Quantum Recurrent Embedding Neural Networks[1]

Presenter: Mingruij Jing
Date: 09252025

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

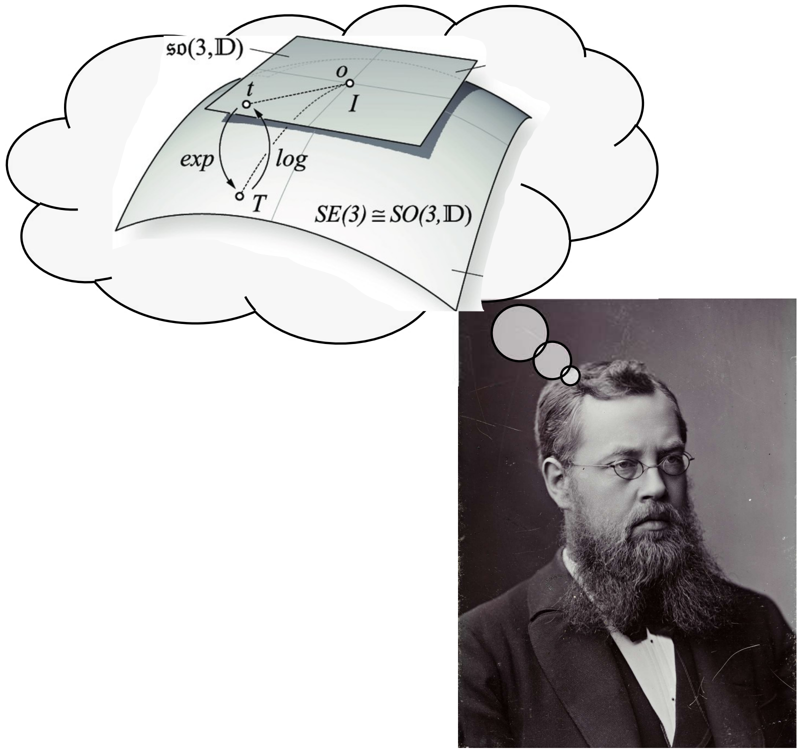
[1] Quantum Recurrent Embedding Neural Network https://arxiv.org/pdf/2506.13185 Email: shengyzhang@tencent.com; felixxinwang@hkust-gz.edu.cn
[2] Ragone, M., Bakalov, B.N., Sauvage, F. et al. A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits. Nat Commun 15, 7172 (2024). https://doi.org/10.1038/s41467-024-49909-3
[3] Fontana, E., Herman, D., Chakrabarti, S. et al. Characterizing barren plateaus in quantum ansätze with the adjoint representation. Nat Commun 15, 7171 (2024). https://doi.org/10.1038/s41467-024-49910-w

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**

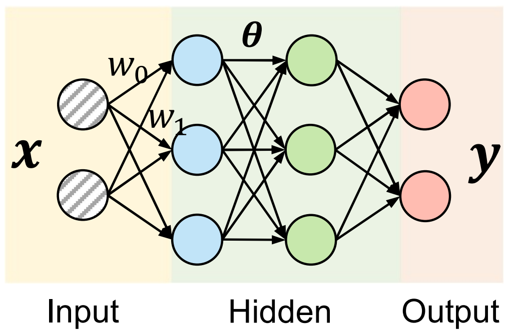
- **Concluding remarks**
  - **Take-Home message**



