

$$\mathbb{E}_{\boldsymbol{\theta}}\left[\partial_{l,k}\mathcal{L}(\rho, O)\right] = 0 \qquad \text{Var}_{\boldsymbol{\theta}}\left[\partial_{l,k}\ell(\rho, O)\right] \in \mathcal{O}\left(\sum_j \frac{1}{d_{\mathfrak{g}_j}^2}\right)$$

- $\ell(\rho, O) = \text{Tr}\left(U(\boldsymbol{\theta})\rho U^{\dagger}(\boldsymbol{\theta})O\right)$; $H_{\mathfrak{g}}$ is the projection of H onto $\mathfrak{g}$.

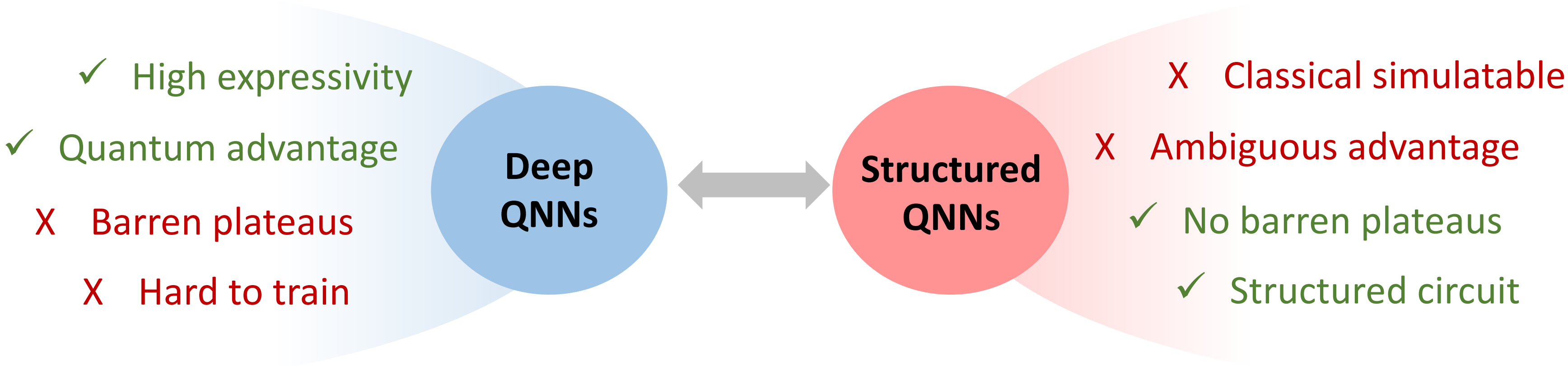
The Lie algebraic theory of QNNs unifies the study of various sources of barren plateaus (BP).



[1] Ragone, Michael, et al. "A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits." *Nature Communications* (2024).
[2] Fontana, Enrico, et al. "Characterizing Barren Plateaus in Quantum ansätze." *Nature Communication* (2024).

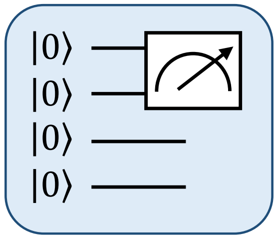
Standing at the crossroads

- Trade-off between expressivity of QNNs, and classical simulatability [1]

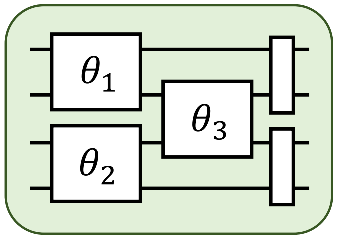


- Strategies for design deep and scalable QNNs that can avoid BP:

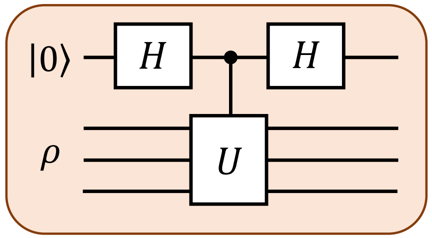
Localized states and measurements



Relatively simple circuit



Algorithmic structure

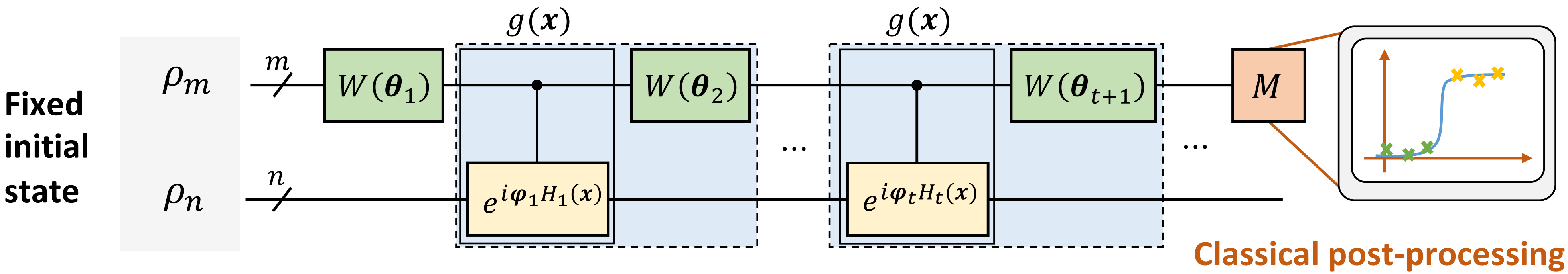


[1] Cerezo, Marco, et al. "Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing." *Nature Communications* (2025).

Quantum Recurrent Embedding Neural Networks

Circuit framework of QRENN

Circuit Model of QRENN



- Assuming $\{\boldsymbol{\theta}: W(\boldsymbol{\theta})\}$ spans $\mathrm{SU}(2^m)$; Hamiltonian $[H_t(\boldsymbol{x}),\ H_\tau(\boldsymbol{x})] = 0$, for any $t \neq \tau$; $m \in O(\log n)$

Theorem

- The DLA of QRENN can be decomposed into

$$\mathfrak{g}_{\mathrm{QRENN}} \simeq \mathfrak{c} \oplus \mathfrak{su}(2^m)^{\oplus r},$$

where $\mathfrak{c} := \mathrm{span}_{\mathbb{R}}\{iI_m \otimes H_t(\boldsymbol{x}) : t \in [T]\}$

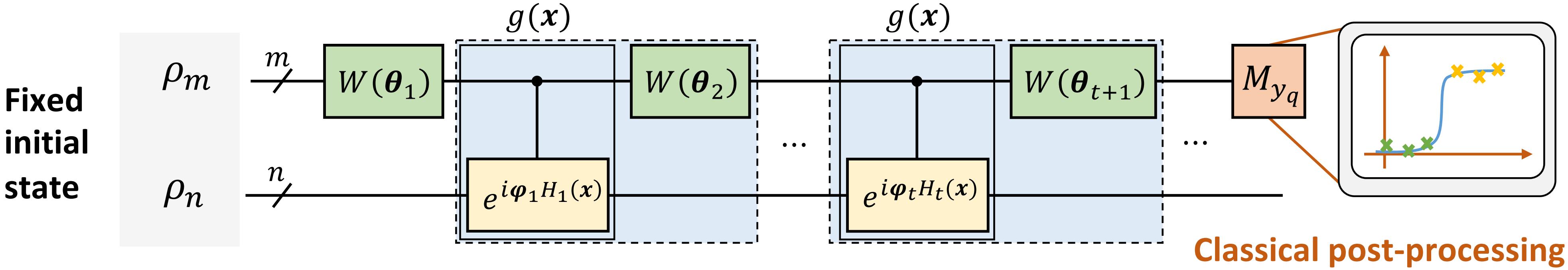
The control embedding makes the DLA a direct sum of $\mathfrak{su}(2^m)$

