Quantum Convolutional Neural Networks

Iris Cong

Soonwon Choi

Mikhail D. Lukin

arXiv:1810.03787

Berkeley Quantum Information Seminar

October 16th, 2018

Why quantum machine learning?

Machine learning: interpret and process large amounts of data

Unmanned Vehicle

Genomics

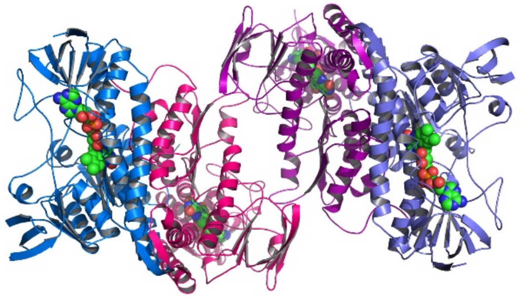


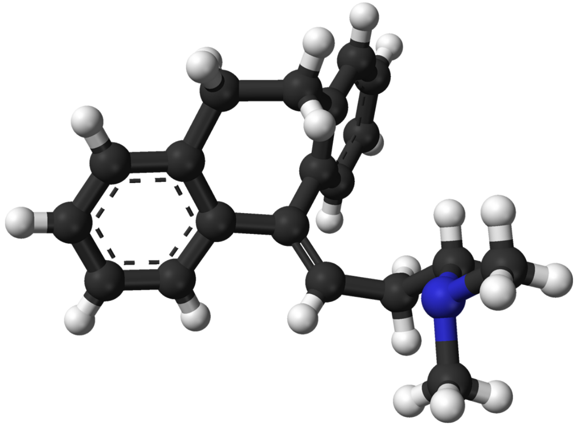
Image recognition

= Cat

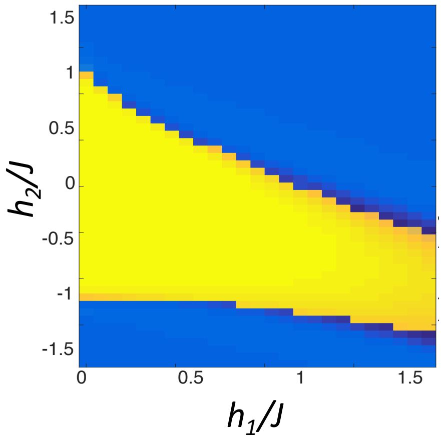
= Dog

Quantum physics: many-body interactions → extremely large complexity

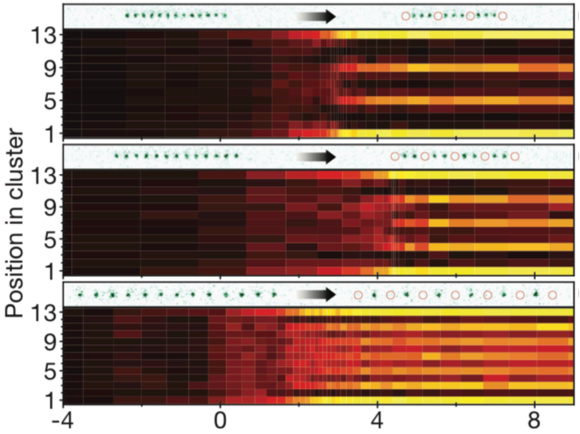
Quantum chemistry



Quantum phases of matter



Near-term quantum computers/ quantum simulators

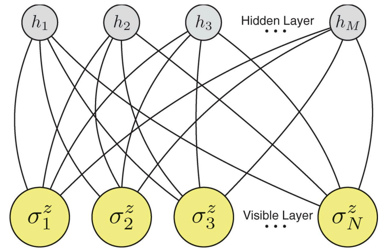


Machine learning + Quantum many-body physics = Exciting!

Machine Learning and Quantum Physics

- (Classical) Machine learning to study quantum physics
- *Quantum machines* to perform classical or quantum computation
  - Quantum speedup? supremacy?

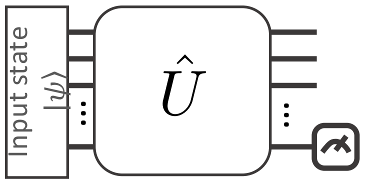
Neural network



Carleo and Troyer, Science (2017)
Carrasquilla and Melko, N. Phys. (2017)
...

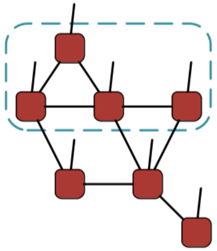
and many others…

Quantum circuits



Farhi and Neven (2018)

Tensor networks



Gao, Zhang, Duan (2017)
Glasser et al., PRX (2018) …

Machine Learning and Quantum Physics

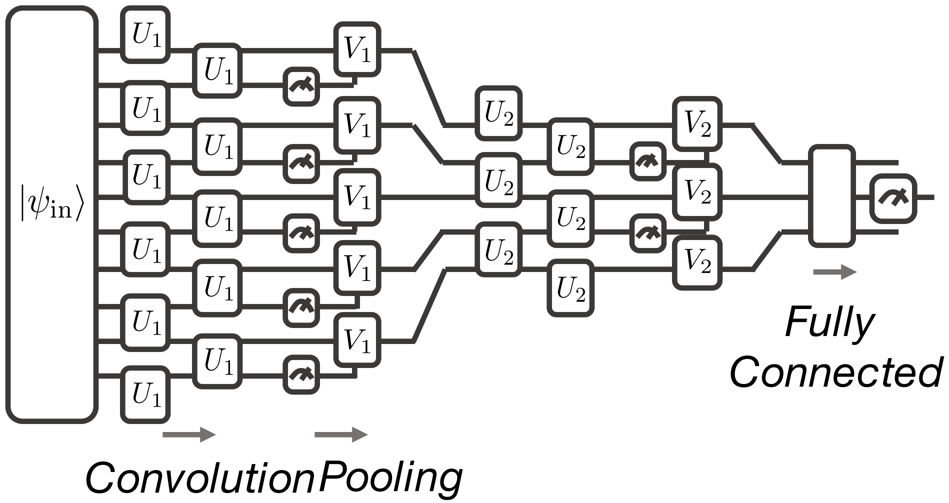
- Many open questions:
  - Why/how does quantum machine learning work?
  - Concrete circuit models suitable for near-term implementation?
  - Relationship to quantum many-body physics?
  - Relationship with quantum information theory?

➡ *Goal of our work*

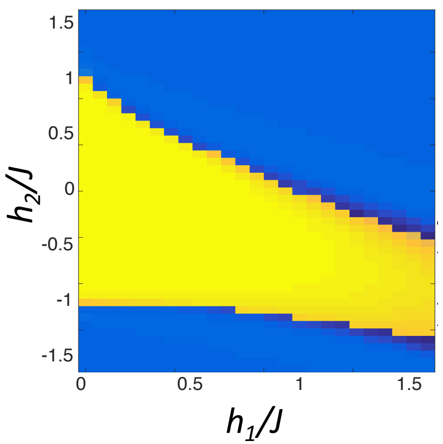
Main Contributions

Concrete and efficient circuit model for quantum classification problems

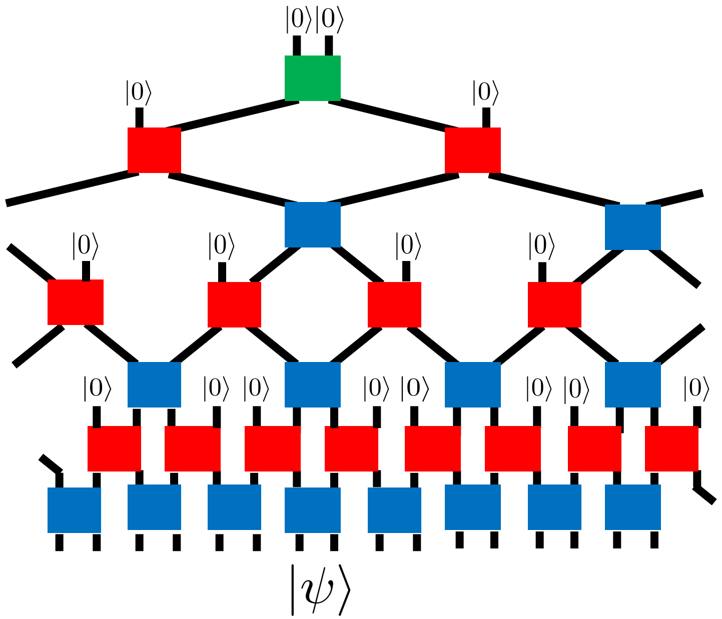
- Good connection to existing ML techniques



Application: Quantum Phase Recognition



Theoretical Explanation: RG Flow, MERA, Quantum Error Correction

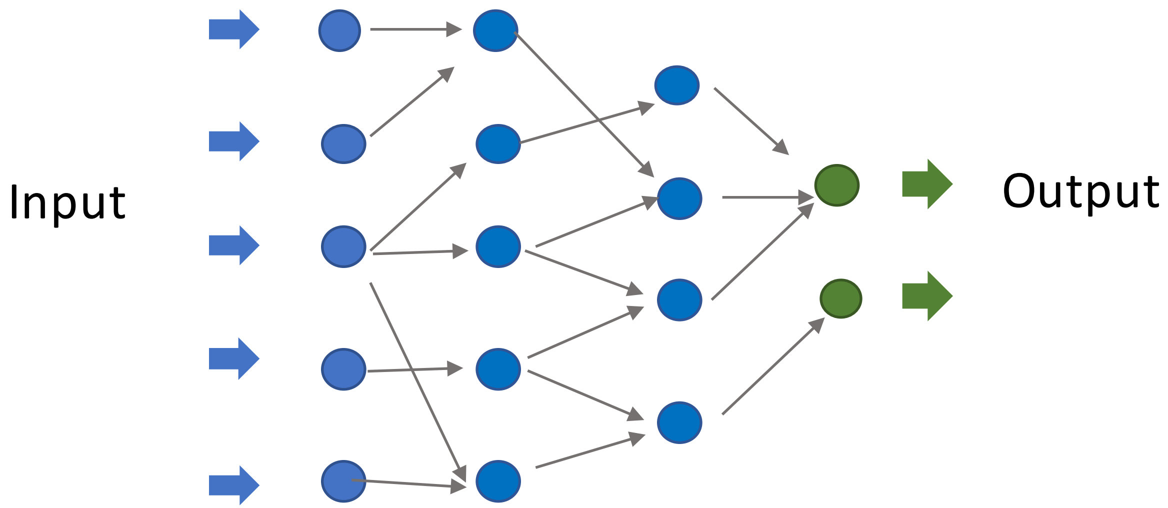


Overview

- **Review of (Classical) Convolutional Neural Networks (CNNs)**
- Quantum Convolutional Neural Networks
- Application to quantum phase recognition
  - Example: 1D symmetry-protected topological (SPT) phase
  - Theoretical explanation: RG flow using MERA + quantum error correction
- Summary and Outlook