Preliminary of Quantum Machine Learning

Quantum Neural Networks

Classical NNs

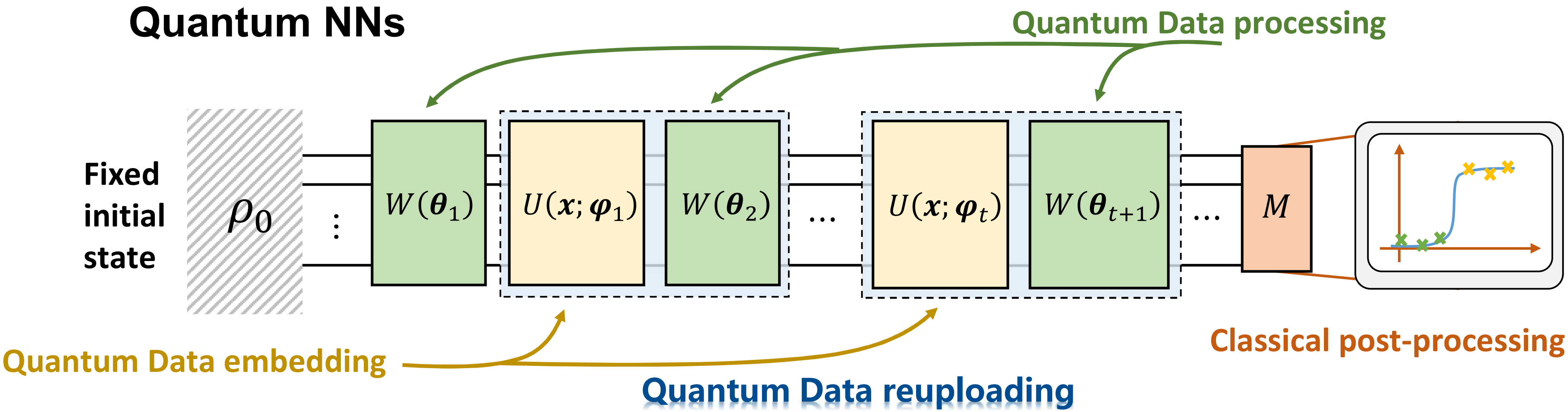


Data vector $\boldsymbol{x} = [x_0, \cdots, x_d]^T$

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

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

Quantum NNs



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],

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

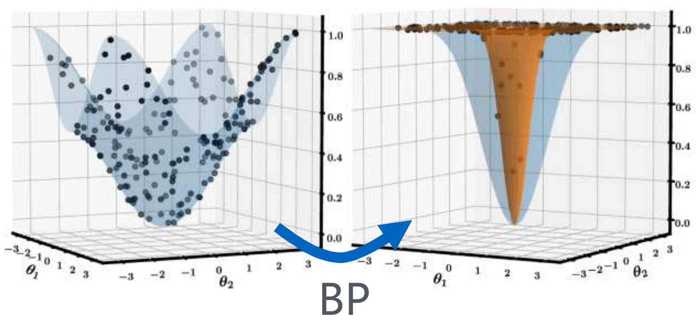
- In the finite case, $\mathfrak{g} = \mathfrak{c} \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_M$ where each $\mathfrak{g}_j$ is simple and $\mathfrak{c}$ is the center of $\mathfrak{g}$.

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

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

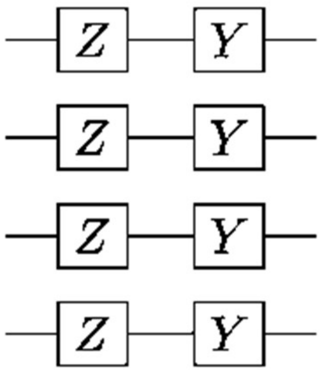
- $\mathcal{L}(\rho, O) = \mathrm{Tr}\left(U(\boldsymbol{\theta})\rho U^\dagger(\boldsymbol{\theta})\mathrm{O}\right)$; $H_{\mathfrak{g}}$ is the projection of H onto $\mathfrak{g}$.

The Lie algebraic theory of QNNs unifies the study of the trainability of QNNs up to a uniform initialization, i.e., Barren Plateaus (BP).



[1] Allcock, Jonathan, et al. "On the dynamical Lie algebras of quantum approximate optimization algorithms." *arXiv preprint arXiv:2407.12587* (2024).
[2] Fontana, Enrico, et al. "The Adjoint Is All You Need: Characterizing Barren Plateaus in Quantum Ans¥" atze." *Nature Communication* (2023).

