Approximately Equivariant QNN for p4m Group Symmetry in Images

Quantum Techniques in Machine Learning 2023

Su Yeon Chang [EPFL, CERN]
Bertrand Le Saux [ESA], Sofia Vallecorsa [CERN], Michele Grossi [CERN]

21 November 2023

Exploiting symmetry in variational quantum machine learning

Johannes Jakob Meyer,[1] Marian Mularski,[1,2] Elies Gil-Fuster,[1,3] Antonio Anna Mele,[1] Francesco Arzani,[1] Alissa Wilms,[1,2] and Jens Eisert[1,4,3]

[1] *Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2] *Porsche Digital GmbH, 71636 Ludwigsburg, Germany*
[3] *Fraunhofer Heinrich Hertz Institute, 10587 Berlin, Germany*
[4] *Helmholtz-Zentrum Berlin für Materialien und Energie, 14109 Berlin, Germany*

(Dated: May 12, 2022)

Variational quantum machine learning is an extensively studied application of near-term quantum computers. The success of variational quantum learning models crucially depends on finding a suitable parametrization of the model that encodes an inductive bias relevant to the learning task. However, precious little is known about guiding principles for the construction of suitable parametrizations. In this work, we holistically explore when and how symmetries of the learning problem can be exploited to construct quantum learning models with outcomes invariant under the symmetry of the learning task. Building on tools from representation theory, we show how a standard gateset can be transformed into an equivariant gateset that respects the symmetries of the problem at hand through a process of gate symmetrization. We benchmark the proposed methods on two toy problems that feature a non-trivial symmetry and observe a substantial increase in generalization performance. As our tools can also be applied in a straightforward way to other variational problems with symmetric structure, we show how equivariant gatesets can be used in variational quantum eigensolvers.

Theoretical Guarantees for Permutation-Equivariant Quantum Neural Networks

Louis Schatzki,[1,2,*] Martín Larocca,[3,4,†] Quynh T. Nguyen,[3,5] Frédéric Sauvage,[3] and M. Cerezo[1,‡]

[1] *Information Sciences, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[2] *Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, Illinois, 61801, USA*[†]
[3] *Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[4] *Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[5] *Harvard Quantum Initiative, Harvard University, Cambridge, Massachusetts 02138, USA*

Despite the great promise of quantum machine learning models, there are several challenges one must overcome before unlocking their full potential. For instance, models based on quantum neural networks (QNNs) can suffer from excessive local minima and barren plateaus in their training landscapes. Recently, the nascent field of geometric quantum machine learning (GQML) has emerged as a potential solution to some of those issues. The key insight of GQML is that one should design architectures, such as equivariant QNNs, encoding the symmetries of the problem at hand. Here, we focus on problems with permutation symmetry (i.e., the group of symmetry S_n), and show how to build S_n-equivariant QNNs. We provide an analytical study of their performance, proving that they do not suffer from barren plateaus, quickly reach overparametrization, and can generalize well from small amounts of data. To verify our results, we perform numerical simulations for a graph state classification task. Our work provides the first theoretical guarantees for equivariant QNNs, thus indicating the extreme power and potential of GQML.

earning task / improve generalization

Theory for Equivariant Quantum Neural Networks

Quynh T. Nguyen,[1,2,*] Louis Schatzki,[3,4,*] Paolo Braccia,[1,5] Michael Ragone,[1,6] Patrick J. Coles,[1] Frédéric Sauvage,[1] Martín Larocca,[1,7] and M. Cerezo[3]

[1] *Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[2] *Harvard Quantum Initiative, Harvard University, Cambridge, Massachusetts 02138, USA*
[3] *Information Sciences, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[4] *Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[5] *Dipartimento di Fisica e Astronomia, Università di Firenze, Sesto Fiorentino (FI), 50019 , Italy*
[6] *Department of Mathematics, University of California Davis, Davis, California 95616, USA*
[7] *Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

Most currently used quantum neural network architectures have little-to-no inductive biases, leading to trainability and generalization issues. Inspired by a similar problem, recent breakthroughs in classical machine learning address this crux by creating models encoding the symmetries of the learning task. This is materialized through the usage of equivariant neural networks whose action commutes with that of the symmetry. In this work, we import these ideas to the quantum realm by presenting a general theoretical framework to understand, classify, design and implement equivariant quantum neural networks. As a special implementation, we show how standard quantum convolutional neural networks (QCNN) can be generalized to group-equivariant QCNNs where both the convolutional and pooling layers are equivariant under the relevant symmetry group. Our framework can be readily applied to virtually all areas of quantum machine learning, and provides hope to alleviate central challenges such as barren plateaus, poor local minima, and sample complexity.

- GQML : **Geometric Qua**

→ Leverage the ideas of GI … inductive bias

- E.g. Classifying connecte … ınder the permutation

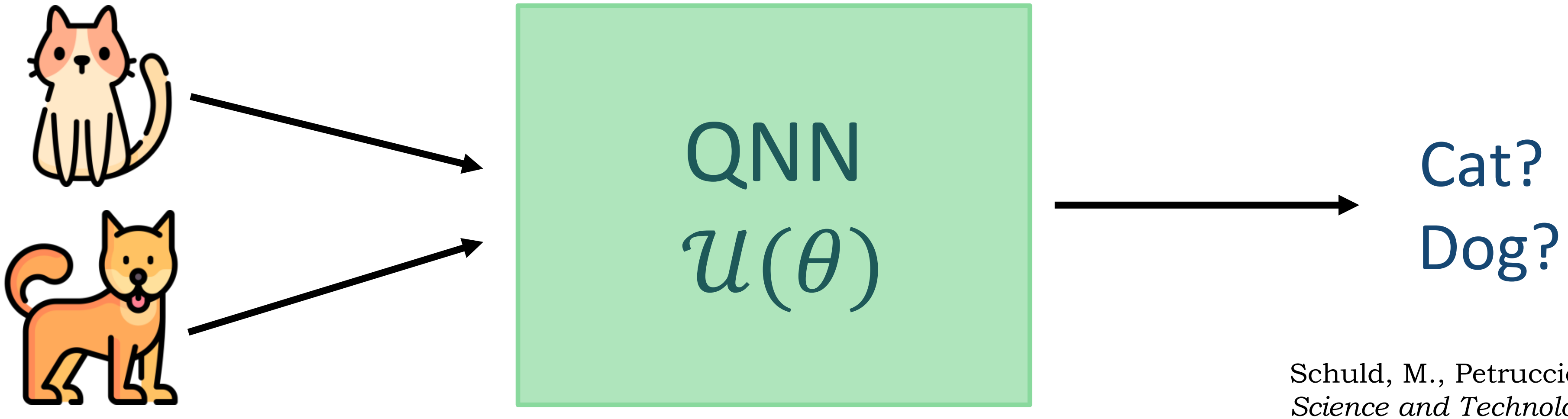
→ **Construct quant … image classification**

Supervised Quantum Machine Learning

- Classical data $x \in \mathcal{X}$ associated with label $\ell \in \mathcal{Y}$ following underlying function $f : \mathcal{X} \rightarrow \mathcal{Y}$
- Supervised learning : Find y_θ which is as close as possible to f
- Final prediction y_θ for an input feature with the observable O

$$y_\theta(x) = \langle \psi(x)| \mathcal{U}^\dagger(\theta) O\, \mathcal{U}(\theta)\, |\psi(x) \rangle$$