Simple algebraic structure and hard to simulate classically ☺

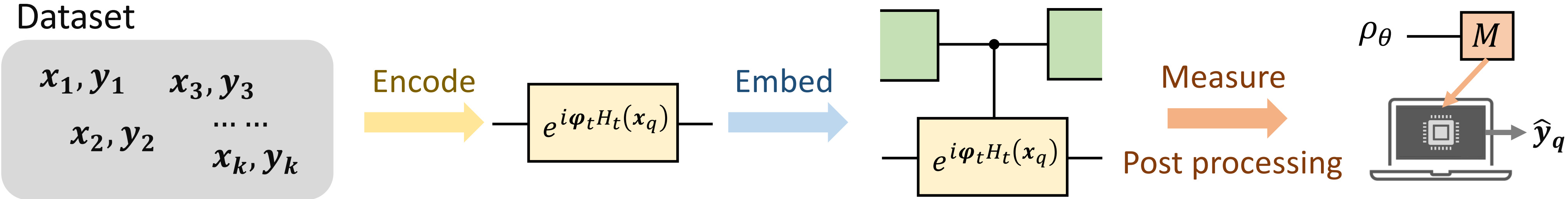
- r is the number of distinct *joint eigenspaces* from $\{H_t(\boldsymbol{x})\}_t$

What can we do with QRENN?

Circuit Model of QRENN



- By choosing suitable measurement M, we can solve quantum supervised learning tasks using QRENN.
- The circuit contains the QSTV primitives; $H_t(\boldsymbol{x})$ are generally not sparse; Decision function can have high degrees (BQP-complete even with $m \in \mathcal{O}(\log n)$ [1][2])



[1] Montanaro, Ashley, and Changpeng Shao. "Quantum and classical query complexities of functions of matrices." *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. 2024.

[2] Gharibian, Sevag, and François Le Gall. "Dequantizing the quantum singular value transformation: hardness and applications to quantum chemistry and the quantum PCP conjecture." Proceedings of the 54th annual ACM SIGACT symposium on theory of computing. 2022.

Quantum supervised learning

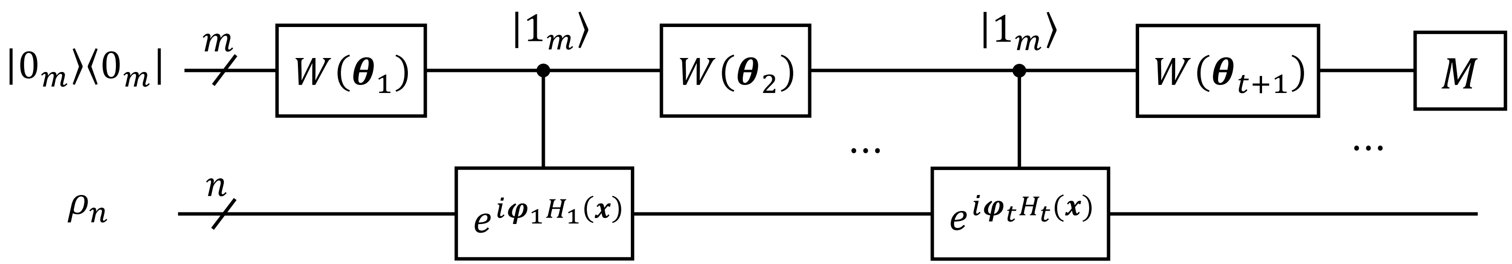
- Given a batch training set $\mathcal{T} = \{(y_q, \boldsymbol{x}_q)\}_q$ with $|\mathcal{T}| = Q$

$$\mathrm{MSE} = \frac{1}{Q}\sum_{q=1}^{Q}\left(y_q - \mathrm{Tr}\left(U(\boldsymbol{x}_q;\boldsymbol{\theta},\boldsymbol{\varphi})\rho_0 U(\boldsymbol{x}_q;\boldsymbol{\theta},\boldsymbol{\varphi})^\dagger M\right)\right)^2$$

- **Hard to estimate on quantum devices**.
- Inspired from quantum hypothesis testing, design $M_1, M_2, \cdots M_k$ forming POVMs. We define the total loss

