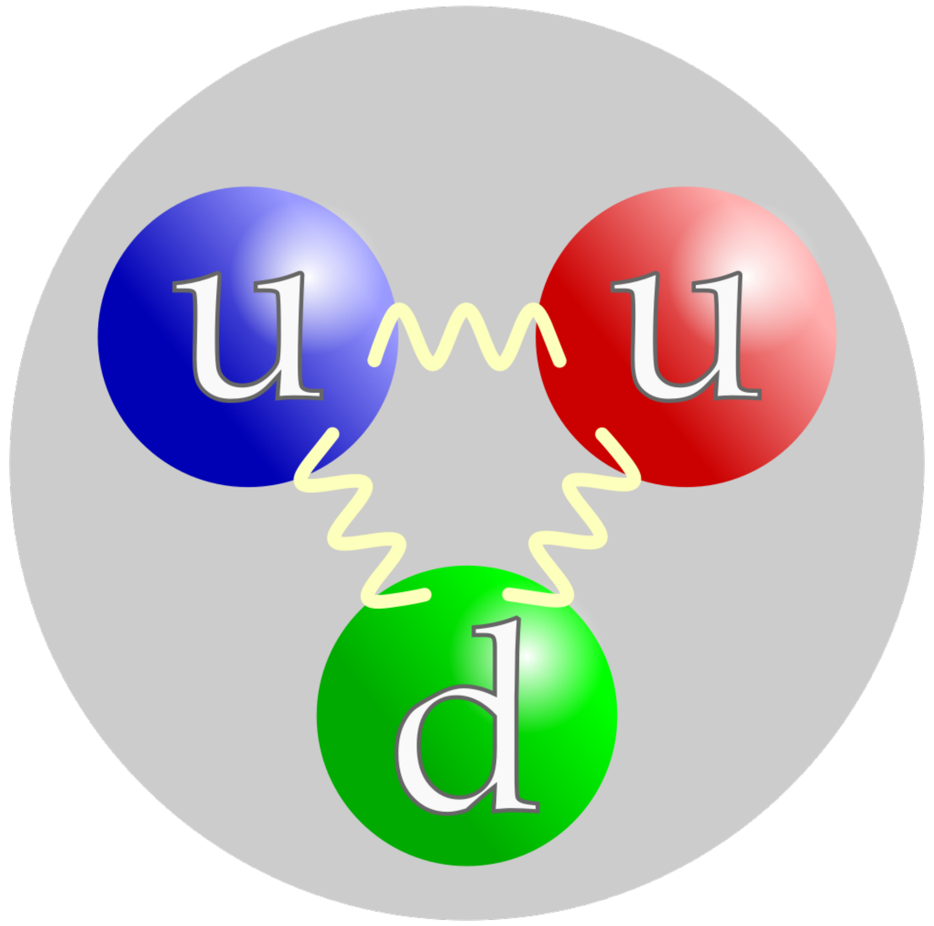
Quantum data learning for quantum simulations in high-energy physics

Lento Nagano (ICEPP, The University of Tokyo)

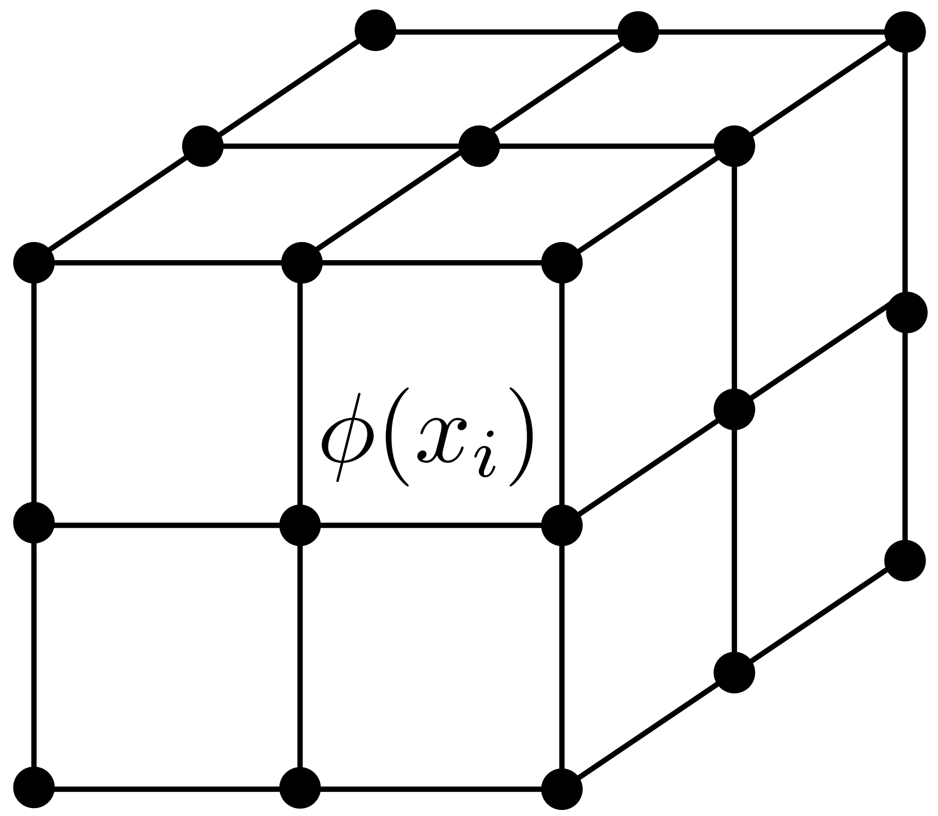
Based on [arXiv:2306.17214] (to appear in PRR)
with A. Miessen, F. Tacchino, T. Onodera, I. Tavernelli (IBM), K. Terashi (ICEPP, UT)