- Models agnostic to the underlying symmetry of $\mathcal{X}$



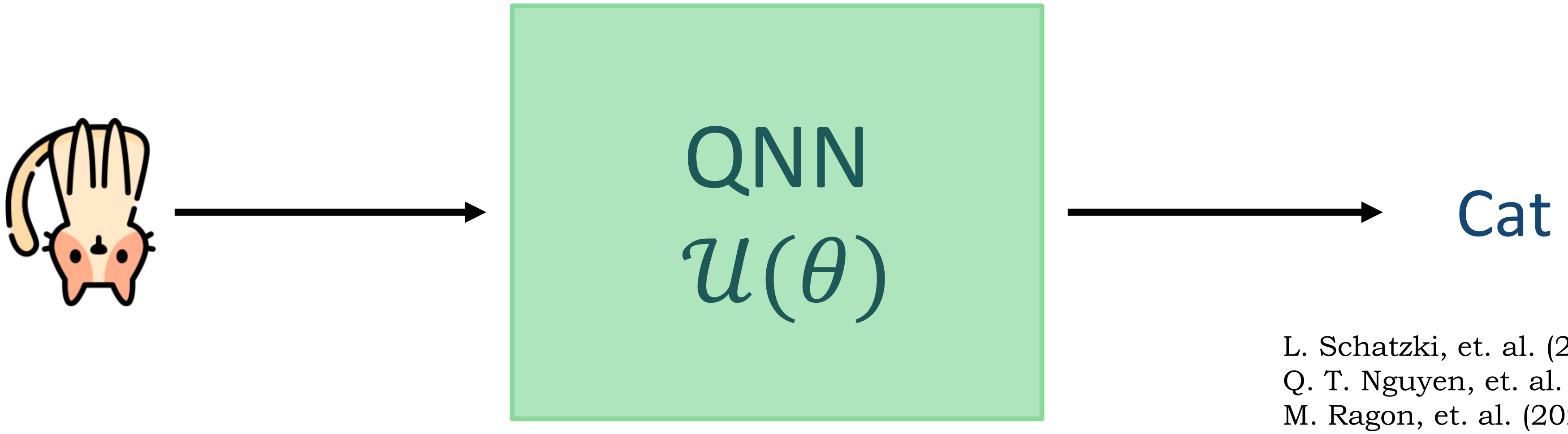
Schuld, M., Petruccione, F. (2021). *Quantum Science and Technology*. Springer, Cham.

Geometric Quantum Machine Learning

- Consider a symmetry group $\mathfrak{G}$ on the data space $\mathcal{X}$
- **$\mathfrak{G}$ – Invariance** : For all $x \in \mathcal{X}$ and $g \in \mathfrak{G}$

$$\boxed{y_\theta(g[x]) = y_\theta(x)}$$

→ Require **1)** Equivariant data embedding **2)** Equivariant QNN **3)** Invariant Measurement



L. Schatzki, et. al. (2022). arXiv:2210.09974
Q. T. Nguyen, et. al. (2022). arXiv:2210.08566
M. Ragon, et. al. (2022). arXiv:2210.07980

Symmetries induced by data embedding

- Quantum feature map $\psi: \mathcal{X} \to \mathcal{H}$

→ Embed classical data into quantum state

- $\mathfrak{G}$-Equivariant embedding : Induces a unitary quantum action $\boldsymbol{V_s[g]}$

$$|\psi(g[x])\rangle = V_s[g]|\psi(x)\rangle$$

- $V_s[g]$ = **Representation** of g on $\mathcal{H}$ induced by ψ

Equivariant ansatz

- Focus on gates generated by a fixed generator G

$$R_G(\theta) = \exp(-i\theta G)$$

- **$\mathfrak{G}$ – equivariant gate** : For all $g \in \mathfrak{G}$ and $\theta \in \mathbb{R}$

$$\boxed{[R_G(\theta), V_s[g]] = 0 \ \leftrightarrow [G, V_s[g]] = 0}$$

→ $\mathcal{U}(\theta)$ is $\mathfrak{G}$ – equivariant if and only if it consists of equivariant quantum operators

Twirling method :

- Arbitrary generator A → Construct a projector onto all symmetry group element
- *Twirling operator* : $\mathcal{T}_{\mathfrak{G}}[A] = \frac{1}{|\mathfrak{G}|}\sum_{g\in\mathfrak{G}} V_s[g]^\dagger A V_s[g]$

→ $\left[\mathcal{T}_{\mathfrak{G}}[A], V_s[g]\right] = 0$ for any $g \in G$

Invariant Measurement

- **$\mathfrak{G}$ – Invariant Observable** : For all $g \in \mathfrak{G}$

$$\boxed{V_s^\dagger[g] O V_s[g] = O}$$

Invariance from equivariance

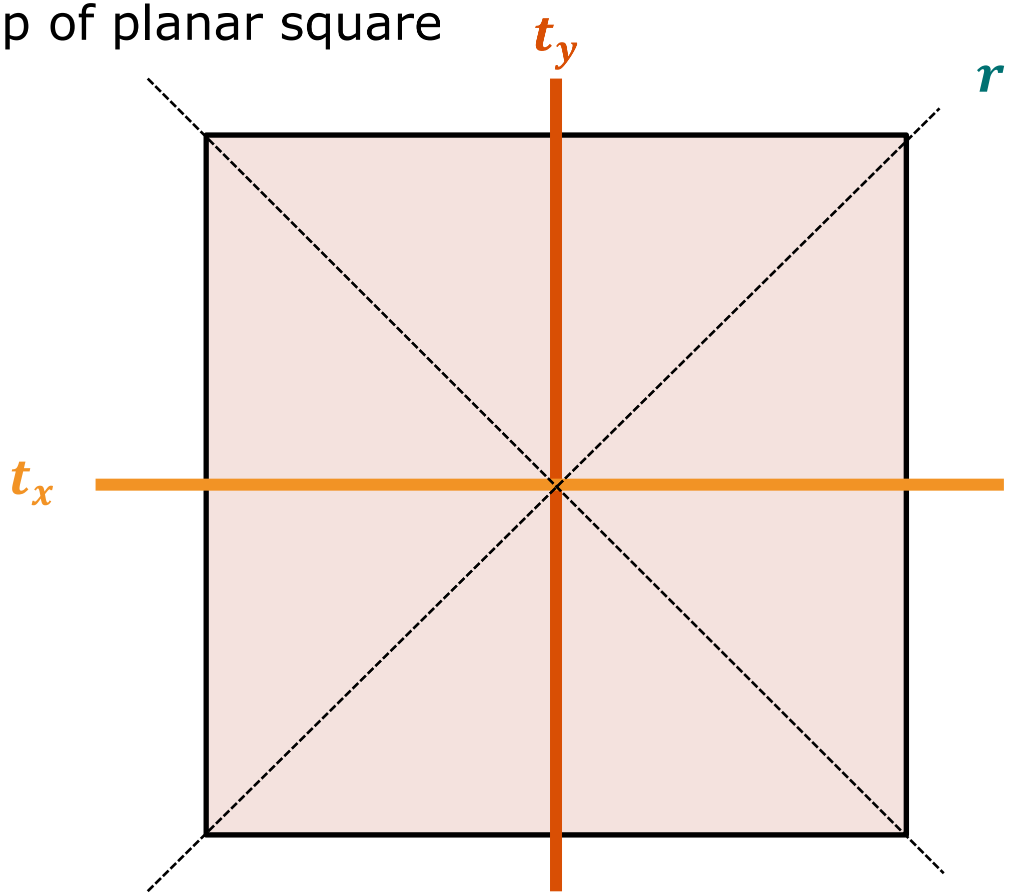
- The final prediction $y_\theta(x)$ is $\mathfrak{G}$ -invariant if it consists of a $\mathfrak{G}$ -equivariant QNN and measurement.