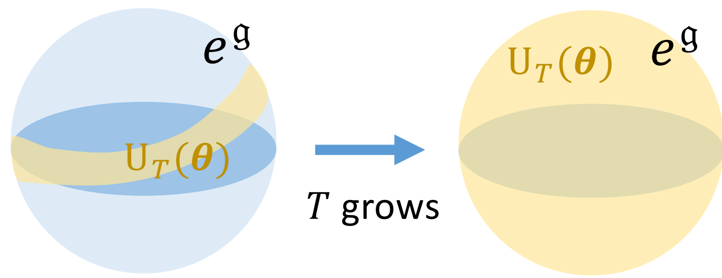
$$\mathcal{L}(\boldsymbol{\theta},\boldsymbol{\varphi}) = 1 - \frac{1}{Q}\sum_{q=1}^{Q}\mathrm{Tr}\left(U(X_q;\boldsymbol{\theta},\boldsymbol{\varphi})\rho_0 U(X_q;\boldsymbol{\theta},\boldsymbol{\varphi})^\dagger M_{y_q}\right)$$

Main theorem on trainability



$$M_0 = \frac{I_{2^m} - Z^{\otimes m}}{2} \otimes I_{2^n};$$
$$M_1 = \frac{I_{2^m} + Z^{\otimes m}}{2} \otimes I_{2^n}$$

- To reach max. expressivity, assuming **sufficiently deep** model.

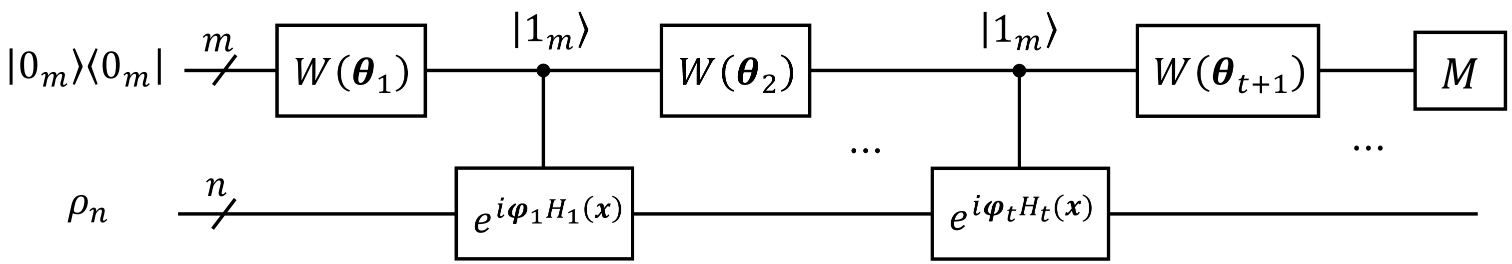


- For QNNs with depth $O(poly(n))$ [2], the circuit achieves 2-design of the compact Lie group and, hence, $\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\varphi}}[\partial_{t,\mu}\mathcal{L}] = 0$ [1]

- By setting the number of slot T polynomially in n, random initialization can give elements fully mixed in $e^{\mathfrak{g}}$

[1] Ragone, Michael, et al. "A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits." *Nature Communications* (2024).
[2] Fontana, Enrico, et al. "Characterizing Barren Plateaus in Quantum ansätze." *Nature Communication* (2024).

Main theorem on trainability



$$M_0 = \frac{I_{2^m} - Z^{\otimes m}}{2} \otimes I_{2^n};$$

$$M_1 = \frac{I_{2^m} + Z^{\otimes m}}{2} \otimes I_{2^n}$$

Theorem

- For $m \in O(\log n)$, if ρ_n has sufficiently large 'overlap', i.e., $\Omega\left(\frac{1}{poly(n)}\right)$ with the union of the image spaces of $\{H_t\}_t$, where $H_t = H_t(\boldsymbol{x})$, then,

$$\mathrm{Var}_{\boldsymbol{\theta},\boldsymbol{\varphi}}[\partial_{t,\mu}\mathcal{L}] \geq \Omega\left(\frac{1}{poly(n)}\right)$$