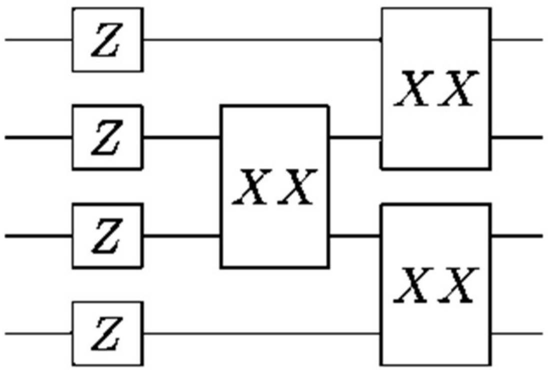
Examples of DLA



Simple local PQC

$$\mathcal{G} = \{Z_j, Y_j\}_{j=1}^{n}$$

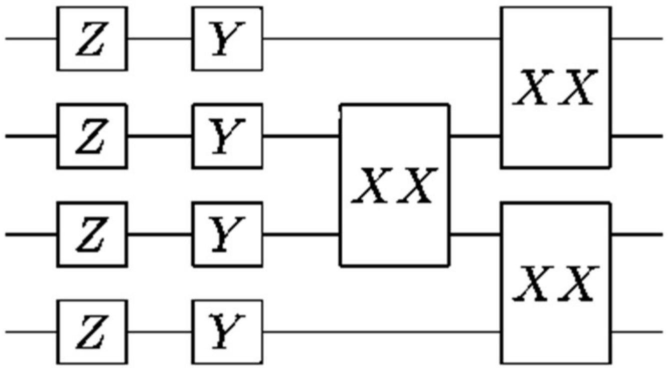
$$\mathfrak{g} = \mathfrak{su}(2)^{\oplus n}$$



Matchgate circuit

$$\mathcal{G} = \{Z_j\}_{j=1}^{n} \cup \{X_j X_{j+1}\}_{j=1}^{n-1}$$

$$\mathfrak{g} = \mathfrak{so}(2n)$$



Universal circuit

$$\mathcal{G} = \{Z_j, Y_j\}_{j=1}^{n} \cup \{X_j X_{j+1}\}_{j=1}^{n-1}$$

$$\mathfrak{g} = \mathfrak{su}(2^n)$$

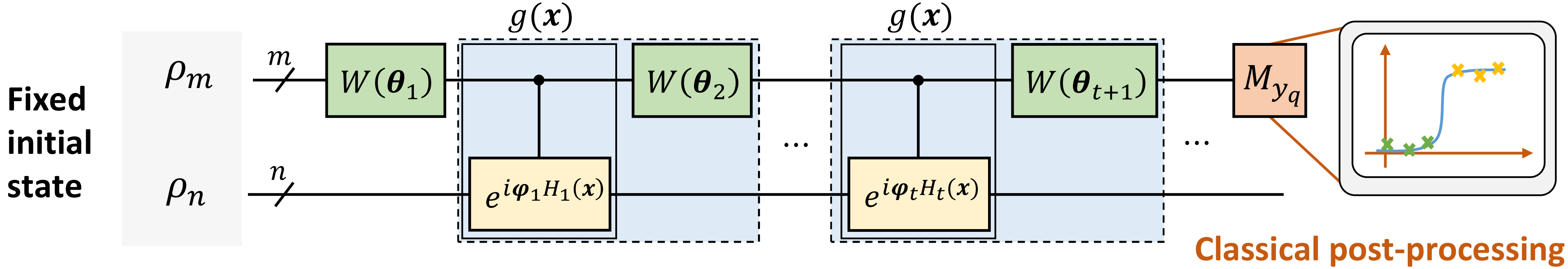
High dimension of $\mathfrak{g}_j$ $\Rightarrow$ High possibility to have BP!

香港科大(广州)
HKUST(GZ)

Quantum Recurrent Embedding Neural Networks

Circuit framework of QRENN

Circuit Model of QRENN



- Assuming $W(\boldsymbol{\theta})$ approximate $SU(2^m)$; Hamiltonian $[H_t(\boldsymbol{x}),\ H_\tau(\boldsymbol{x})] = 0$, for any $t \neq \tau$

- The DLA of QRENN can be decomposed into

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

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

The intersections of the eigenspaces of H_t and H_τ can be decomposed into direct sum

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

Quantum supervised learning

香港科大(广州) HKUST(GZ)