Review of (Classical) CNN

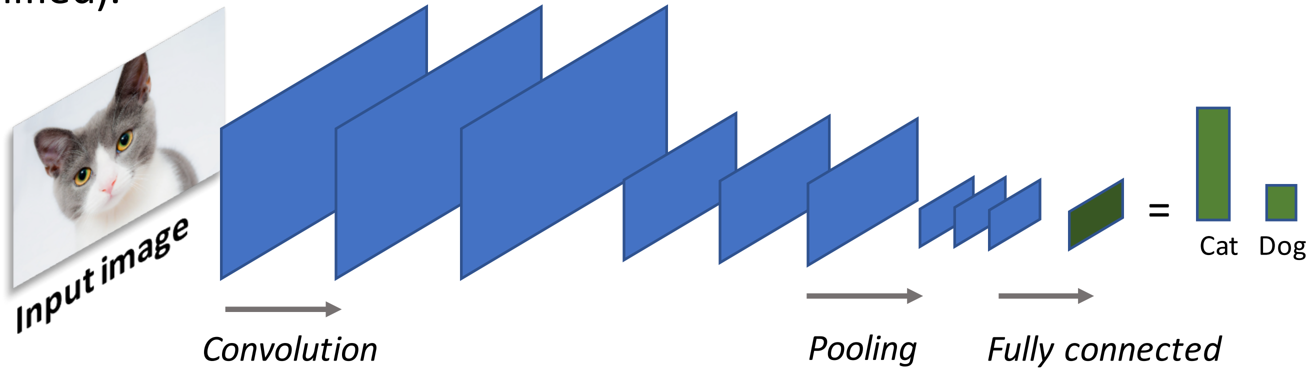
- Particular class of deep, feed-forward neural network



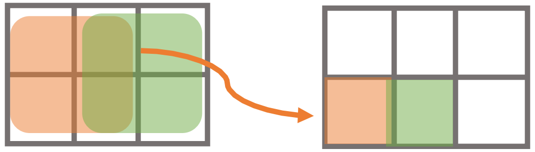
Review of (Classical) CNN

- A structured network – multiple *layers* of image processing

Example (simplified):



Convolution layers



$$y_{i,j} = w_0 x_{i,j} + w_1 x_{i+,j} + w_2 x_{i,j+1} + w_3 x_{i+1,j+1}$$

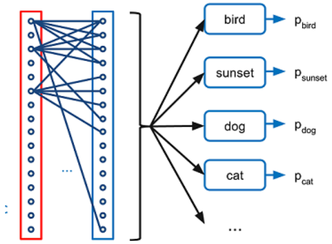
Pooling layers

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

→

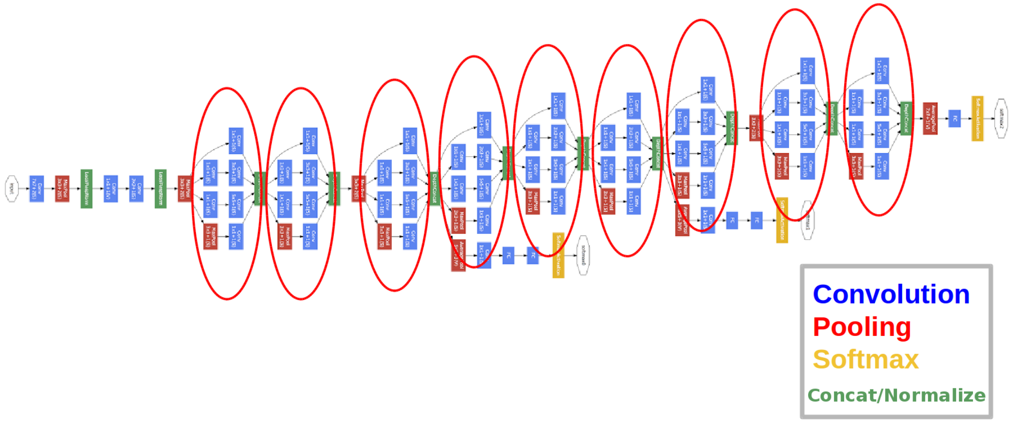
6	8
3	4

Fully connected

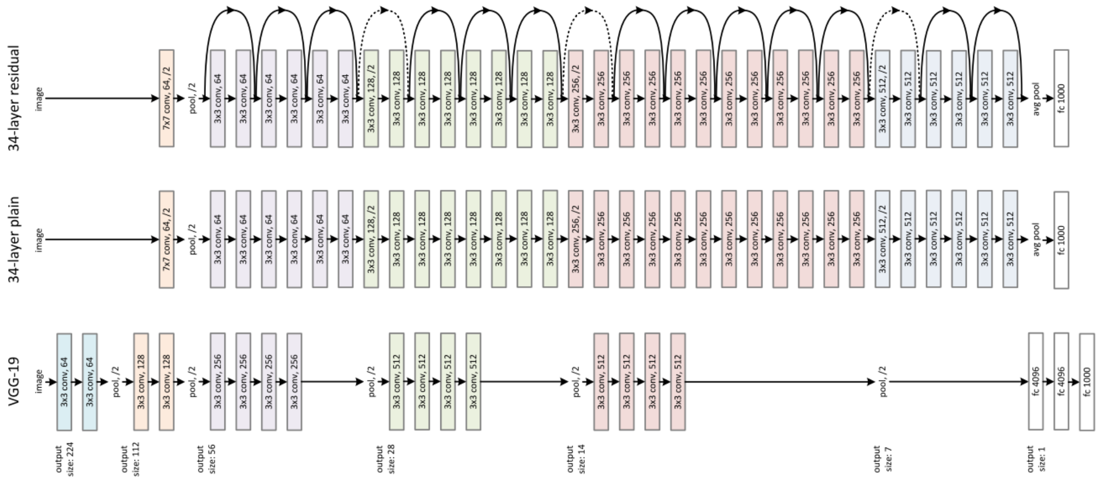


Review of (Classical) CNN

Many ImageNet Competition Winners!



GoogLeNet (2014)



ResNet (2015)

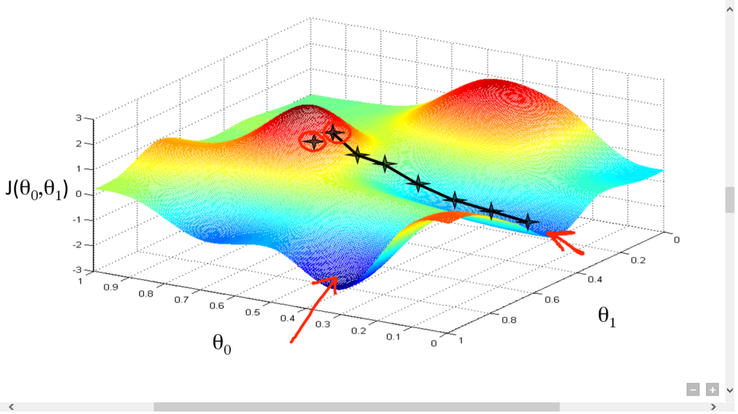
And many others…

Learning Procedure for (Classical) CNN

- Gradient-based supervised learning
- Training data: labeled $\{(\mathbf{x_i}, y_i)\}$
- Common error functions
- Initialize to random weights
- Update **w** to **w' = w** – $\eta\, \nabla_{\mathbf{w}} E$, η = learning rate

$$E(\{(\vec{x_i}, y_i)\}, \mathbf{w}) = \frac{1}{2N} \sum_{i=1}^{N} (y_i - f_{\mathbf{w}}(\vec{x_i}))^2$$

$$\mathrm{KL}(\{(\vec{x_i}, y_i)\}, \mathbf{w}) = \sum_{i=1}^{N} y_i \log \frac{y_i}{f_{\mathbf{w}}(x_i)}$$

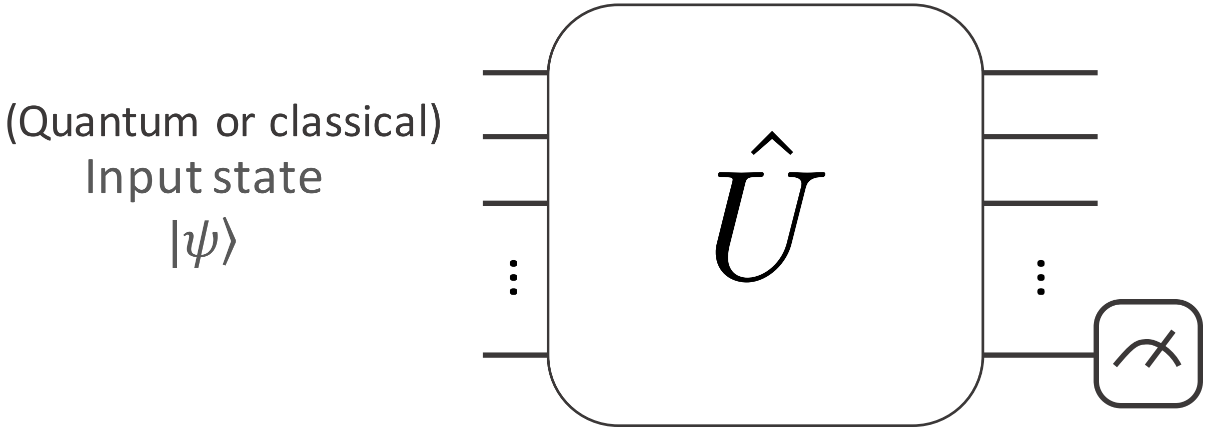


Overview

- Review of (Classical) Convolutional Neural Networks (CNNs)
- **Quantum Convolutional Neural Networks (QCNNs)**
- Application to quantum phase recognition
  - Example: 1D symmetry-protected topological (SPT) phase
  - Theoretical explanation: RG flow using MERA + quantum error correction
- Summary and Outlook

Prior Works: Generic Quantum Circuit

- Farhi and Neven, 2018
- Problems:
  - Too many parameters
  - Subject to overfitting
  - Need all-to-all connectivity
  - Lack of physical intuition?