$$\begin{aligned}
y_\theta(g[x]) &= \langle \psi(g[x]) | \mathcal{U}^\dagger(\theta) O\, \mathcal{U}(\theta) | \psi(g[x]) \rangle \\
&= \left\langle \psi(x) \middle| V_s^\dagger \mathcal{U}^\dagger(\theta) O\, \mathcal{U}(\theta) V_s \middle| \psi(x) \right\rangle \\
&= \left\langle \psi(x) \middle| \mathcal{U}^\dagger(\theta) V_s^\dagger O\, V_s \mathcal{U}(\theta) \middle| \psi(x) \right\rangle \\
&= \langle \psi(x) | \mathcal{U}^\dagger(\theta) O\, \mathcal{U}(\theta) | \psi(x) \rangle = y_\theta(x)
\end{aligned}$$

→ Model respects the label symmetries

p4m symmetry

- Plane (Wallpaper) symmetry group : Symmetry group of planar square
- Consists of 8 components
  - Identity e
  - **Rotation** r, r^2, r^3 of $90°, 180°, 270°$ around origin
  - **Reflection** t_x, t_y in the x, y axis
  - Reflection in the two diagonal

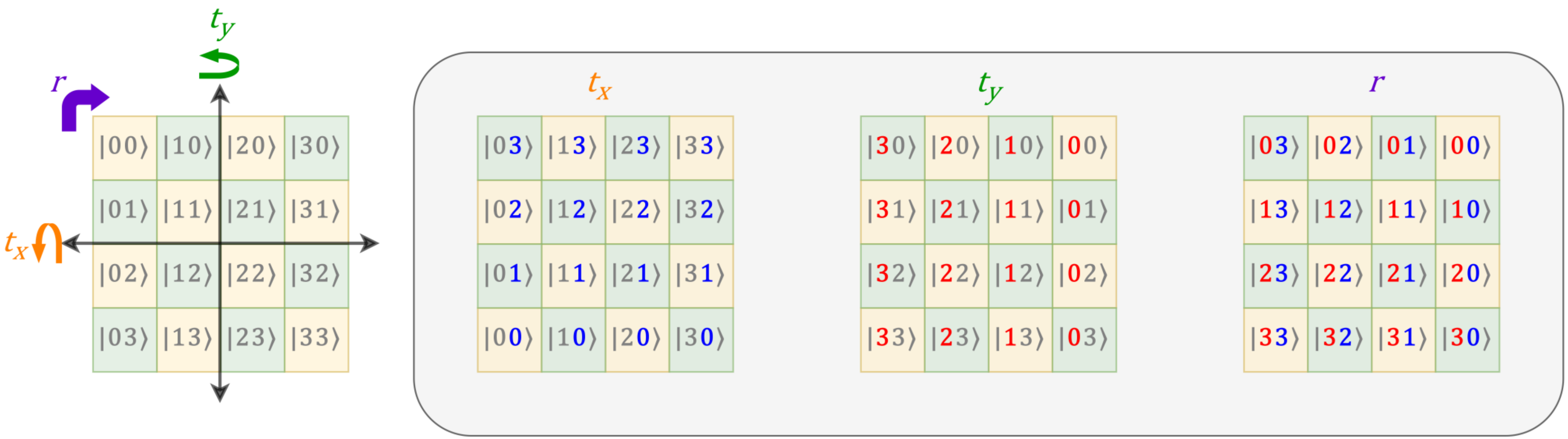


Quantum Embedding for p4m symmetry

<u>Coordinate-Aware Amplitude Embedding method</u>

- Consider a classical image of size $\boldsymbol{L} \times \boldsymbol{L}$, $\mathbf{x} = \{\mathbf{x}_{00}, \mathbf{x}_{01}, \ldots, \mathbf{x}_{L-1L-1}\}$.
- Explicitly denote the coordinates of each pixels with $2n$ qubits, $n = \log_2 L$:

$$|\psi(x)\rangle = \frac{1}{\mathcal{N}} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} x_{ij} \underset{q_{0:n}}{|\boldsymbol{i}\rangle} \underset{q_{n:2n}}{|\boldsymbol{j}\rangle}$$