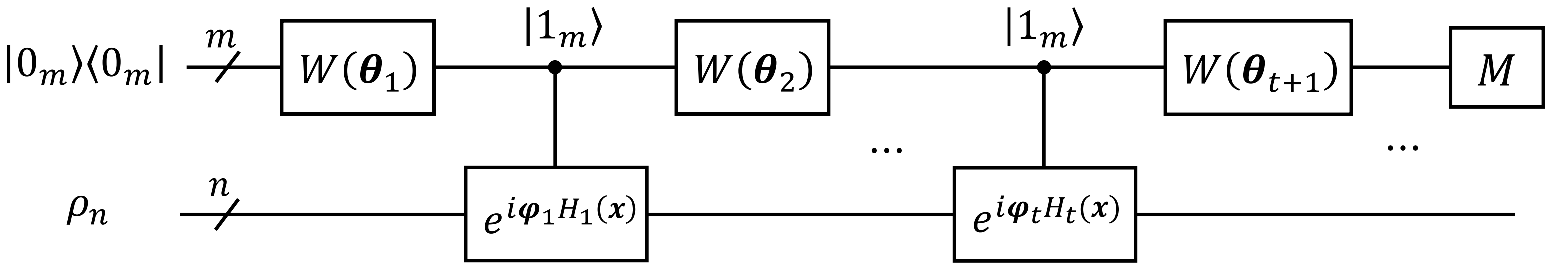
- Information about a quantum system is stored in its Hamiltonian
- Given a batch training set $\mathcal{T} = \{(y_q, X_q)\}_q$. Originally use MSE loss:

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

- **Hard to analyse gradient**, through experimental no BP.
- Inspired from hypothesis testing, design $M_1, M_2, \cdots M_k$ forming POVM. We define

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

Main theorem on trainability



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

- For sufficiently deep QRENN (scales $O\big(poly(n)\big)$)[3], the circuit achieve 2-design of the compact Lie group and, hence, $\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\varphi}}\left[\partial_{t,\mu}\mathcal{L}\right] = 0$ [2].

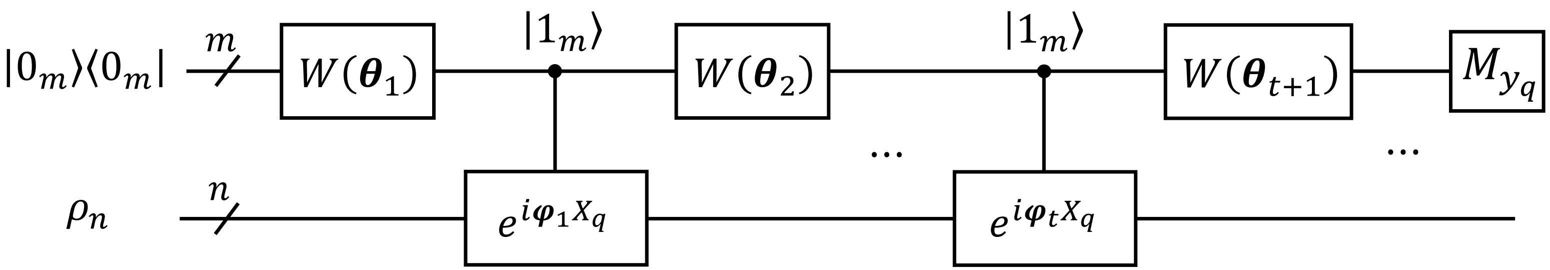
Theorem

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

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

[2] Fontana, Enrico, et al. "The Adjoint Is All You Need: Characterizing Barren Plateaus in Quantum Ans¥" atze." *Nature Communication* (2023).
[3] Ragone, Michael, et al. "A Lie algebraic theory of barren plateaus for deep parameterized quantum circuits." *Nature Communications* (2024)

Sketch of proof



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

- Taking the derivative to the loss function:

$$\partial_\Omega \mathcal{L} = -\frac{1}{Q}\sum_{q=1}^{Q} \mathrm{Tr}\left(U_{g^-}^\dagger(\boldsymbol{x}_q) i\rho_0 U_{g^-}(\boldsymbol{x}_q)[\Omega, U_{g^+}(\boldsymbol{x}_q) iM_{y_q} U_{g^+}^\dagger(\boldsymbol{x}_q)]\right)$$

- Averaging over the group ⇒ $\mathbb{E}_{g^\pm\sim\mu^{\otimes 2}}[\partial_\Omega \mathcal{L}] = 0,$ ⇒ $\mathrm{Var}[\partial_\Omega \mathcal{L}] = \mathbb{E}_{g^\pm\sim\mu^{\otimes 2}}[(\partial_\Omega \mathcal{L})^2].$