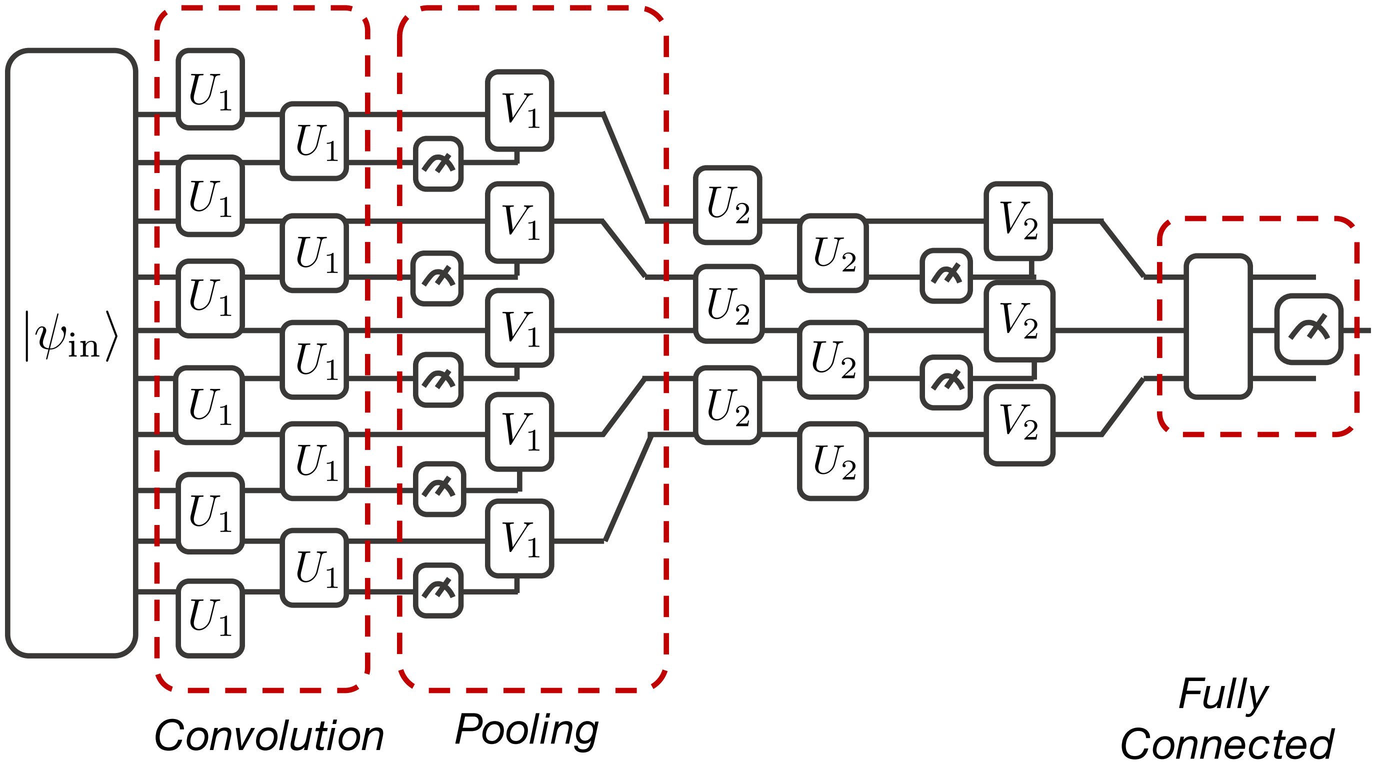


⇨ *Impose "CNN structure"*

Quantum CNN Architecture

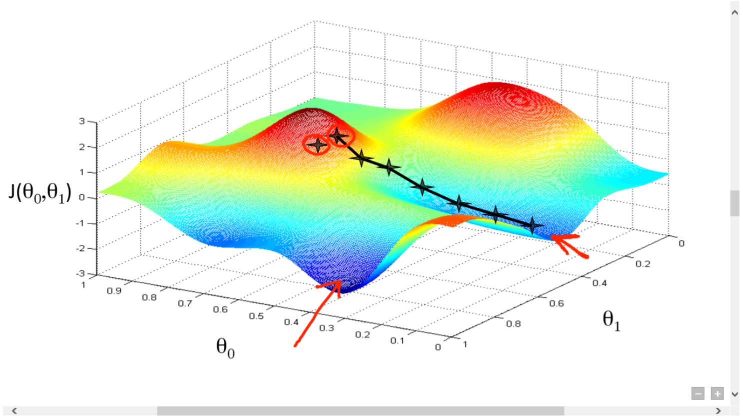
Same types of layers:

1. Convolution
   - Local unitaries, trans. inv., 1D, 2D, 3D …
2. Pooling
   - Reduce system size
   - Final unitary depends on meas. outcomes
3. Fully connected
   - Non-local measurement
   - E.g. for quantum phase recognition: string operator measurement



Learning Procedure for Quantum CNN

- Parameterize unitaries in convolution, pooling, and fully-connected layers
- Initialize to random unitaries
- Similar gradient-based learning



Overview

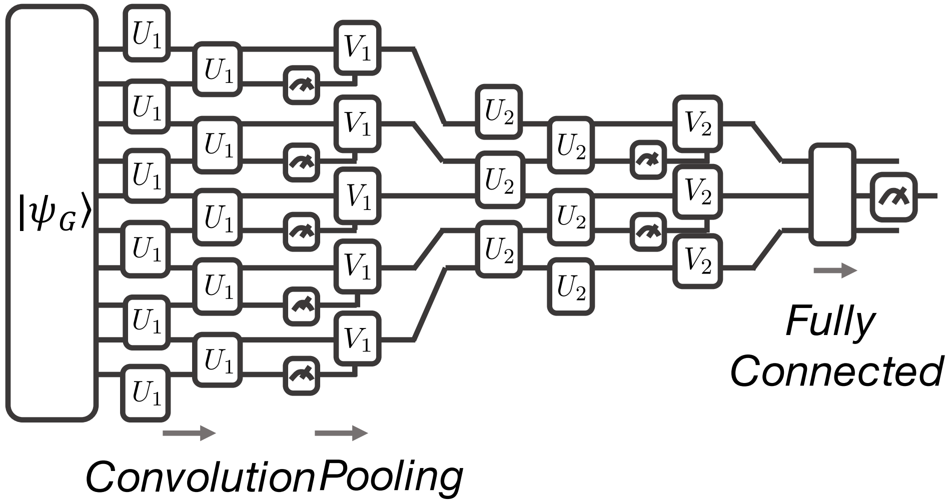
- Review of (Classical) Convolutional Neural Networks (CNNs)
- Quantum Convolutional Neural Networks
- **Application to quantum phase recognition**
    - Example: 1D symmetry-protected topological (SPT) phase
    - Theoretical explanation: RG flow using MERA + quantum error correction
- Summary and Outlook

Application: Quantum Phase Recognition

Problem: Given quantum many-body system in (unknown) ground state $|\psi_G\rangle$, does $|\psi_G\rangle$ belong to a particular quantum phase $\mathcal{P}$?

Direct analog of image classification, but *intrinsically quantum problem*

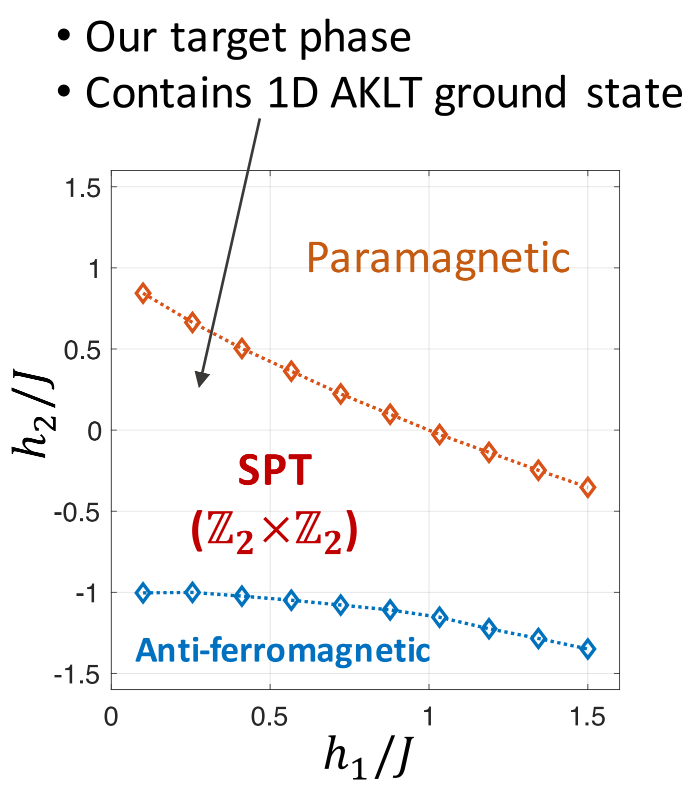
Claim: Quantum CNN is very efficient in quantum phase recognition



$$= \begin{cases} 1 & \text{if } |\psi_G\rangle \in \mathcal{P} \\ 0 & \text{if } |\psi_G\rangle \notin \mathcal{P} \end{cases}$$

Example: 1D ZXZ Model ($G = \mathbb{Z}_2 \times \mathbb{Z}_2$ SPT)

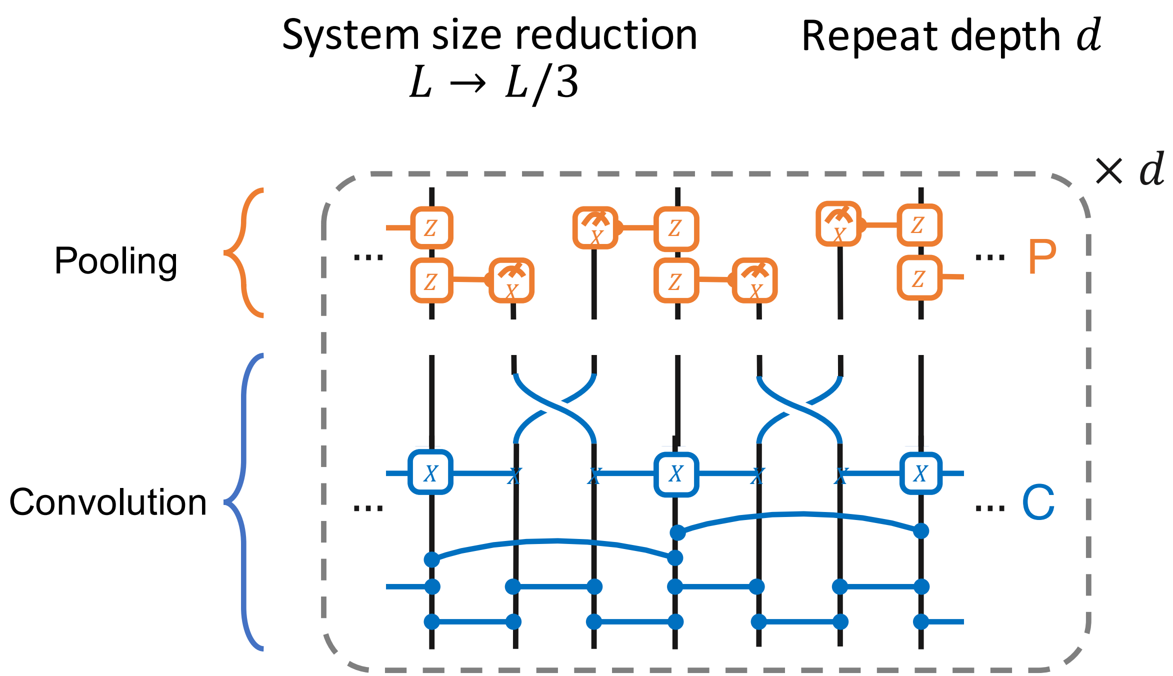
- $H = -J\sum_i Z_i X_{i+1} Z_{i+2} - h_1 \sum_i X_i - h_2 \sum_i X_i X_{i+1}$
- $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry: $\hat{X}_e = \prod_{i \in even} \hat{X}_i$, and $\hat{X}_o = \prod_{i \in odd} \hat{X}_i$
- When $h_1 = h_2 = 0$, g.s. = 1D cluster state (stabilizer code)
- When $h_2 = 0$, exactly solvable by Jordan-Wigner
- Simple example → good analytical guess for QCNN circuit, demonstrate learning later

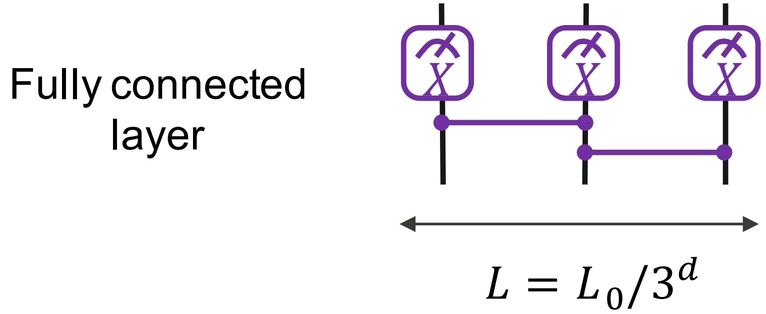


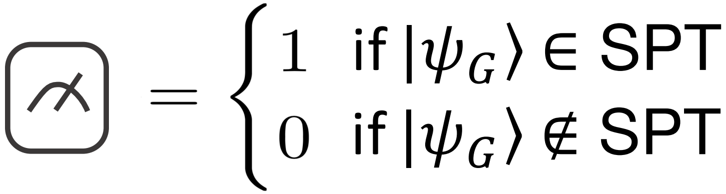
* Phase diagram is obtained from iDMRG with bond dimension 150.