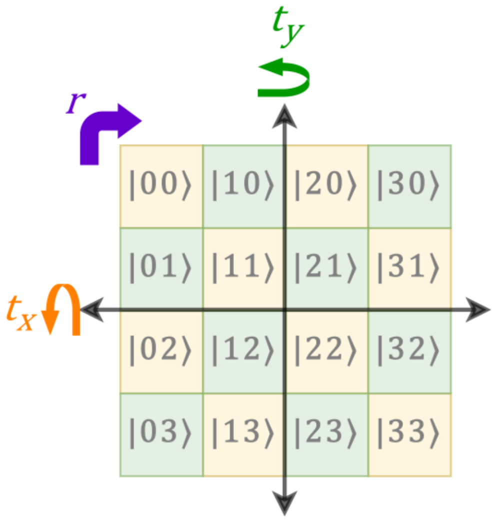
Induced representation of CAA embedding

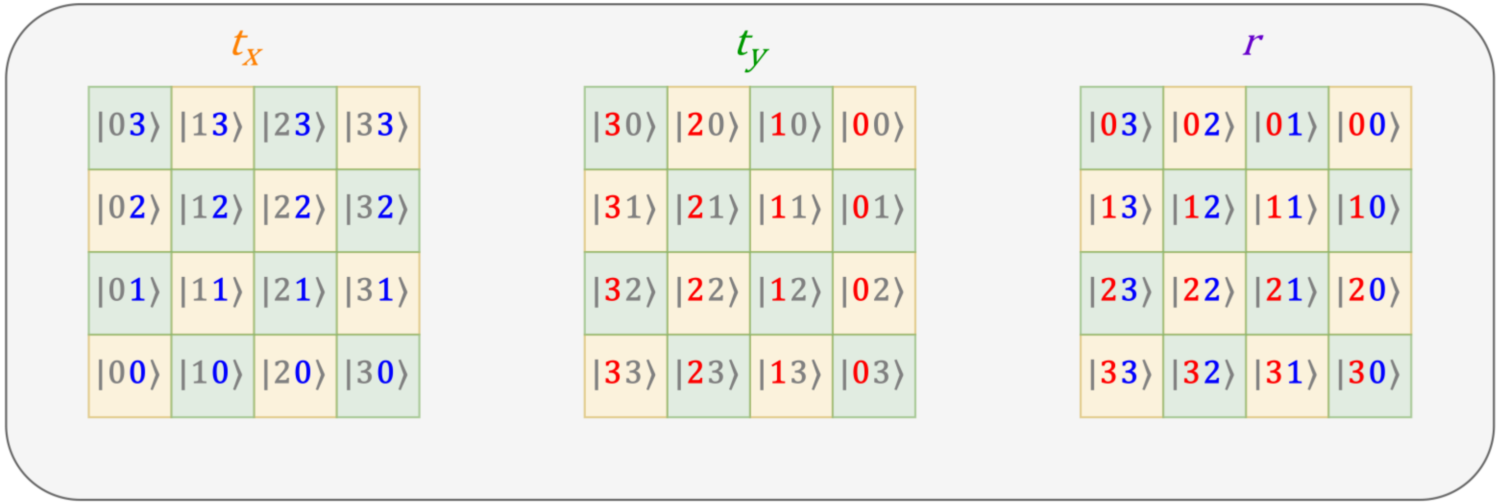
- Induced Symmetry representation

$$V_x = \mathbb{I}_n \otimes X^{\otimes n}, = X_{n:2n}, \qquad V_y = X^{\otimes n} \otimes \mathbb{I}_n = X_{1:n}$$
$$V_r = \left(\bigotimes_{i=1}^{n} SWAP_{i,i+n}\right)(X^{\otimes n} \otimes \mathbb{I}_n) = V_r{}'V_y$$

→ Find operators that commute with V_x, V_y, V_r

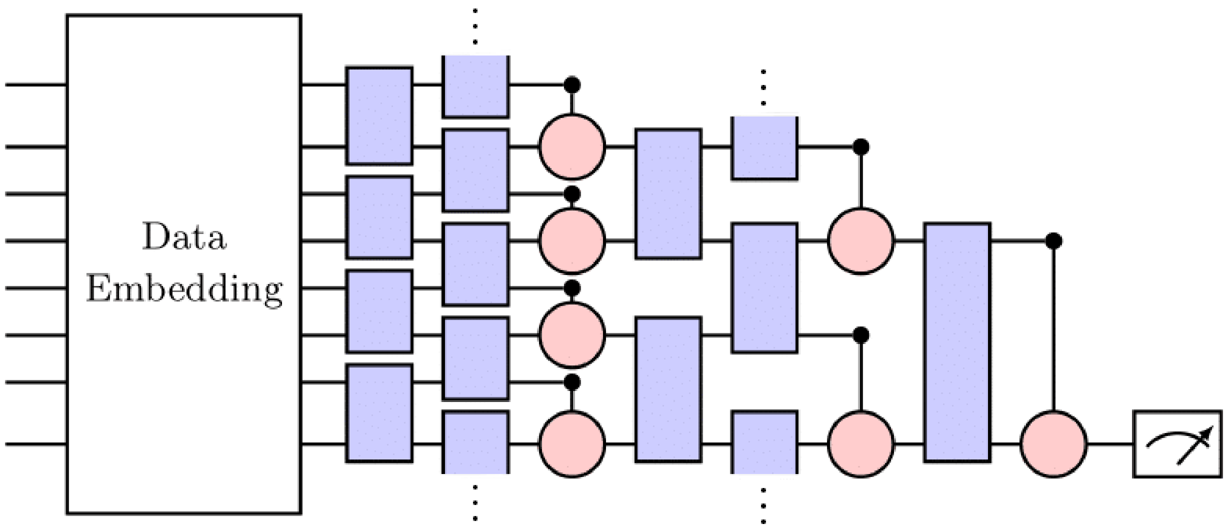
$$\boxed{U \in \mathrm{comm}\left\{X_{1:n}, X_{n:2n}, \bigotimes_{i=1}^{n} SWAP_{i,i+n}\right\}}$$





Quantum Convolutional Neural Network

- Quantum analogue of classical convolutional neural network
- Consists of : **1)** Convolutional filters **2)** Pooling layers → Hierarchical architecture
- (Partial) Translational Invariance : Identical parameters for all the filters in each layer
- Allows to avoid barren plateau



Cong, I., Choi, S. & Lukin, M.D. *Nat. Phys.* **15**, 1273–1278 (2019)
T. Hur et. al. (2022). *Quantum Mach. Intell.* **4**, 3

Equivariant Quantum Ansatzs

- For V_r' : Repeat the same parameter on q_i and q_{i+n} for even n
- For V_x, V_y: Consider <u>two different cases</u>

1) Gates constrained to $q_{1:n}$ **<u>OR</u>** $q_{n:2n}$
- Consider 2-body quantum gate → Quantum gates that commute with $X \otimes X$

$$X_1, X_2, X_1X_2, Y_1Y_2, Z_1Z_2 \in \mathfrak{comm}(X \otimes X)$$

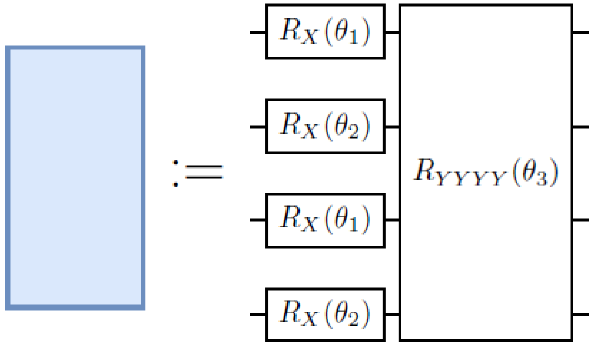
→ Equivariant gate set $G_{s,1} = \{X_1, X_2, Y_1Y_2, Z_1Z_2\}$

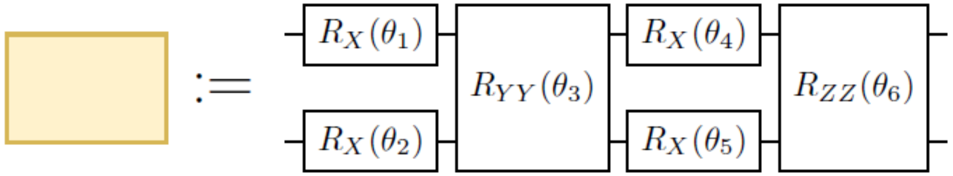
2) Gates applied to $q_{1:n}$ **<u>AND</u>** $q_{n:2n}$
- Quantum gates that commute with $X^n \otimes \mathbb{I}^n$ → Unbiased weight of P_σ applied on $q_{1:n}$ and $q_{n:2n}$
- Even number of Pauli Y and Pauli Z gates applied on $q_{1:n}$ and $q_{n:2n}$
- Requires at least 4 qubit gates → Search quantum gates that **<u>commute with $XX\mathbb{I}\mathbb{I}$</u>**