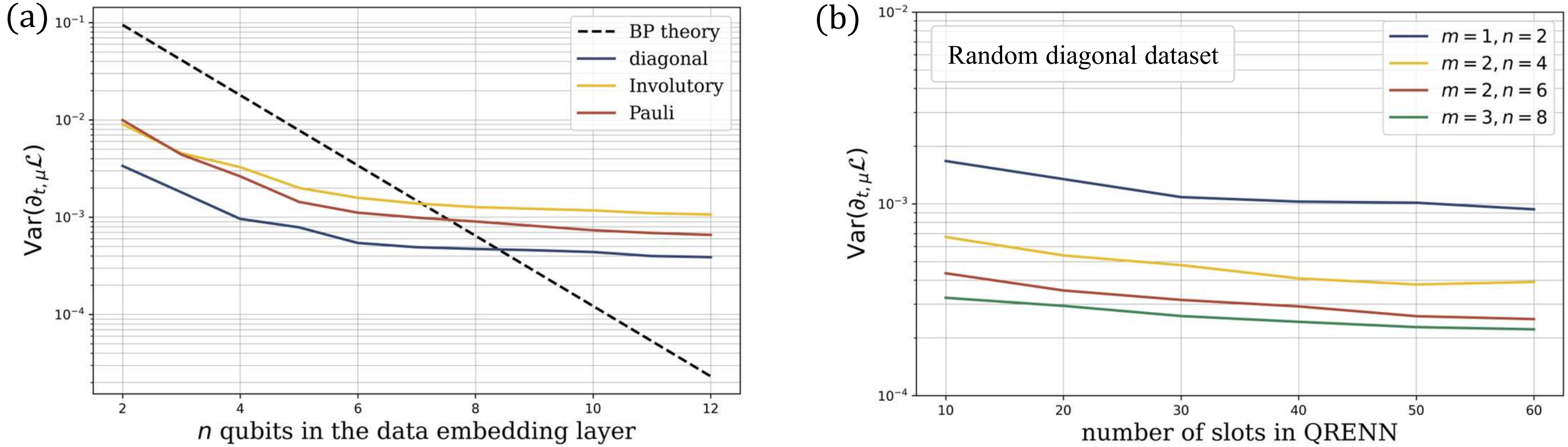
- Use the property of split Casimir operator: $\int_G U_g^{\otimes 2}(A\otimes B)(U_g^\dagger)^{\otimes 2} dg = \sum_\alpha \frac{\mathrm{Tr}(A_{\mathfrak{g}_\alpha} B_{\mathfrak{g}_\alpha})}{d_{\mathfrak{g}_\alpha}} K_{\mathfrak{g}_\alpha} + A_{\mathfrak{c}} \otimes B_{\mathfrak{c}},$

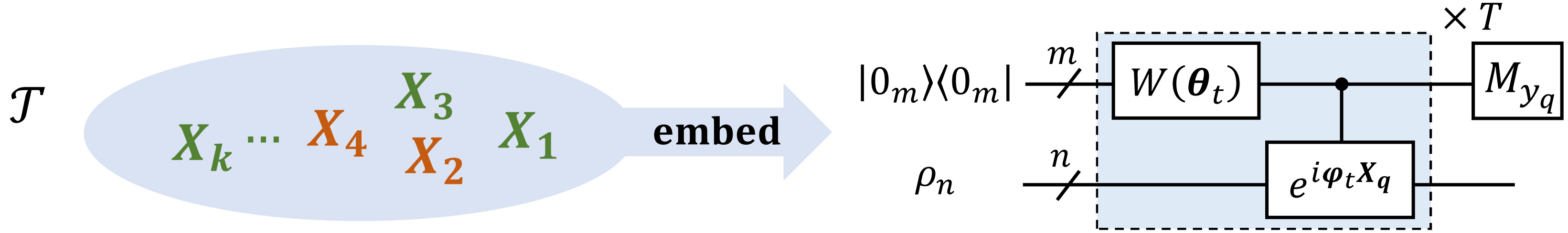
⇒ $$\mathbb{E}_{g^\pm\sim\mu^{\otimes 2}}[(\partial_\Omega \mathcal{L})^2] \geq \frac{1}{Q^2}\sum_q \sum_{\boldsymbol{\lambda}_q} \frac{\left\|(M_{y_q})_{\mathfrak{g}_{\boldsymbol{\lambda}_q}}\right\|_F^2 \left\|\rho_{\mathfrak{g}_{\boldsymbol{\lambda}_q}}\right\|_F^2 \left\|\Omega_{\mathfrak{g}_{\boldsymbol{\lambda}_q}}\right\|_K^2}{d^2_{\mathfrak{g}_{\boldsymbol{\lambda}_q}}} \geq \frac{1}{Q^2}\frac{2^{m+1}\|\Omega_\mu\|_F^2}{(2^{2m}-1)^2}\sum_q \left\|M_{y_q}\right\|_F^2 R^2_{X_q}(\rho_n).$$

$$\left\|\frac{I_{2^m} \pm Z^{\otimes m}}{2}\right\|_F^2 = 2^{m-1} \in \Omega(1/\mathrm{poly}(n)),$$