Example: 1D ZXZ Model ($G = \mathbb{Z}_2 \times \mathbb{Z}_2$ SPT)

- Quantum CNN Circuit:

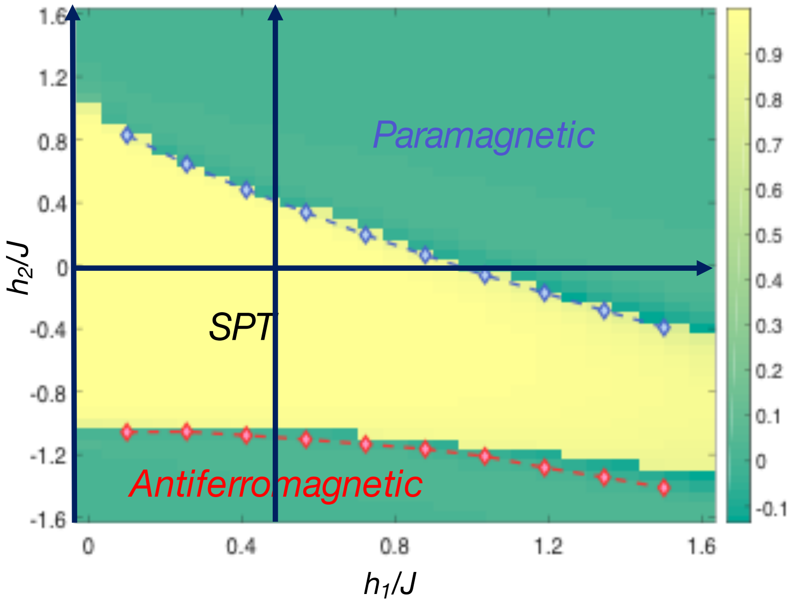




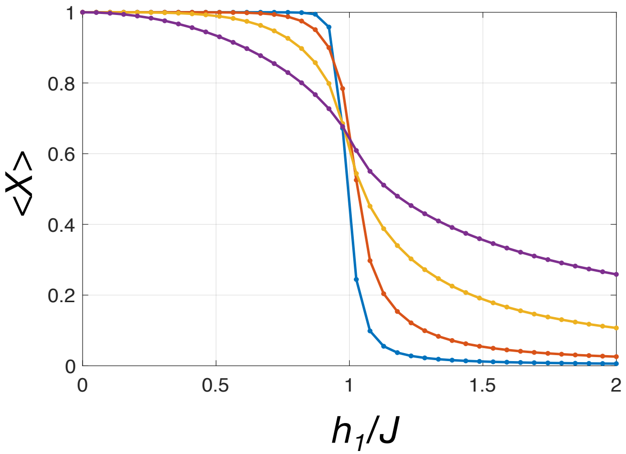


Example: 1D ZXZ Model ($G = \mathbb{Z}_2 \times \mathbb{Z}_2$ SPT)

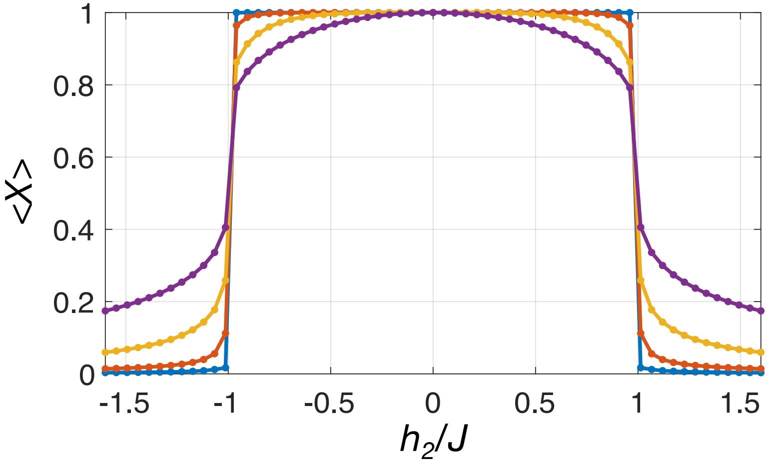
- Results of numerical simulation:



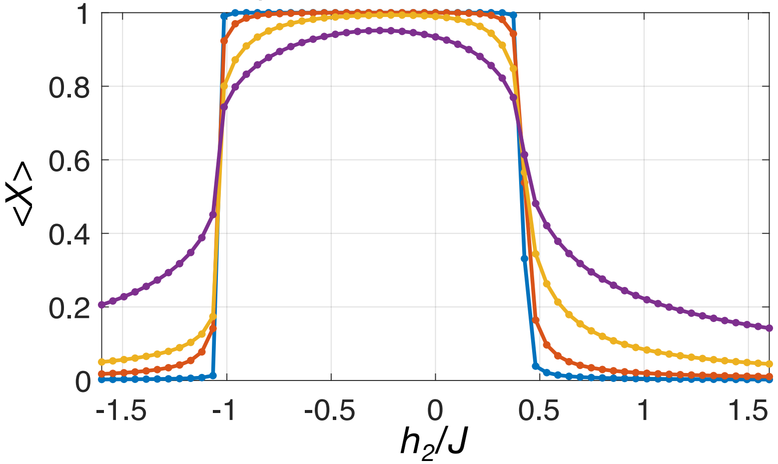
Fixed $h_2 = 0$:



Fixed $h_1 = 0$:



Fixed $h_1 = 0.5J$:



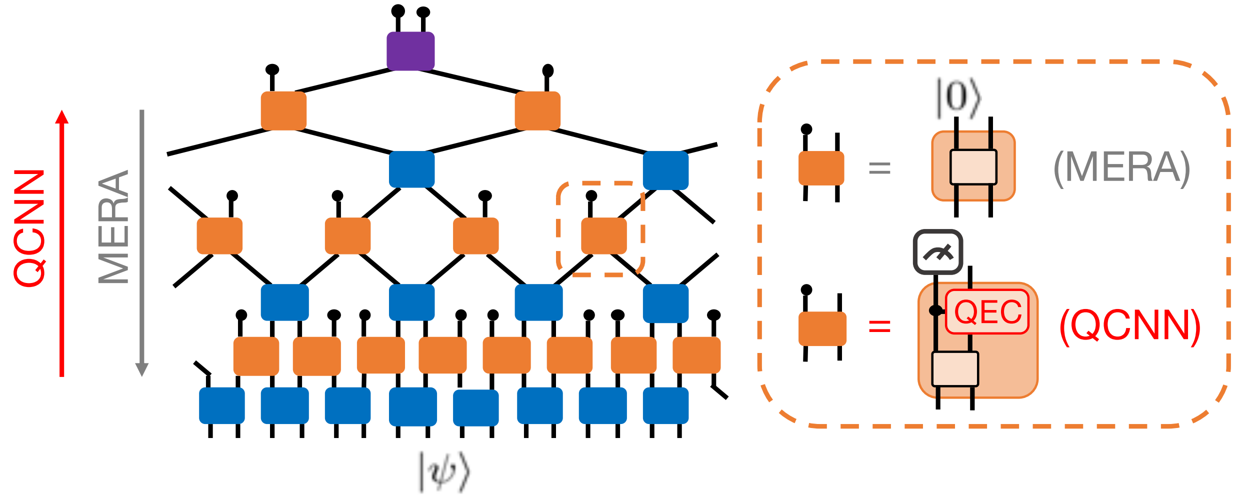
* Phase diagram is obtained from iDMRG with bond dimension 150.
Ground states are obtained from DMRG with bond dimension 130.
2D plot: $N = 45$ spins (depth 2), 1D curves: $N = 135$ spins

Why does it work?

Multiscale Entanglement Renormalization Ansatz (MERA)