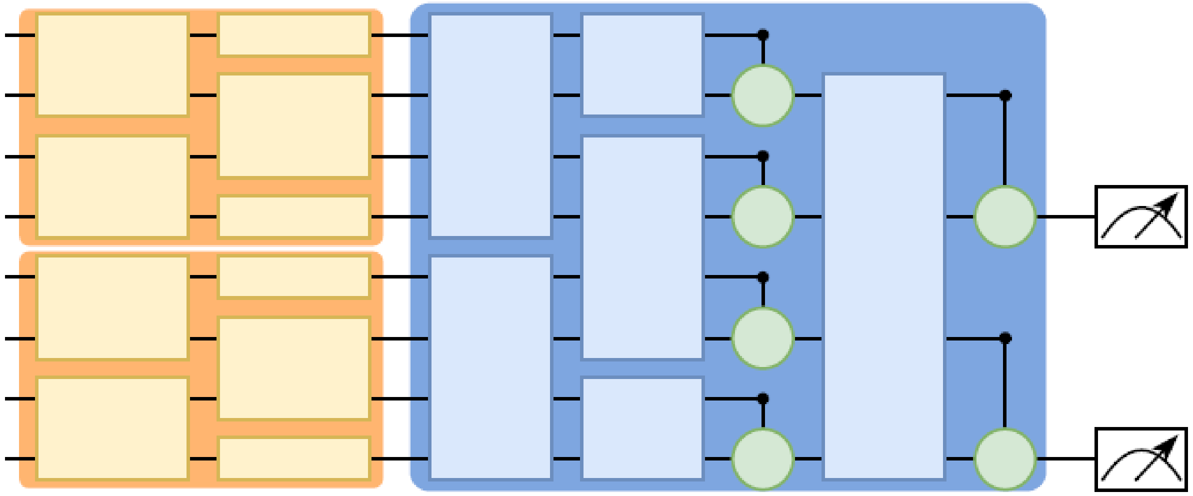
→ Equivariant gate set $G_{s,2} = \{P_\sigma P_\sigma P_{\sigma'} P_{\sigma'} | P_{\sigma,\sigma'} \in \{X, Y, Z\}\}$

Equivariant Quantum CNN

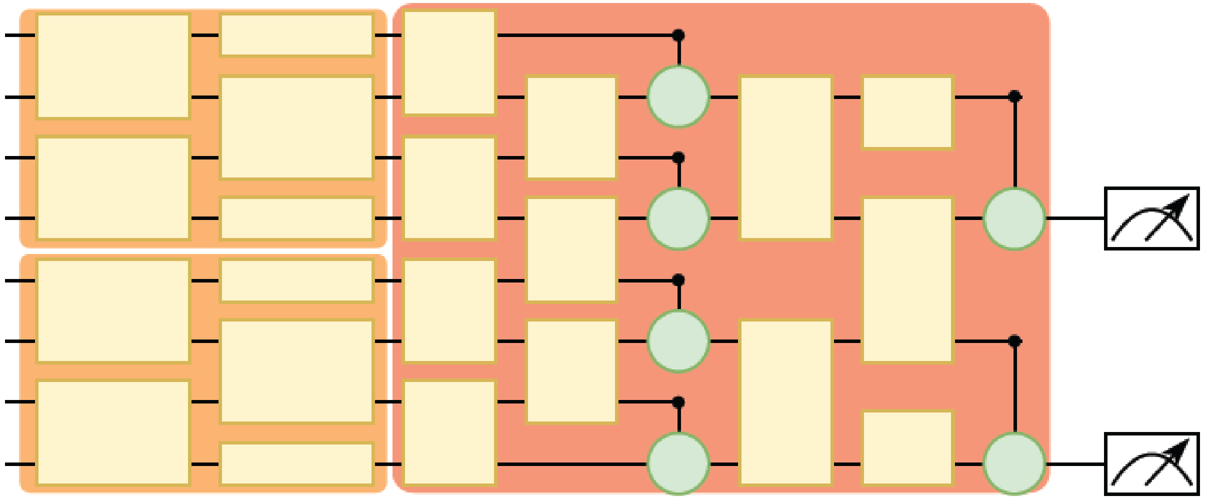
Low angle assumption
Label misassignment error ~ 2%





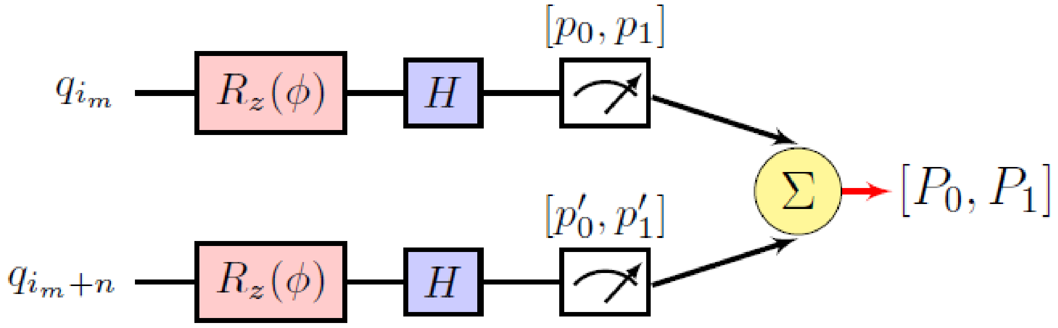


EquivQCNN



ApprEquivQCNN

Approximately Invariant Measurement



M_1: $\phi = 0$

M_2: $\phi \neq 0$

- Final quantum state : $|\psi_f\rangle = r_0 e^{i\theta_0}|00\rangle + r_1 e^{i\theta_1}|01\rangle + r_2 e^{i\theta_2}|10\rangle + r_3 e^{i\theta_3}|11\rangle$
- For $V_r{}'$: Invariance guaranteed by $p_0 + p_0'$
- For V_x : Denote P_0^x for the output of reflected image

Low angle assumption
Label misassignment error 2~6%

$$P_0 - P_0^x = \sin(2\phi)[r_0 r_1 \sin(\theta_1 - \theta_0) + r_2 r_3 \sin(\theta_3 - \theta_2)]$$
$$\leq \frac{1}{2}\boldsymbol{\sin(2\phi)}\,[\sin(\theta_1 - \theta_0) + \sin(\theta_3 - \theta_2)]$$