□

Numerical results on trainability





- Gradient sampling experiments, 500 random initial parameters $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ of the model, ρ_n being fixed.

- We have tested three datasets namely Diagonal, Involutory and Pauli sets. For each dataset, **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)

- Information about a quantum system is stored in its Hamiltonian
- Given a batch training set $\mathcal{T} = \{(y_q, H_q)\}_q$. Inspired from hypothesis testing, design $\{M_1, M_2, \cdots M_k\}$ forming POVMs. We define

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

Problem

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

Given a cluster-Ising model with periodic boundary conditions

$$H(\lambda) = -\sum_{j=1}^{N} X_{j-1}Z_jX_{j+1} + \lambda \sum_{j=1}^{N} Y_jY_{j+1}.$$

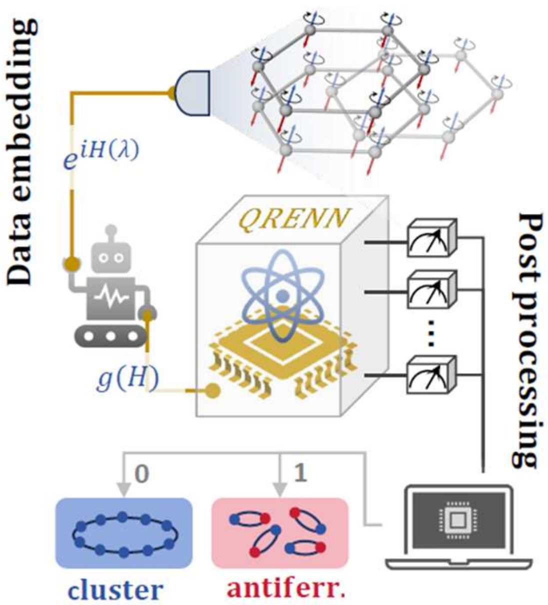
where X, Y and Z are Pauli matrices. SPT phase in the Hamiltonian model [4]:

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

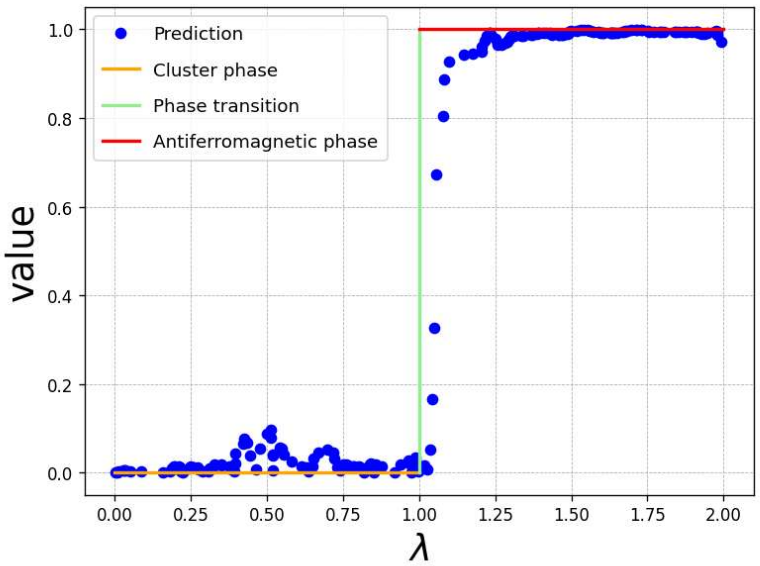
[4] Li, Weikang, Zhi-de Lu, and Dong-Ling Deng. "Quantum neural network classifiers: A tutorial." *SciPost Physics Lecture Notes* (2022): 061.