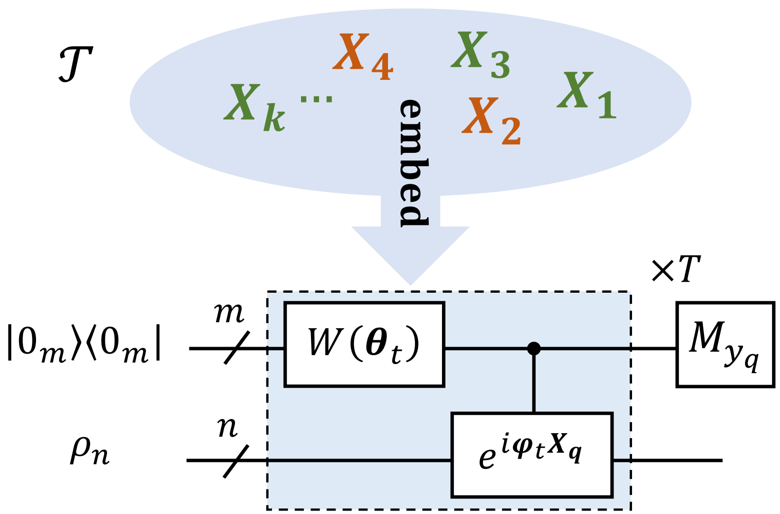
- To prevent from BP, the dimension of data processing register must be restricted to $\log n$ scaled.
- The initial state ρ_n should serve as a 'good' probe that can interact with H_t's
- The overlap assumption is often **used in advanced quantum algorithm design** for achieving advantages.

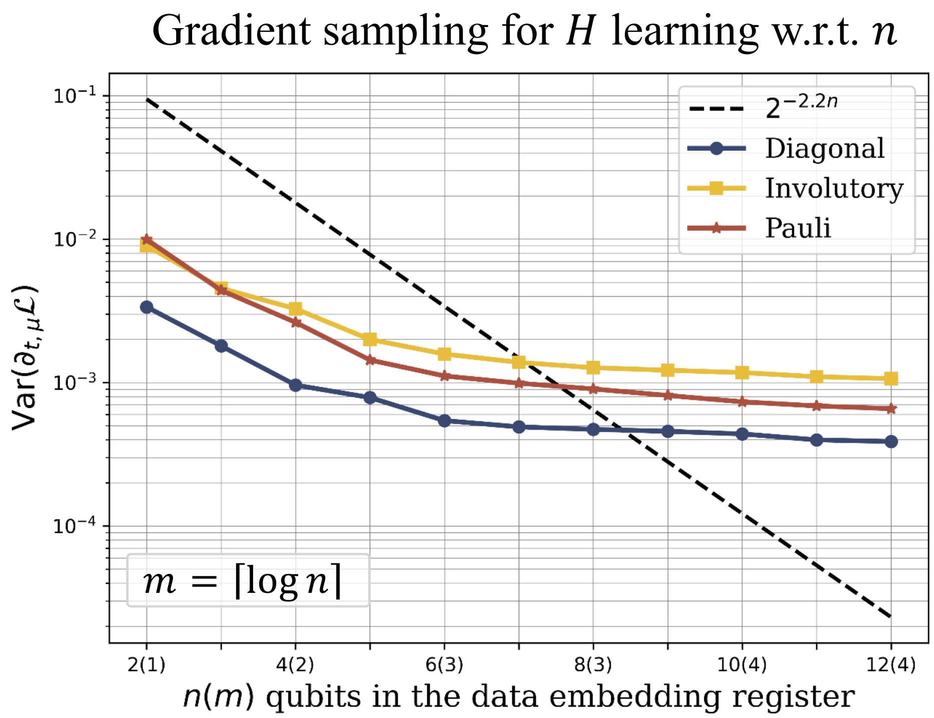
Numerical results on trainability

香港科大(广州) HKUST(GZ)

- **Task:** to classify whether a Hamiltonian is Pauli, involutory or diagonal

Input: X_q → **Output:** class label y_q





- Gradient sampling experiments, 500 random initial parameters $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ of the model, ρ_n being fixed.
- For each dataset in diagonal, involutory and pauli sets, **50 Hamiltonians with feature** is generated and mixed with another **50 random Hermitian matrices** (from Haar unitary).

Supervised learning on quantum data

香港科大(广州) HKUST(GZ)

- Apart from classifying Hamiltonians, what else?