➡ Approximately invariant with an error of ϵ

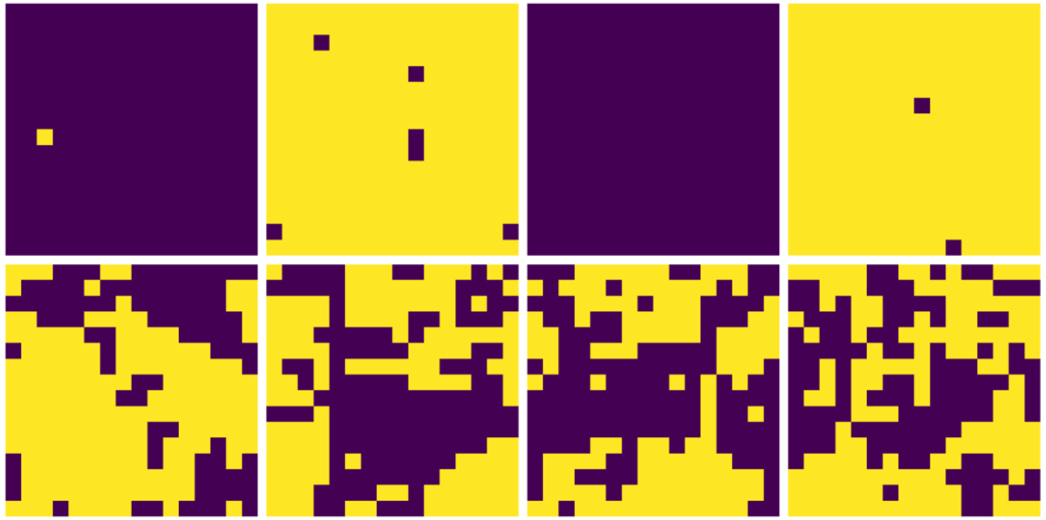
Application

- Test EquivQCNN on two datasets with $L = 16$ → **8 qubits**

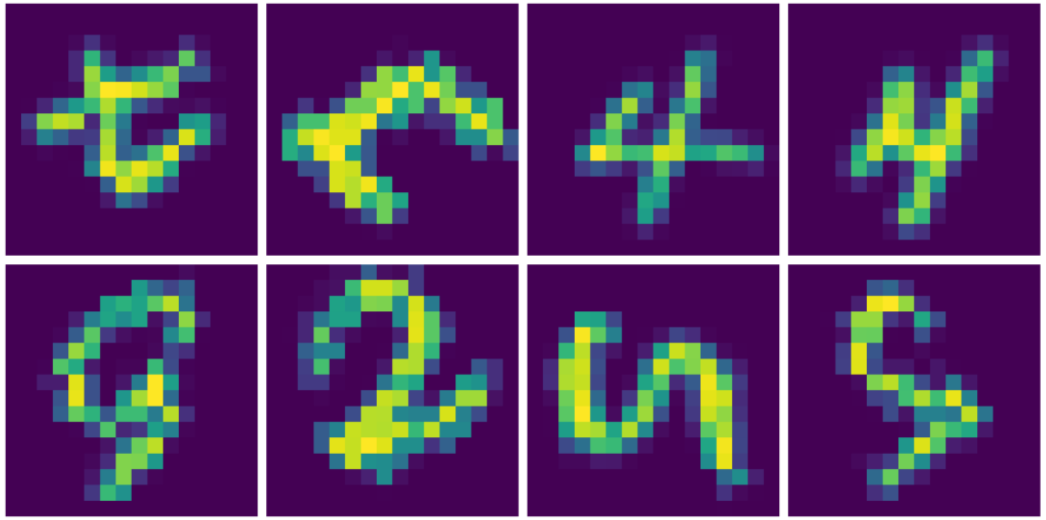
1) Spin distribution in Ising model under the Hamiltonian $\mathcal{H} = -J\sum_{\langle ij \rangle} \sigma_i \sigma_j$

→ Phase detection (Ordered / Disordered phase)

2) Extended MNIST dataset → Augmented with rotated & reflected images



Spin distribution in Ising model



Extended MNIST

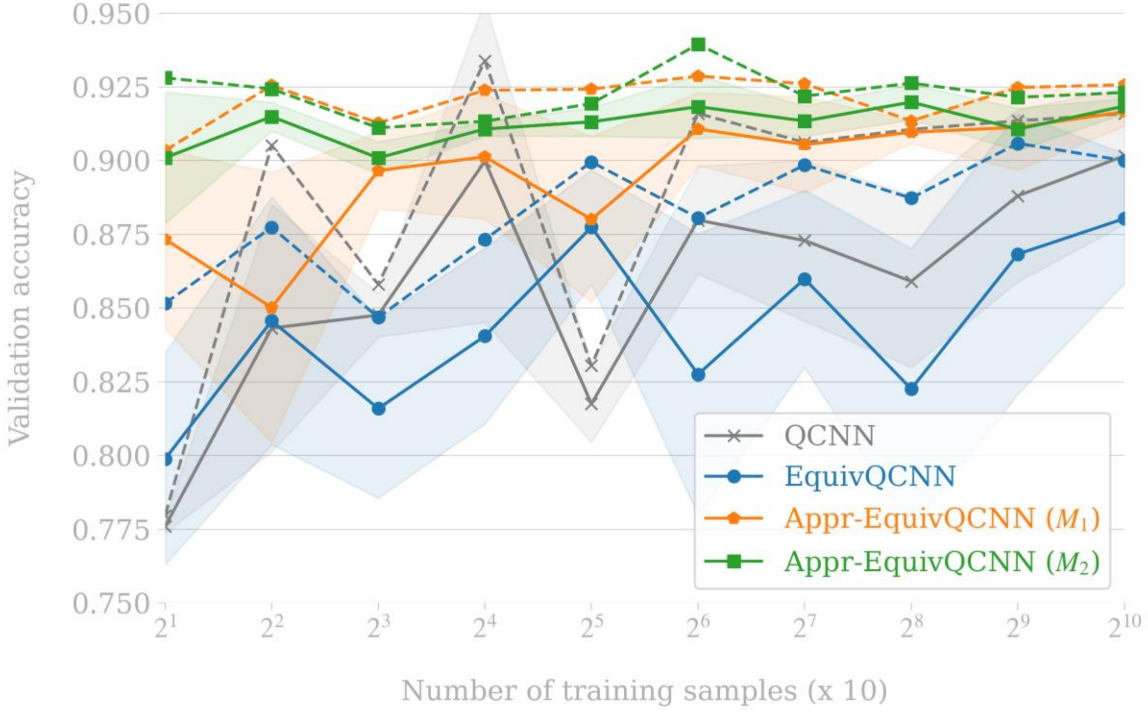
Generalization power of EQCNN

- Measure generalization power with test accuracy for non-equivariant & equivariant QCNN

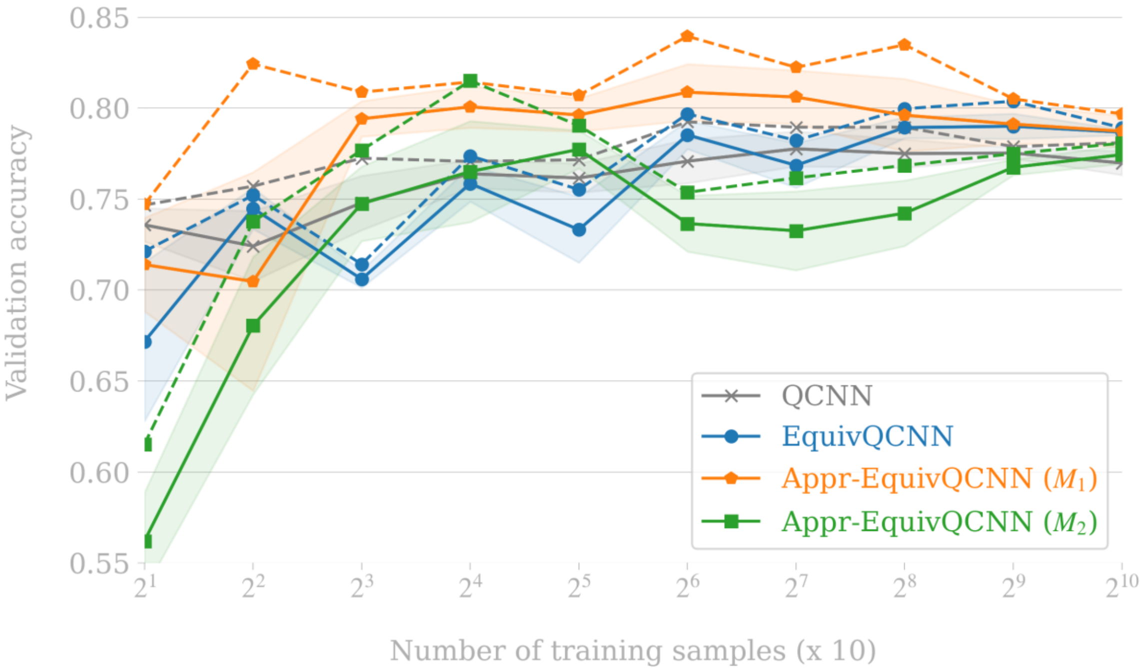
→ Similar number (23 - 25) of parameters for both models