+

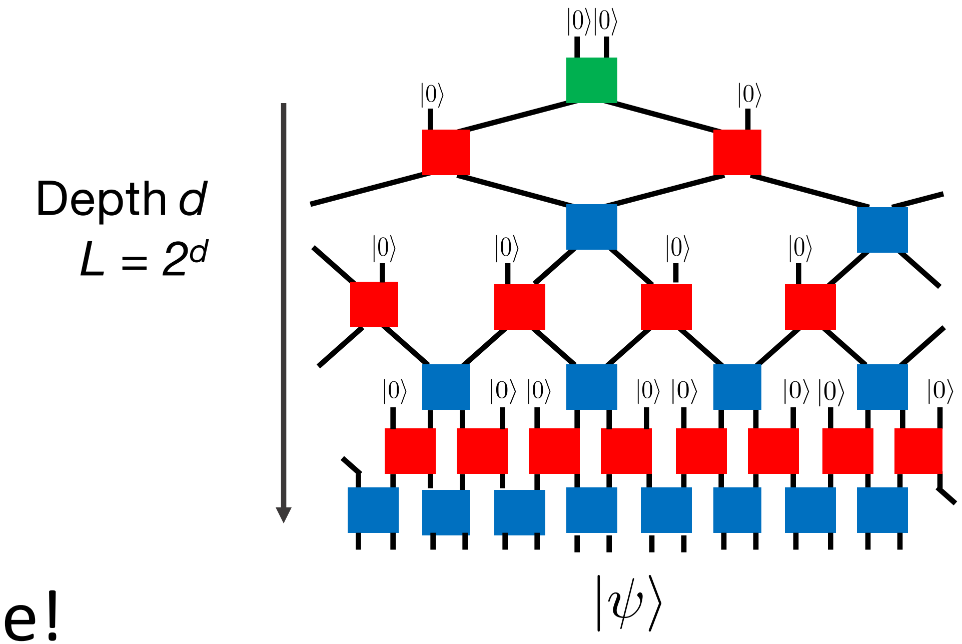
Quantum Error Correction



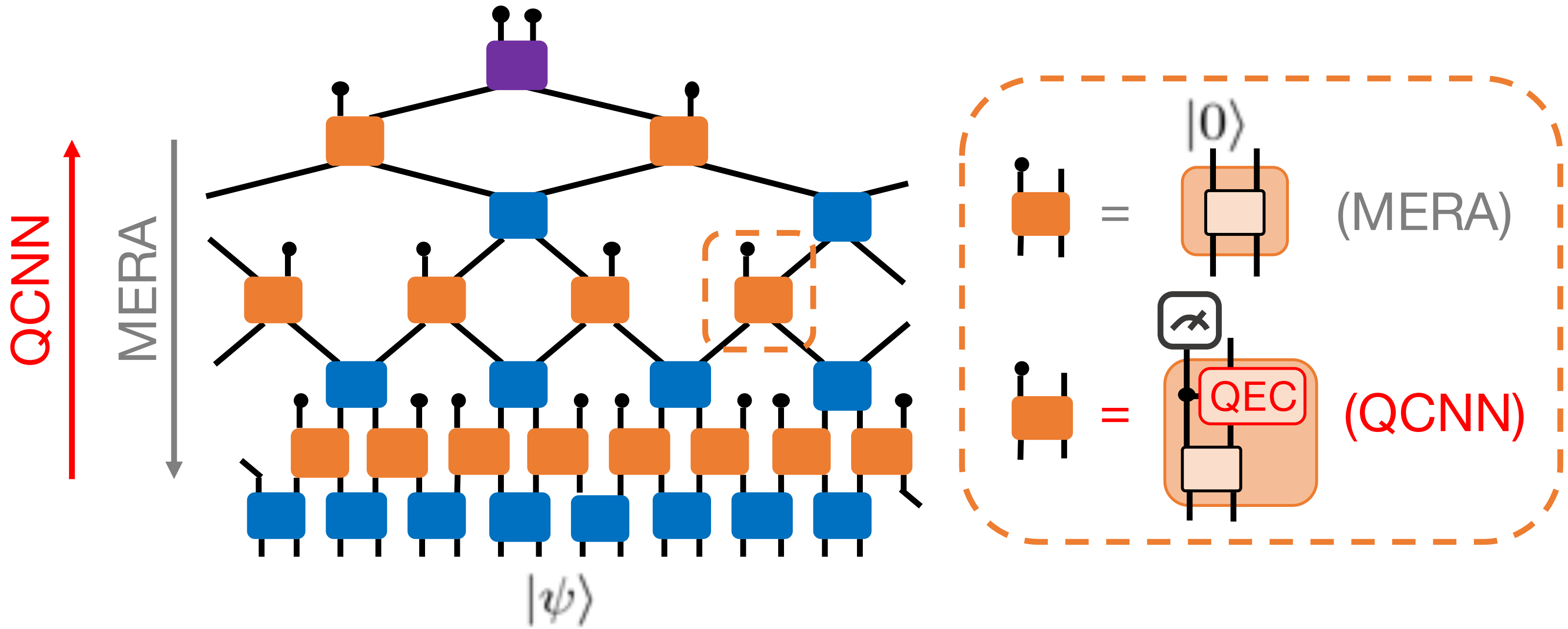
⇨ Circuit simulation of RG flow

Multiscale Entanglement Renormalization Ansatz (MERA)

- Variational ansatz for many-body wavefunction
- Efficient description of:
  - Gapped ground states
  - Critical systems
  - Many interesting physical systems…
- Quantum CNN has very similar structure!



Why does it work?

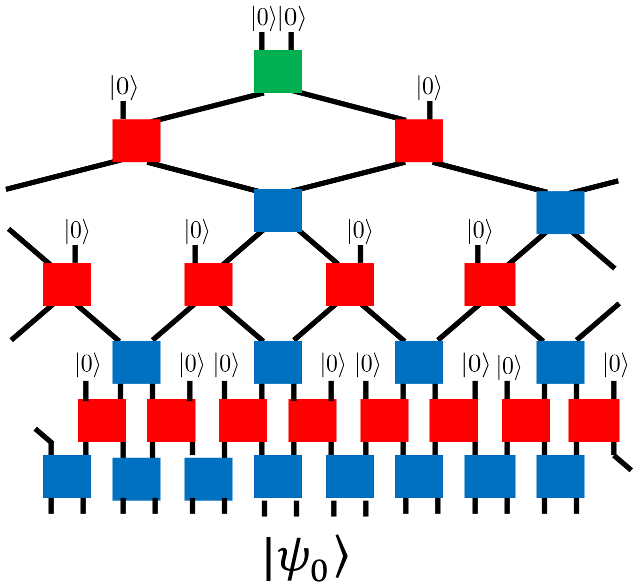


Key differences:

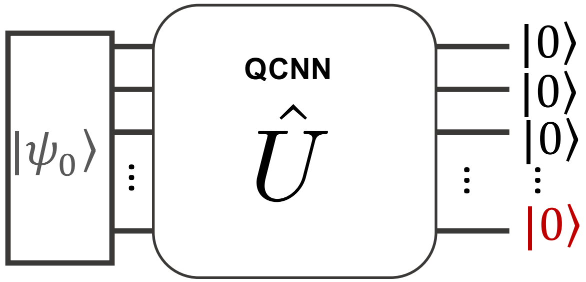
- Opposite direction
- MERA circuit is only defined for ancillas = |0>
- Measurement does not always give |0>!

Why does it work?

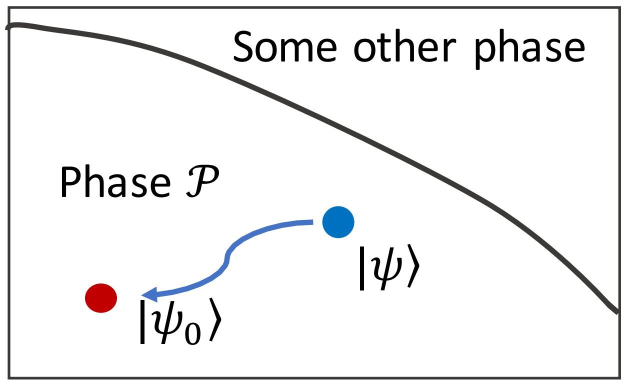
Good starting point!



MERA representation of $|\psi_0\rangle$ → QCNN circuit to exactly recognize $|\psi_0\rangle$ with deterministic intermediate measurement outcomes.



… but NOT sufficient for quantum phase recognition.



Ideally, a *single* QCNN should recognize *ALL* $|\psi\rangle$ in the same phase $\mathcal{P}$.

Key idea: use quantum error correction to induce "flow."

Why does it work?

- Idea: Treat perturbations terms as quantum errors
- Example:

$$H = \underbrace{-J\sum_i Z_i X_{i+1} Z_{i+2}}_{\text{RG fixed point = Quantum code}} \underbrace{- h_1 \sum_i X_i - h_2 \sum_i X_i X_{i+1}}_{\text{Error perturbations}}$$

RG fixed point
= Quantum code

Error perturbations

- This construction is generic: includes all 1D SPT, 2D string-net models

Why does it work?

(cluster state)

- Quantum CNN for $\mathbb{Z}_2\times\mathbb{Z}_2$ SPT phase = MERA for $|\psi_0\rangle$ + QEC

$|\psi_0{}^{(L)}\rangle$ → $|\psi_0{}^{(L/3)}\rangle \otimes |0000\ldots\rangle$

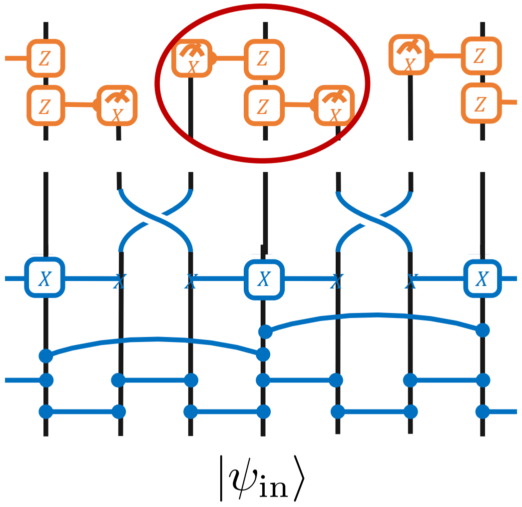
$|\psi^{(L)}\rangle$ → $|\psi'^{(L/3)}\rangle \otimes |01001\ldots\rangle$

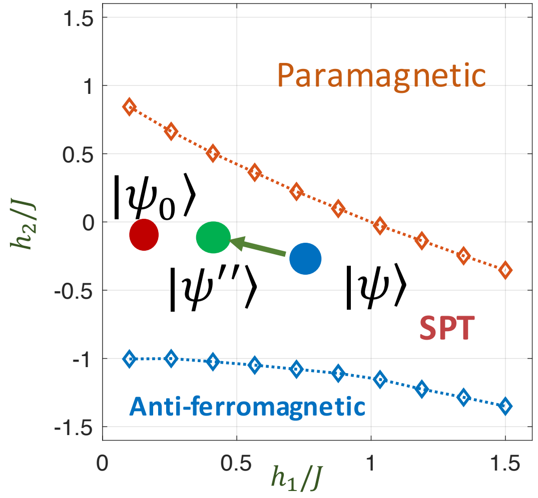
→ $|\psi''^{(L/3)}\rangle \otimes |01001\ldots\rangle$

Quantum error correction

→ Repeat d times

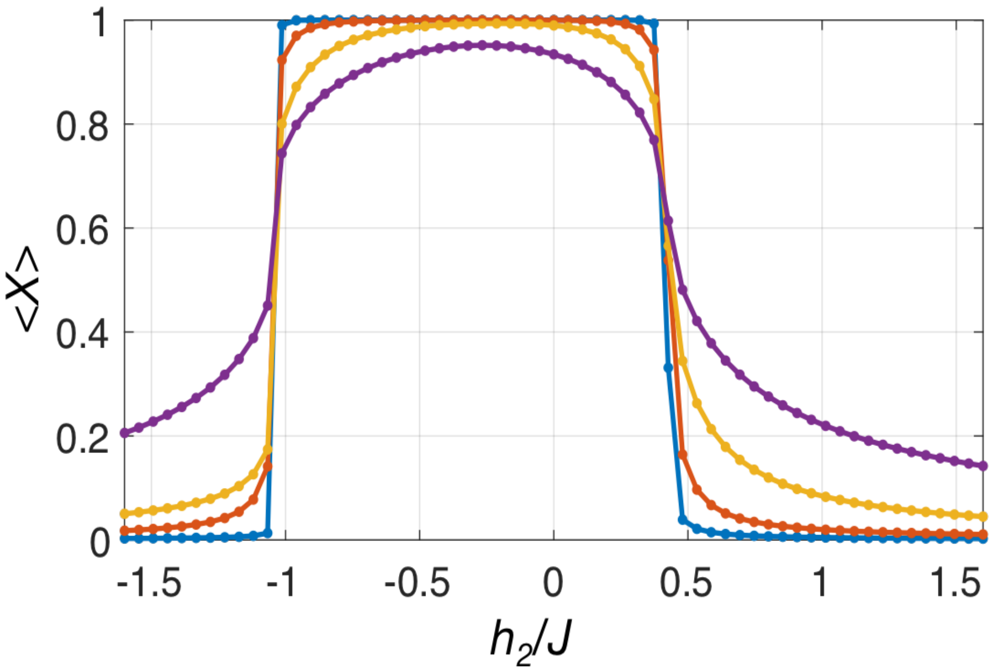
→ Measure string order parameter for $|\psi_0\rangle$