SPT phase detection

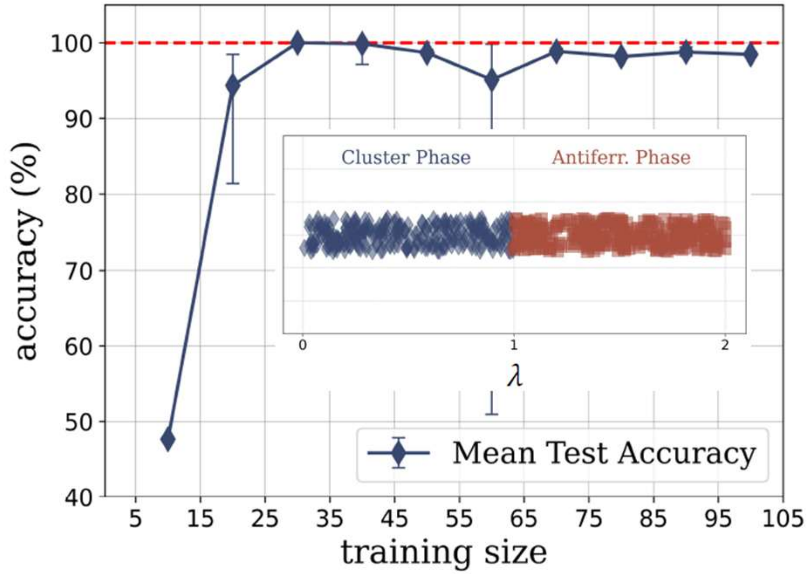
- **QRENN model in learning SPT phase**



Train:



Prediction:



Case 1: $\mathrm{slots} = 10$, $m = 1$ and $n = 8$, initial state $|0\rangle \otimes |+\rangle^{\otimes 8}$. **Outcome**: Training 40 data uniformly generated by sampling $\lambda \in [0,2]$. Achieve 92.32% accuracy on 560 testing data.

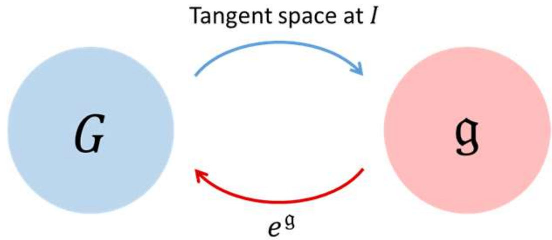
Case 2: $\mathrm{slots} = 10$, $m = 1$ and $n = 8$, initial state $|0\rangle \otimes |+\rangle^{\otimes 8}$. **Outcome**: Train with different data sizes. Find an improvement in performance as training size increases.

Concluding remarks

Preliminary | QRENN | Concluding remarks

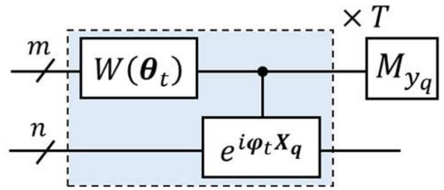
Conclusion

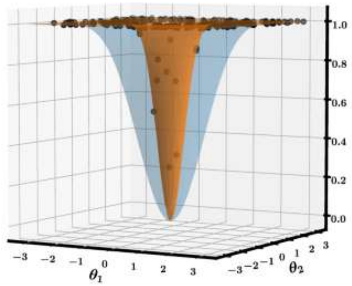
- **Recent developments in the Quantum Machine Learning**
  - From Lie algebra to quantum neural networks
  - Dynamical Lie algebra for barren plateaus

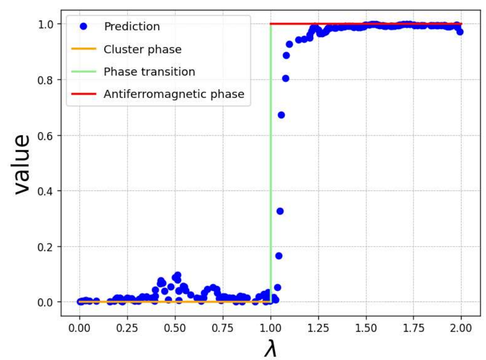


$$\mathfrak{g} = \text{span}_{\mathbb{R}}\langle iH_1, iH_2, \cdots, iH_L\rangle_{Lie}$$

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







~Thanks for watching~

QUAIR Group

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