→ Different number of training samples $N = 20, 40, 80, \ldots, 10240$

Ising model

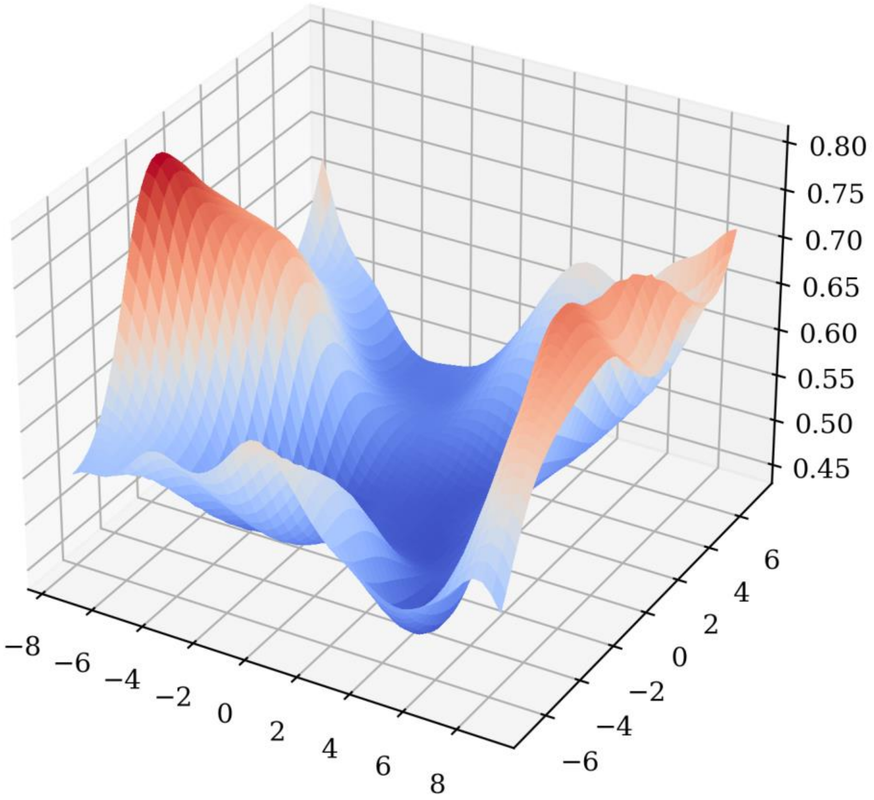


MNIST

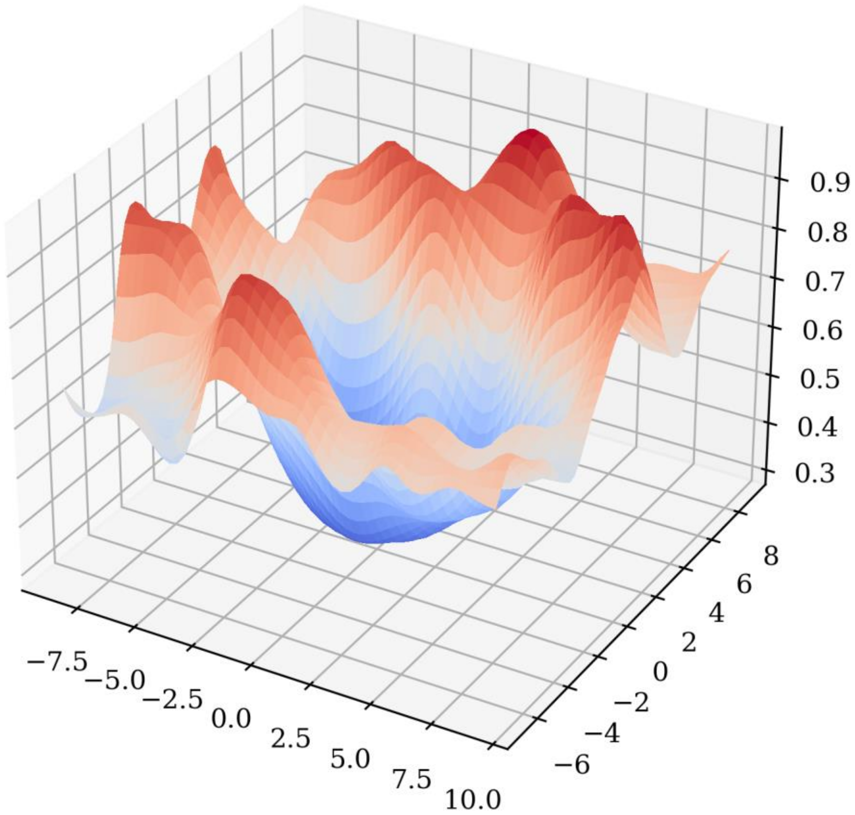


Non-convexity of loss landscape

- Equivariance → Smoother and more convex loss landscape
- Loss landscape plotted for Ising model phase detection