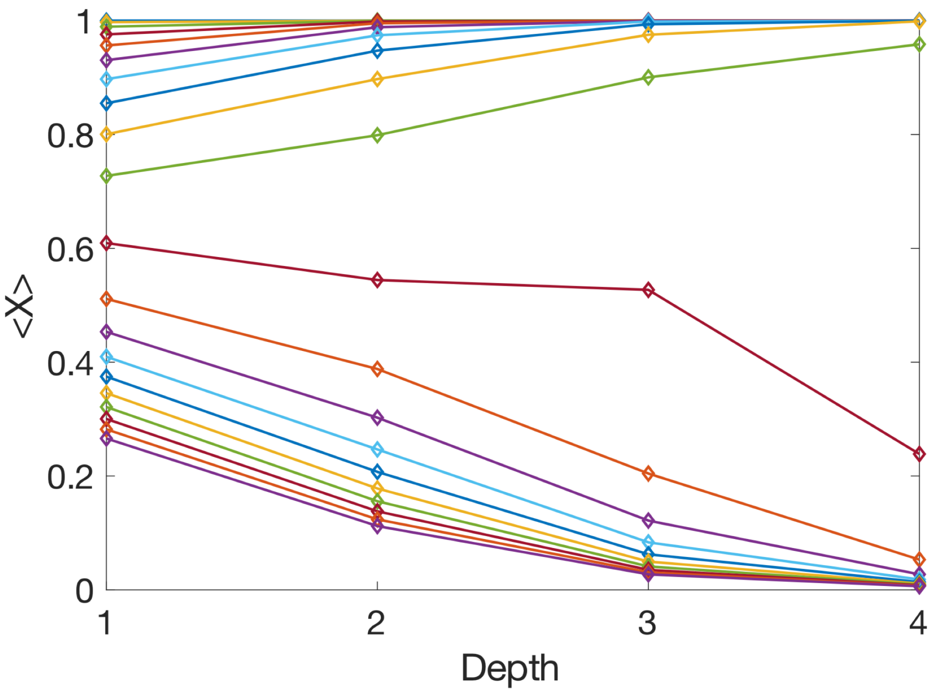




Example: 1D ZXZ Model ($G = \mathbb{Z}_2 \times \mathbb{Z}_2$ SPT)

RG flow in our example (N = 135 spins):

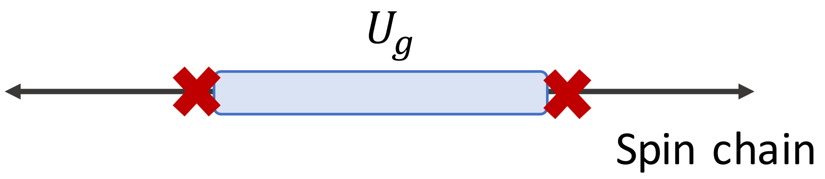




Why does it work?

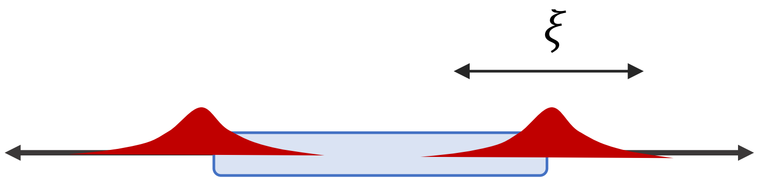
(alternative explanation using string order parameters)

- Identifying an SPT phase = Detecting a string order parameter



Example: $\mathbb{Z}_2 \times \mathbb{Z}_2$ SPT $\quad \langle \mathcal{S}_{ij} \rangle = \langle Z_i X_{i+1} X_{i+3} \dots X_{j-1} Z_j \rangle$

- Problem: away from RG fixed point

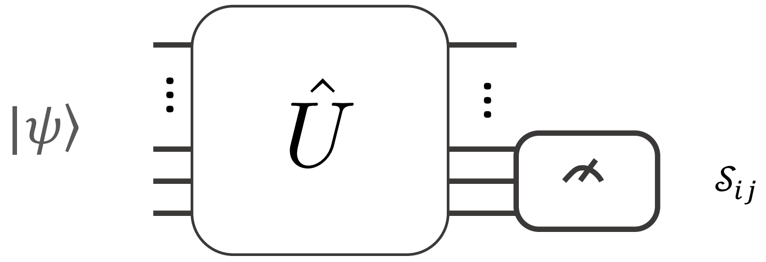


$$\mathcal{S}_{ij} = \mathcal{O}\left(1/\xi^a\right)$$

QCNN: *use more than one string operator*

Why does it work?

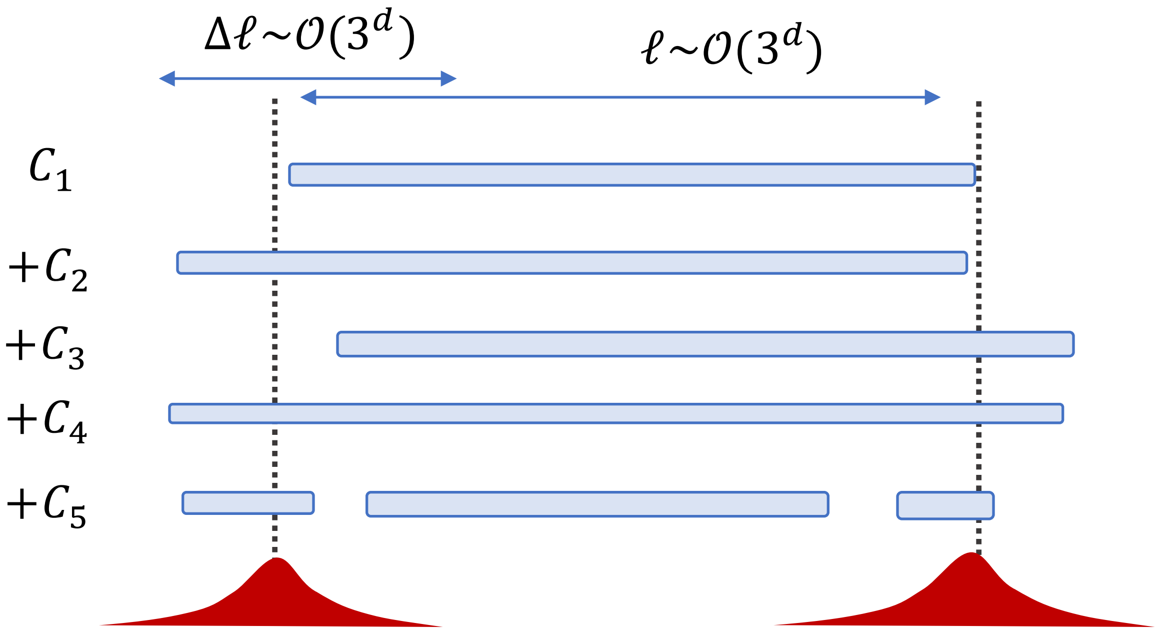
(alternative explanation using string order parameters)



Heisenberg picture

$$\langle\psi|U^\dagger \mathcal{S}_{ij} U|\psi\rangle$$

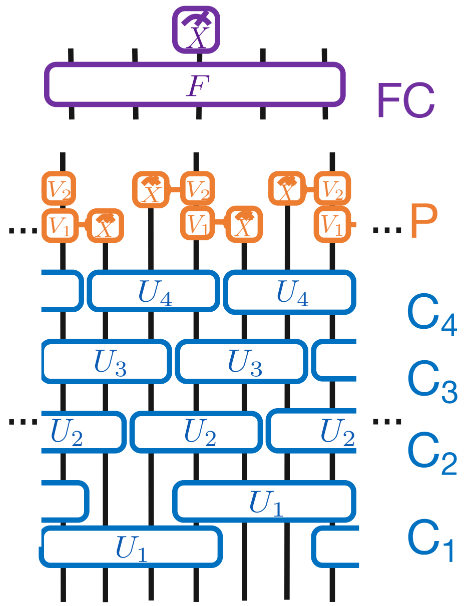
$$\langle\psi|U^\dagger Z_{i-1} X_i Z_{i+1} U|\psi\rangle =$$

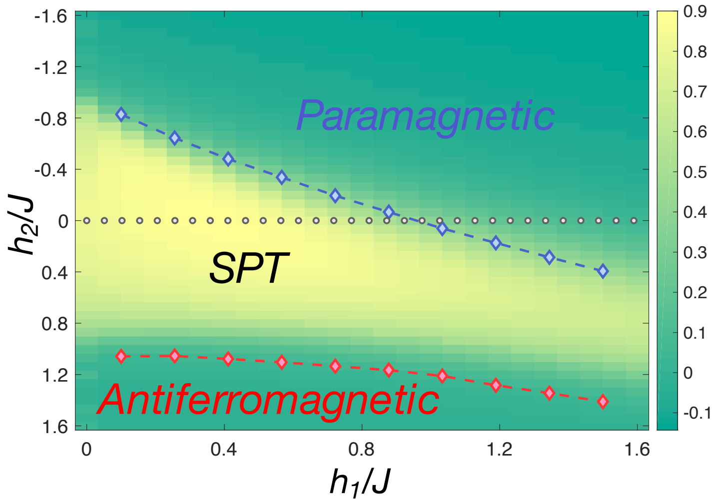


Exponentially many, long string operators!