Problem

Given a cluster-Ising model with periodic boundary conditions

$$H(\lambda) = -\sum_{j=1}^{N} X_{j-1} Z_j X_{j+1} + \lambda \sum_{j=1}^{N} Y_j Y_{j+1}.$$

where X, Y and Z are Pauli matrices. SPT phase in the Hamiltonian model [4]:

- A cluster : $\lambda < 1$
- An antiferromagnetic phase : $\lambda > 1$.

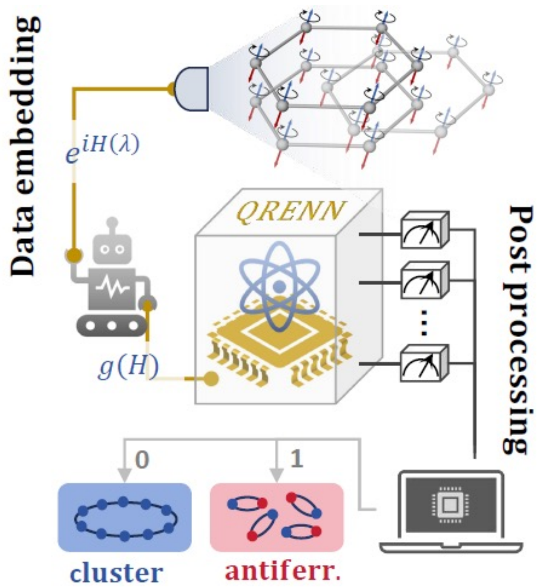
Can we detect different symmetry-protected topological (SPT) phases of physical models via QRENN?

[4] Li, Weikang, Zhi-de Lu, and Dong-Ling Deng. "Quantum neural network classifiers: A tutorial." *SciPost Physics Lecture Notes* (2022): 061.

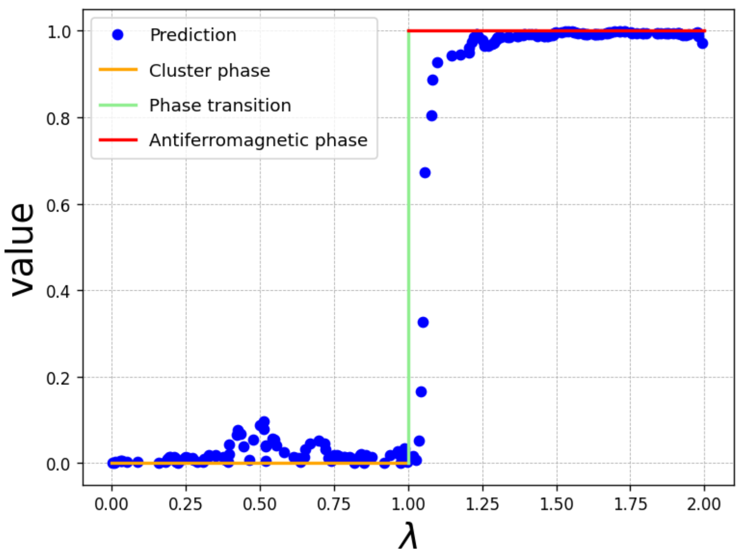
SPT phase detection

香港科大(广州) HKUST(GZ)

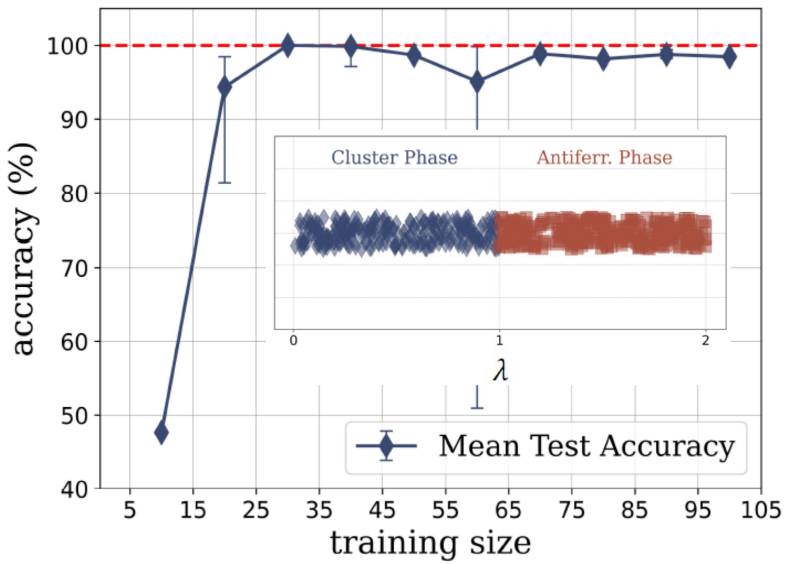
- **QRENN model in learning SPT phase**



Train:



Prediction:



Case 1: slots $= 10$, $m = 1$ and $n = 8$, initial state $|0\rangle \otimes |+\rangle^{\otimes 8}$. **Outcome**: Training 40 data uniformly generated by sampling $\lambda \in [0,2]$. Achieve 98.72% accuracy on 560 testing data.

Case 2: slots $= 10$, $m = 1$ and $n = 8$, initial state $|0\rangle \otimes |+\rangle^{\otimes 8}$. **Outcome**: Train with different data sizes. Find an improvement in performance as training size increases.

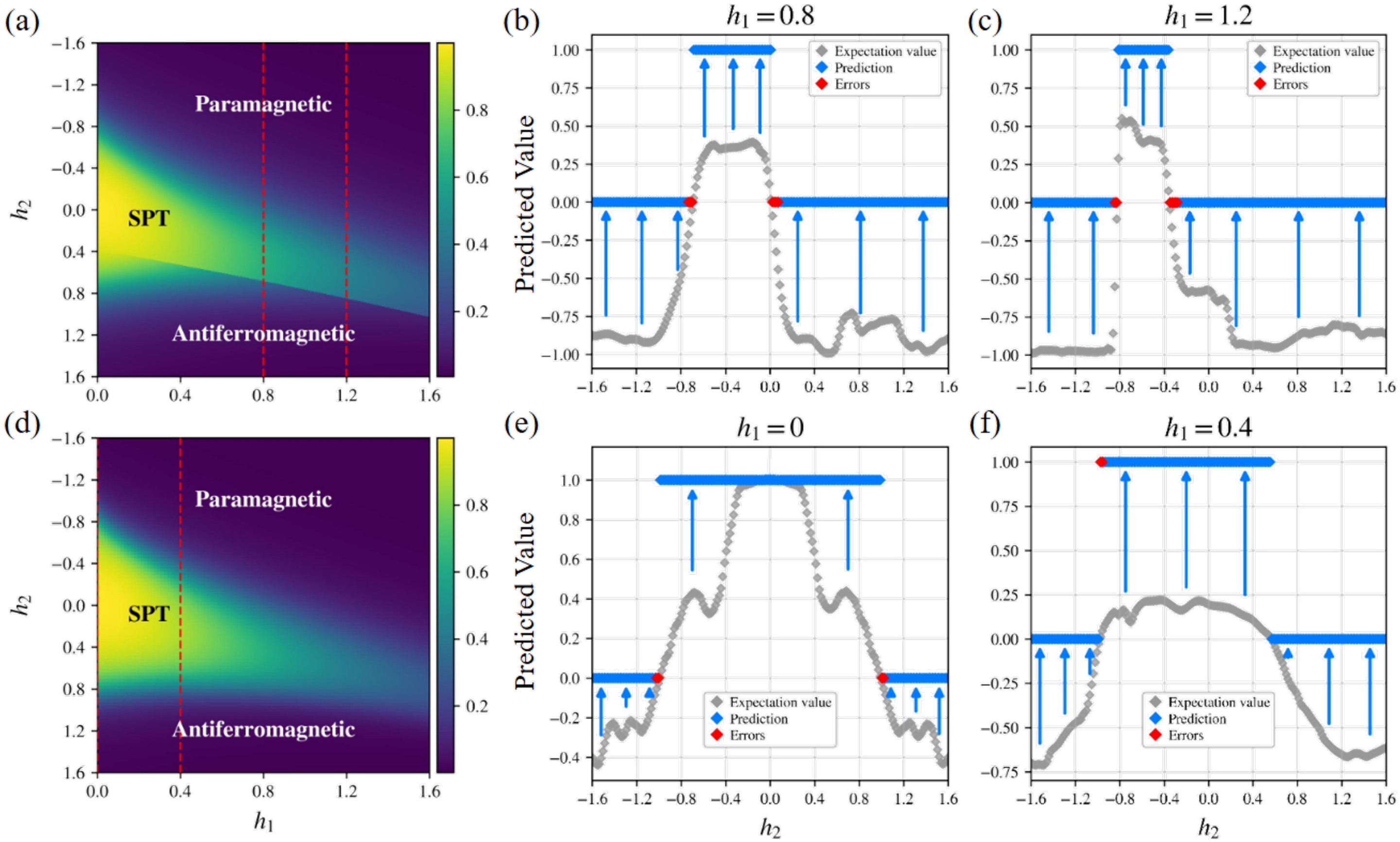
More on SPT phase detection

- Another model with two parameters:

$$H(h_1, h_2) = -\sum_{j=1}^{N} Z_{j-1}X_jZ_{j+1} + h_1\sum_{j=1}^{N} X_j + h_2\sum_{j=1}^{N-1} X_jX_{j+1}$$

Settings: $slots = 60$, $m = 1$ and $n = 7$(a,b,c), $n = 9$(d,e,f) , initial state $|0\rangle \otimes |\psi(h_1)\rangle^{\otimes n}$.

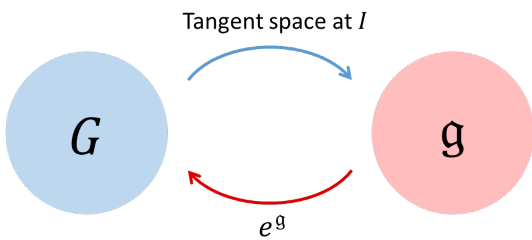
Outcome: Training 100 data uniformly generated by sampling $h_2 \in [-1.6, 1.6]$. Achieve 98% accuracy on 200 testing data.



香港科大(广州)
HKUST(GZ)

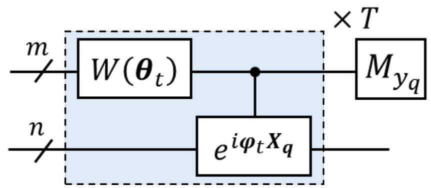