Non-equivariant QCNN

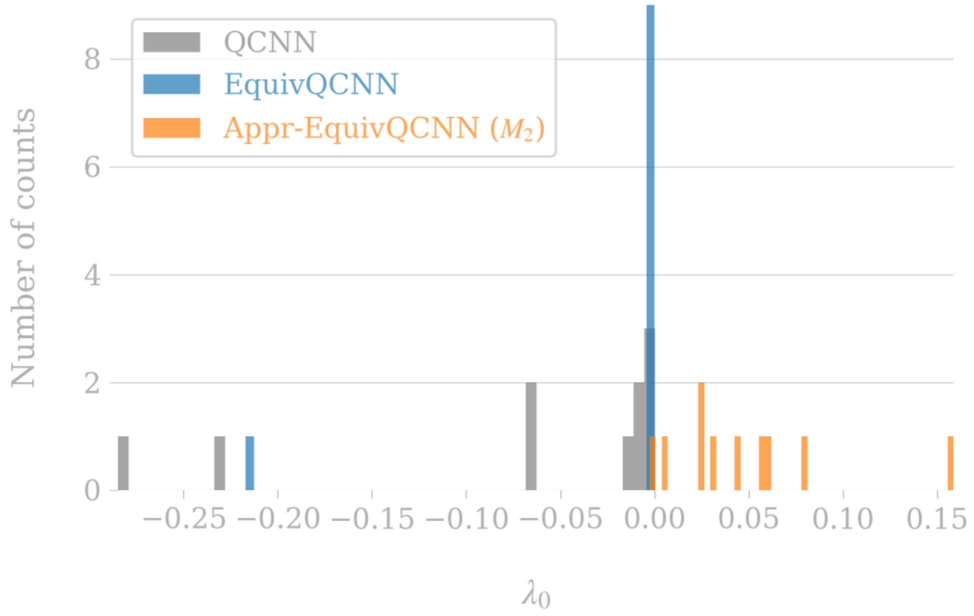


ApprEquivQCNN

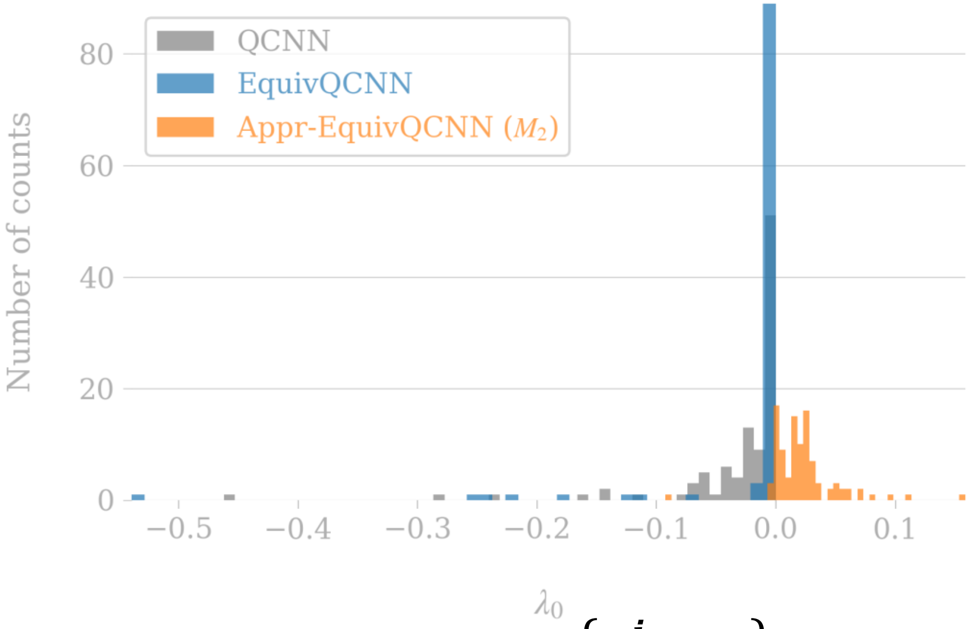
Li, Hao, et al. *Advances in neural information processing systems* 31 (2018).

Non-convexity of loss landscape

- Hessian of f, $\left(H_f\right)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ → Curvature of the loss landscape
- Non-negative eigenvalue λ → Convex function
- Measure the lowest eigenvalue λ_0 for Ising model phase detection

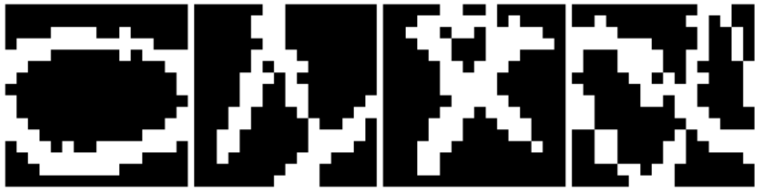


$N = 40$



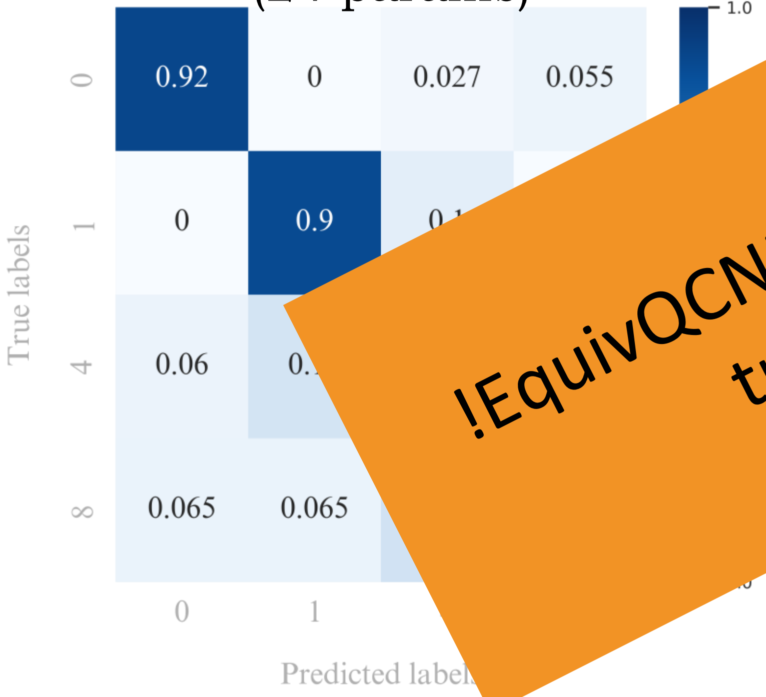
$N = \{2^i \cdot 10\}_{i=1,\ldots,10}$

Multiclass-classification



- 4 class-classification with semeion dataset (16 x 16 pixels handwrittien digits)
- 311 training samples, 313 validation samples

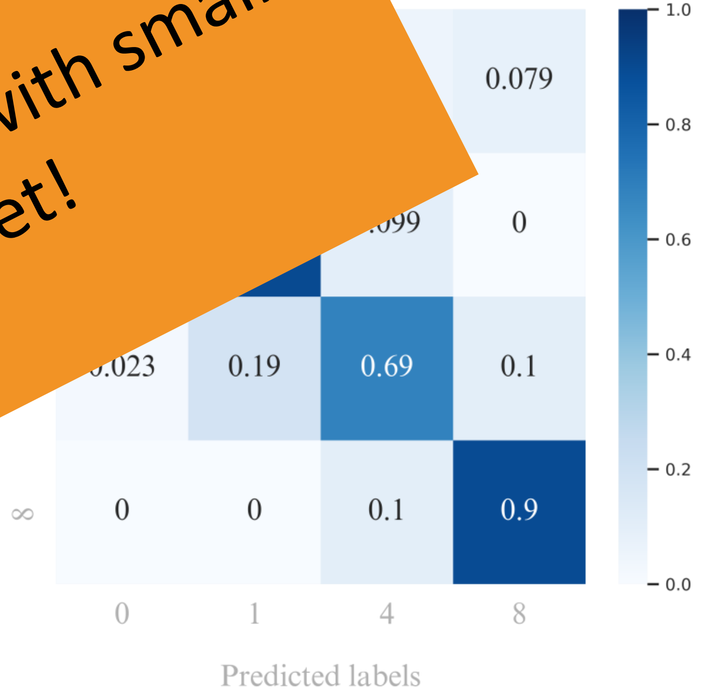
Non-equivariant QCNN
(24 params)



(76.5 ± 1.8)%

!EquivQCNN wins with small training set!

vQCNN M_2
ams)



(82.2 ± 0.2)%

Conclusion

Approximately equivariant QCNN for p4m symmetry

- Construct Equivariant QCNN for planar symmetry on images with a restricted noise
- Better generalization power for EquivQCNN
- Smoother loss landscape proved with eigenvalues of the Hessian

Future works :

1. Quantify the EquivQCNN with other metrics (overparameterization, gradient magnitude)
2. Construct equivariant neural network for RGB images
3. Investigate the impact of noises on GQML.

CERN | QUANTUM TECHNOLOGY INITIATIVE

QUESTIONS?

su.yeon.chang@cern.ch

Codes!