Training: Example

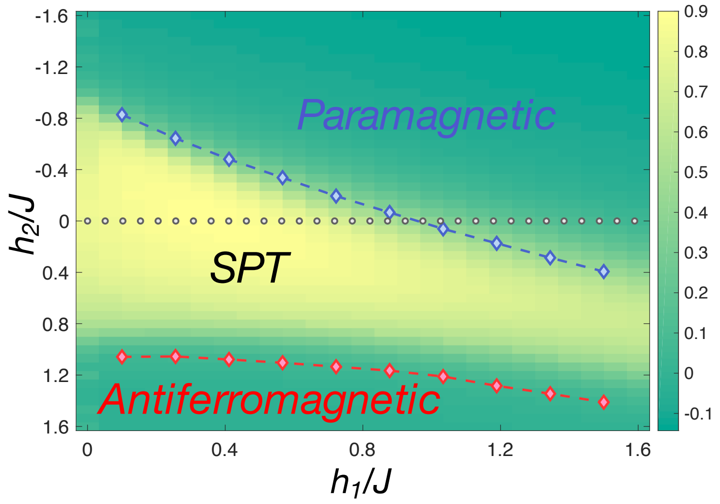
- Circuit structure:
- N = 15 spins (depth 1) for simulations
- Initialize all unitaries to random values
- Train along h_2 = 0 (solvable by JW)
- Gradient descent: $\mathrm{MSE} = \frac{1}{2M}\sum_{\alpha=1}^{M}(y_i - f_{\{U_i,V_j,F\}}(|\psi_\alpha\rangle))^2$





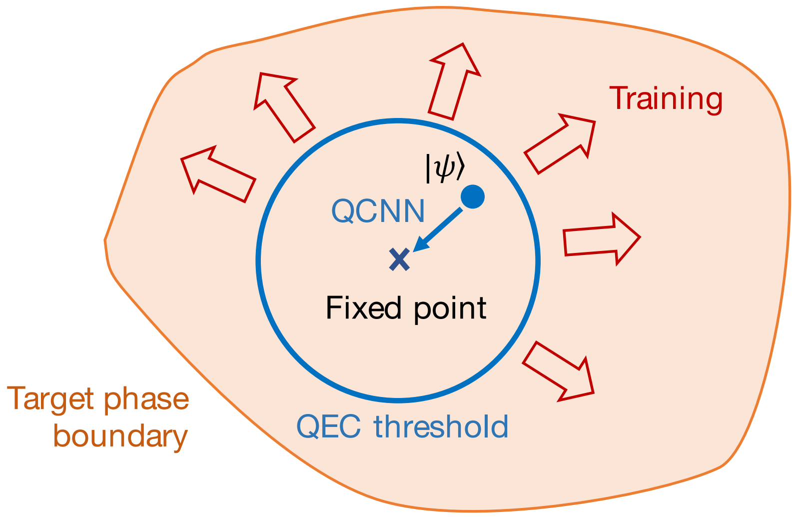
Training: Example

- Observation: training on 1D, JW-solvable set can still produce the correct 2D phase diagram
- Demonstrates how QCNN structure avoids overfitting



Training: General Cases

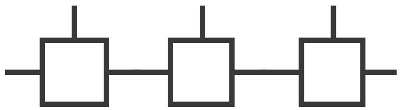
- More generally: training = expanding QEC threshold
- Efficiency: reduced # of parameters ($O(\log(N))$ instead of $O(\exp(N))$)



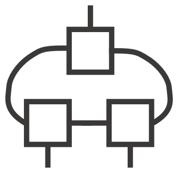
General Recipe - QCNN for quantum phase recognition

- Step 1: Given a quantum phase construct a MPS representation of a RG fixed point and its MERA circuit

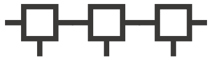
Schuch *et al*, PRB **84**, 165139 (2011)

$|\psi_0\rangle =$ 

(isometric form)

MERA isometry = 

- Step 2: Construct its parent Hamiltonian $H = \sum_i h_i$ with $[h_i, h_j] = 0$.

- Step 3: Design quantum error correction code based on
  - Stabilizers from h_i.
  - Code space from 

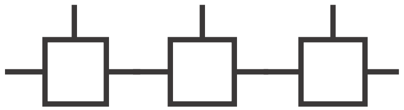
- Step 4: Construct QCNN ansatz based on Step 3.

- Step 5: Optimize QCNN via learning procedures.

Generic Construction of QCNN Circuit

(High-level outline; in progress)

- **Goal**: Given quantum phase $\mathcal{P}$ classified by its RG fixed point, construct quantum CNN circuit to best recognize states in $\mathcal{P}$
- Step 1: Construct tensor network representation of RG fixed point

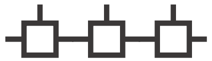
$|\psi_0\rangle =$  (isometric form)

Schuch *et al*, PRB **84**, 165139 (2011)

- Step 2: Construct parent Hamiltonian of this tensor network (nearest-neighbor commuting terms): $H = \sum_i h_i$ with $[h_i, h_j] = 0$.

Generic Construction of QCNN Circuit

(High-level outline; in progress)

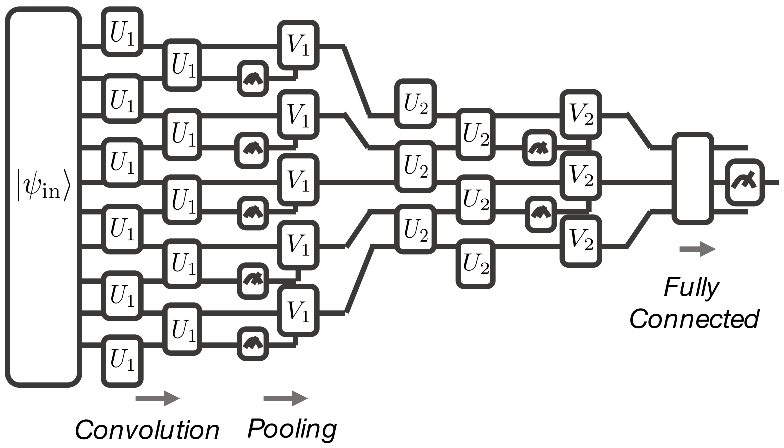
- Step 3: Design quantum error correction code with (generalized) stabilizers which are the h_i. The code space forms the ground state space 
- Step 4: Construct quantum CNN circuit based on Step 3
- Step 5: Optimize quantum CNN based on learning procedures

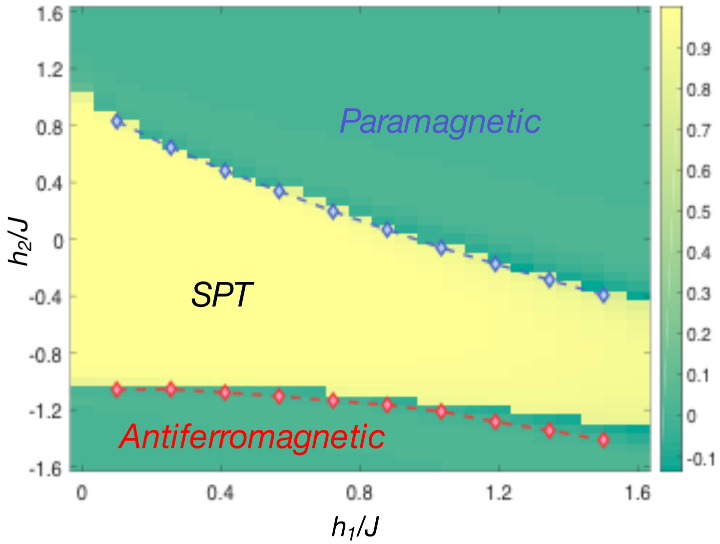
Overview

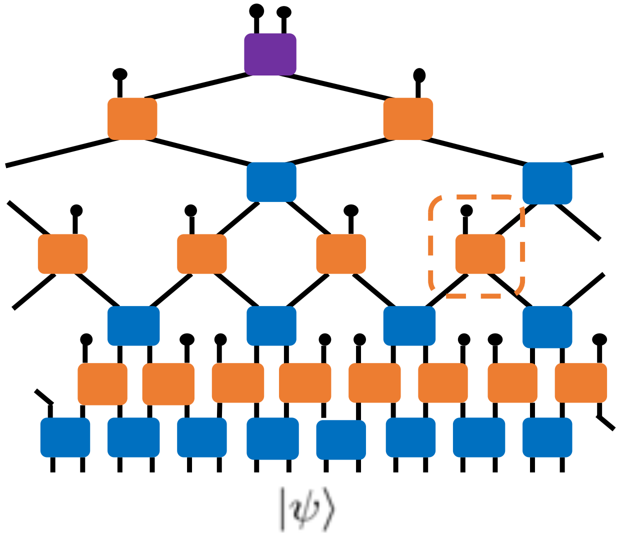
- Review of (Classical) CNN
- Quantum CNN: Architecture and Learning Procedure
- Application: quantum phase recognition
  - Example: 1D cluster state
  - Theoretical explanation: RG flow using MERA + quantum error correction
- **Summary and Outlook**

Summary

- Concrete architecture, learning proposal for quantum classification
- Application to quantum phase recognition: 1D SPT phase ($\mathbb{Z}_2 \times \mathbb{Z}_2$)
- Theoretical explanation: QCNN ≈ MERA + QEC ≈ RG flow







Other Results and Outlook

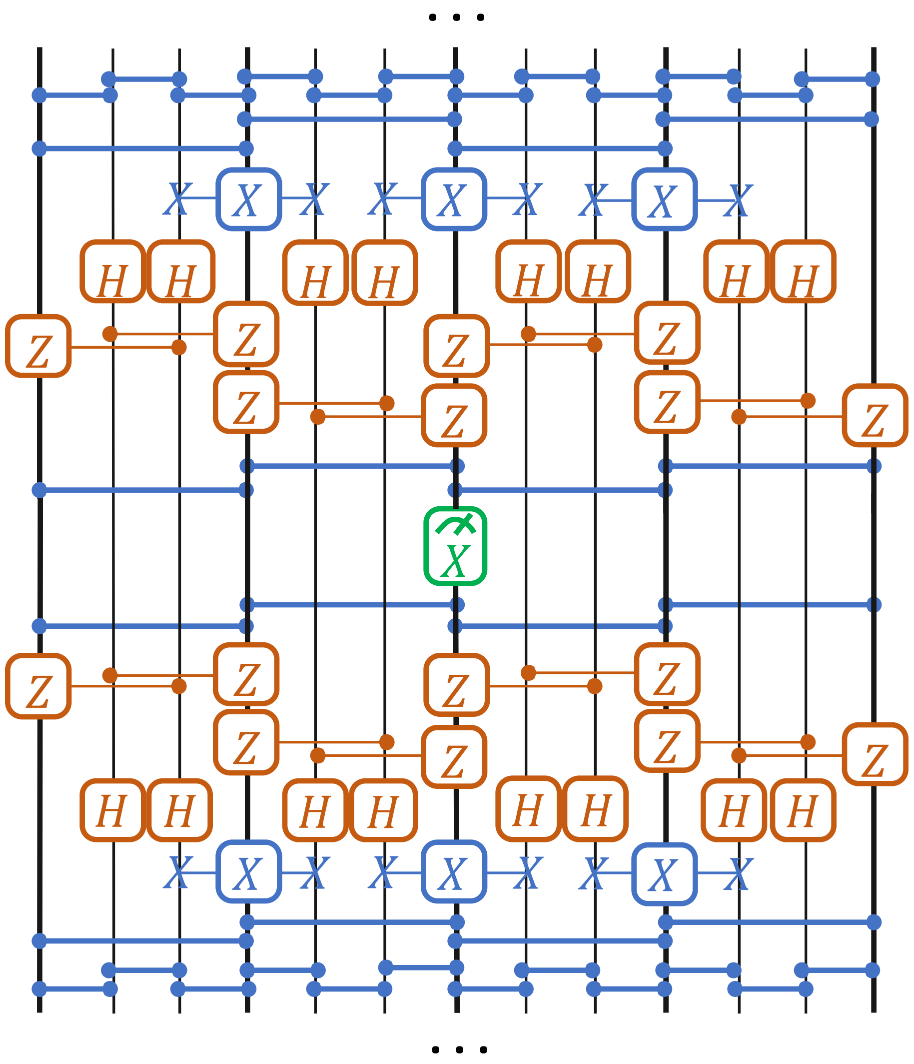
- 2D Quantum CNN circuit: toric code, string-net
- Efficient training procedures inspired by machine learning
- Gapless phases (e.g. spin liquids), holographic codes
- Application: fault-tolerant computation
- Experimental realizations
  - e.g. trapped ions, Rydberg atoms, superconducting circuits

Thanks!