Concluding remarks

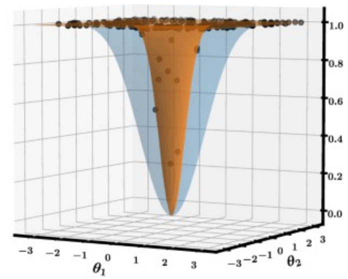
Conclusion



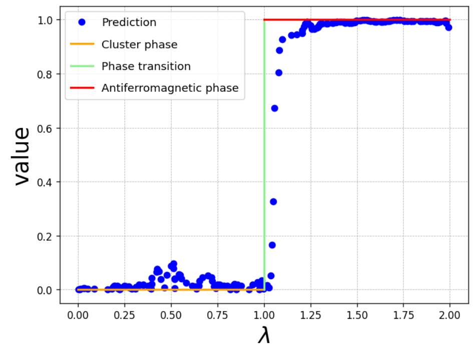
- **Recent developments in the Quantum Machine Learning**
  - Quantum neural networks
  - Dynamical Lie algebra for barren plateaus

$$\mathfrak{g} = \text{span}_{\mathbb{R}}\langle iH_1, iH_2, \cdots, iH_L\rangle_{Lie}$$





- **Quantum Recurrent Embedding Neural Network**
  - Inspiration to QNNs design ⇒ QRENN
  - Can avoid BP in quantum supervised learning
  - Application in SPT phase detection



Outlooks

- **A more rigorous proof on the BQP-completeness of the model.**
- **The control embedding can be hard to realized physically.**
  - Any other simpler embedding method?
- **Other thoughts of Quantum Recurrent Embedding Neural Network**
  - Treat it as a quantum neuron, can we build up larger structures?
  - Application in quantum sensing?
  - Efficient warm-start on initial probe state

~Thanks for watching~

QUAIR Group

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)