Backup Slides

Heisenberg Picture: Topological Measurement

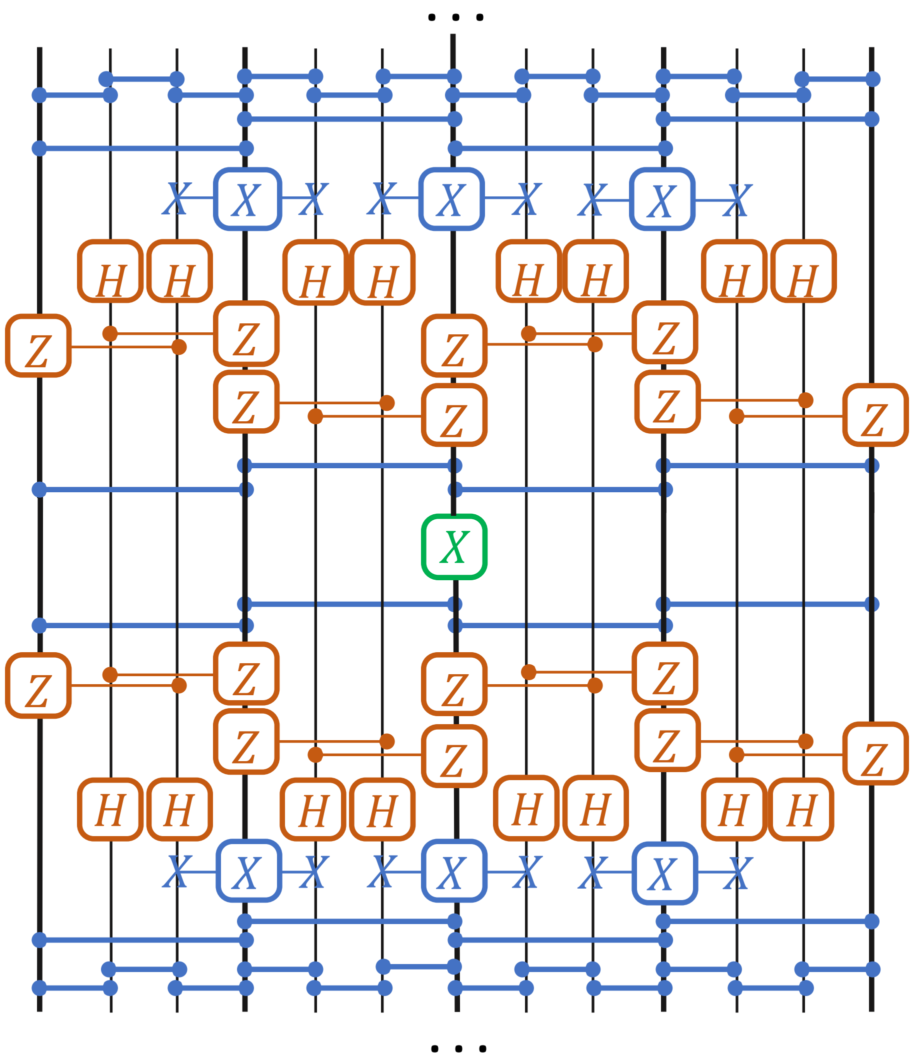
Measured operator =



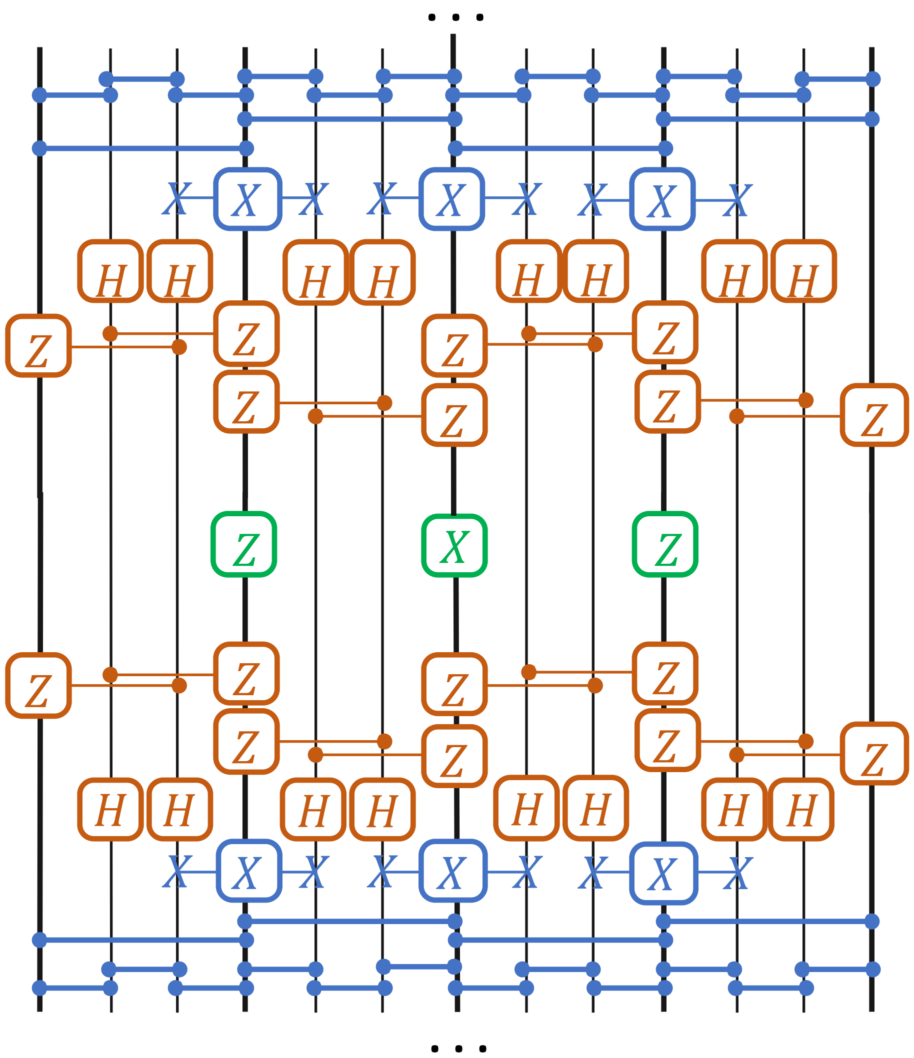
= sum of exponentially many string operators

Detects presence of edge modes!

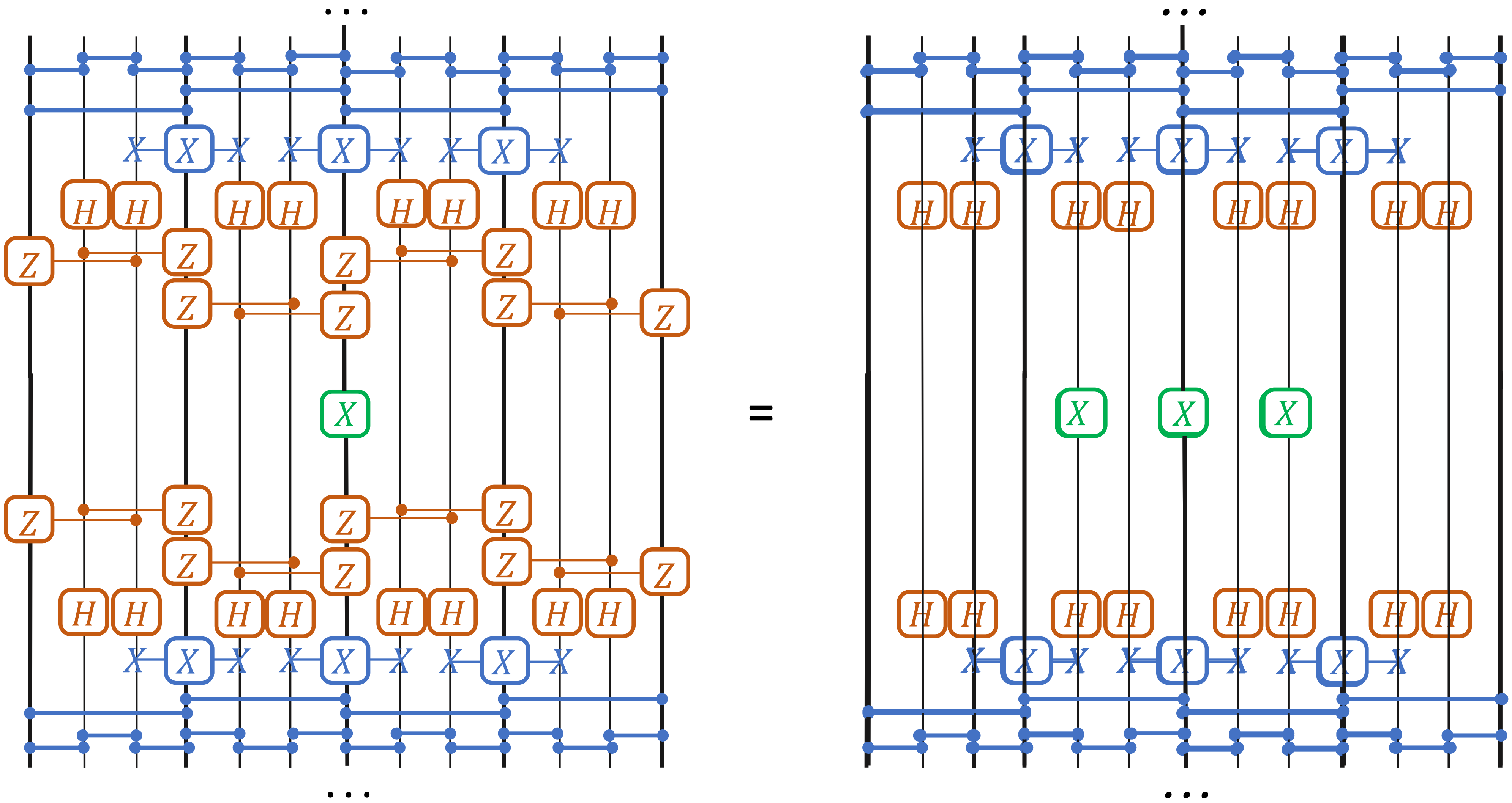
Heisenberg Picture: Topological Measurement



=



Heisenberg Picture: Topological Measurement



Heisenberg Picture: Topological Measurement

- Conclusion:
  - Measured operator is sum of exponentially many string operators
  - Measures probability for all Majorana modes to be "topologically connected"

    → 1 in the SPT phase, 0 in the trivial phase