High-energy physics (HEP)

- motivation: understanding elementary particles and their interactions
- described by gauge theory (U(1)×SU(2)×SU(3))
- quantum chromodynamics (QCD) at a low-energy scale
  - strong coupling → nonperturbative effects (ex. confinement)
  - in general hard to solve analytically…
- numerical simulation of gauge theory: lattice gauge theory
  - conventional Monte-Carlo method (in Lagrangian formalism): topological term, real-time dynamics etc.→ sign problem
  - Hamiltonian simulation (quantum simulation, tensor network)

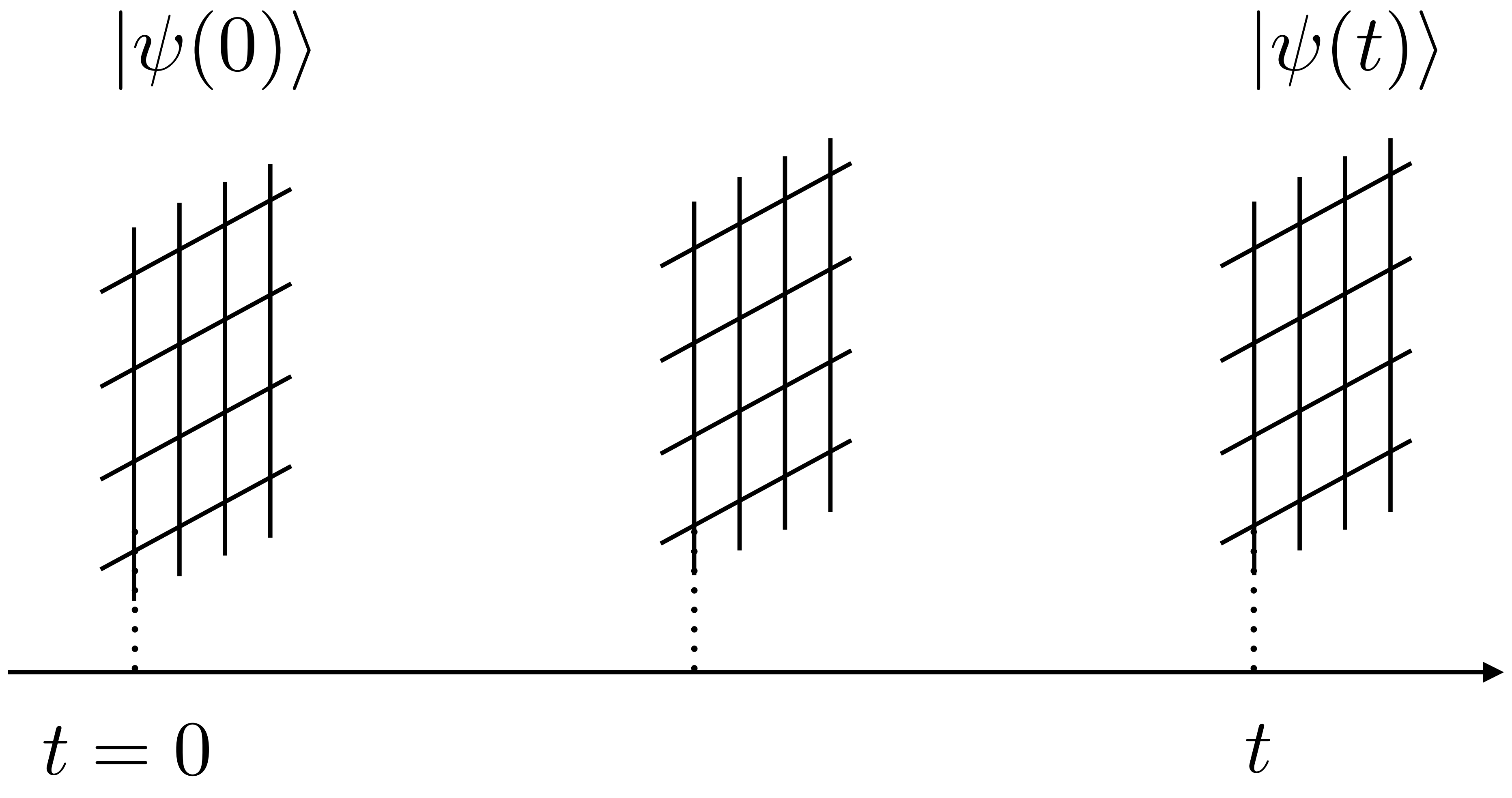


[figure from Wikipedia]

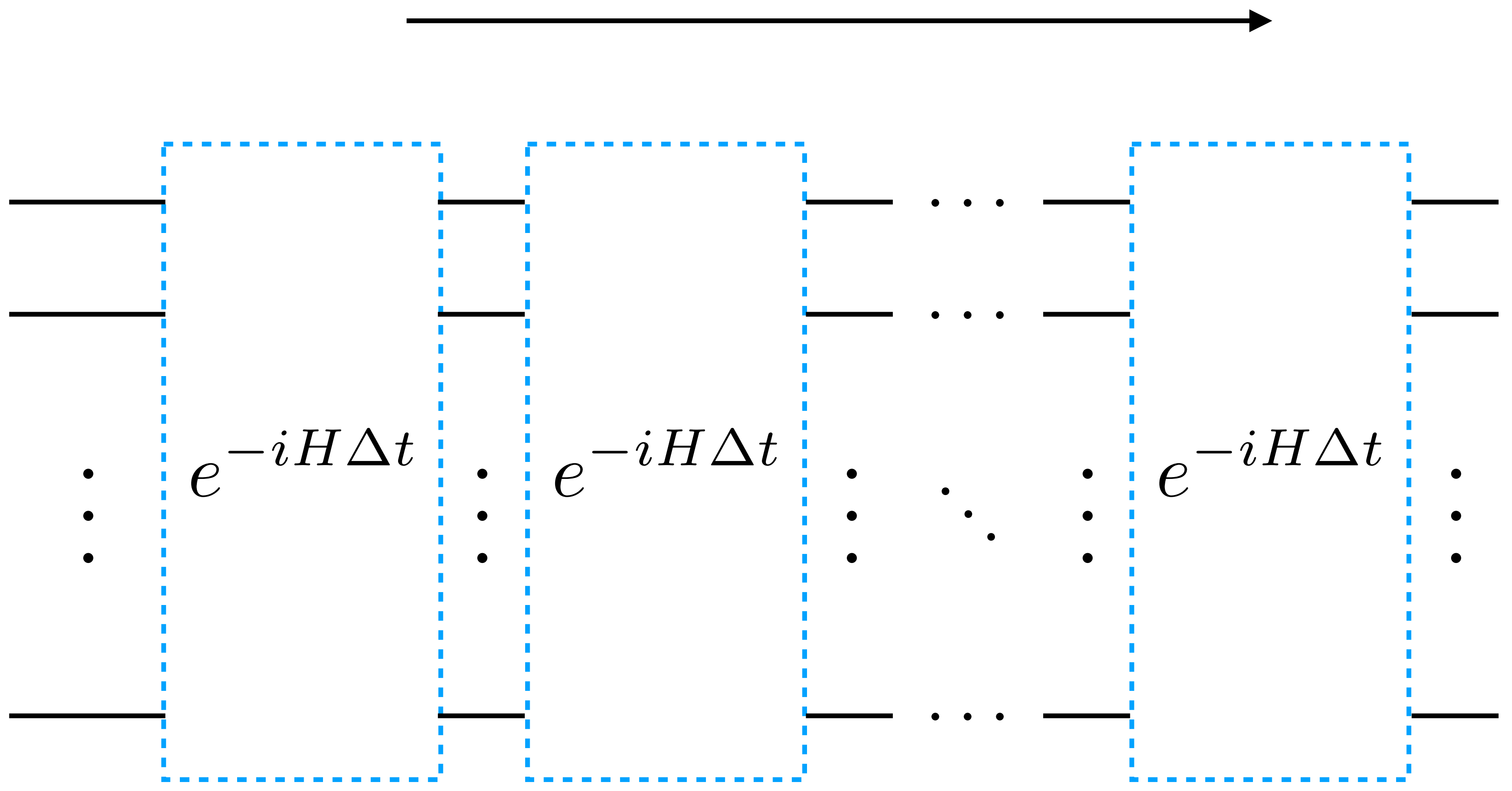


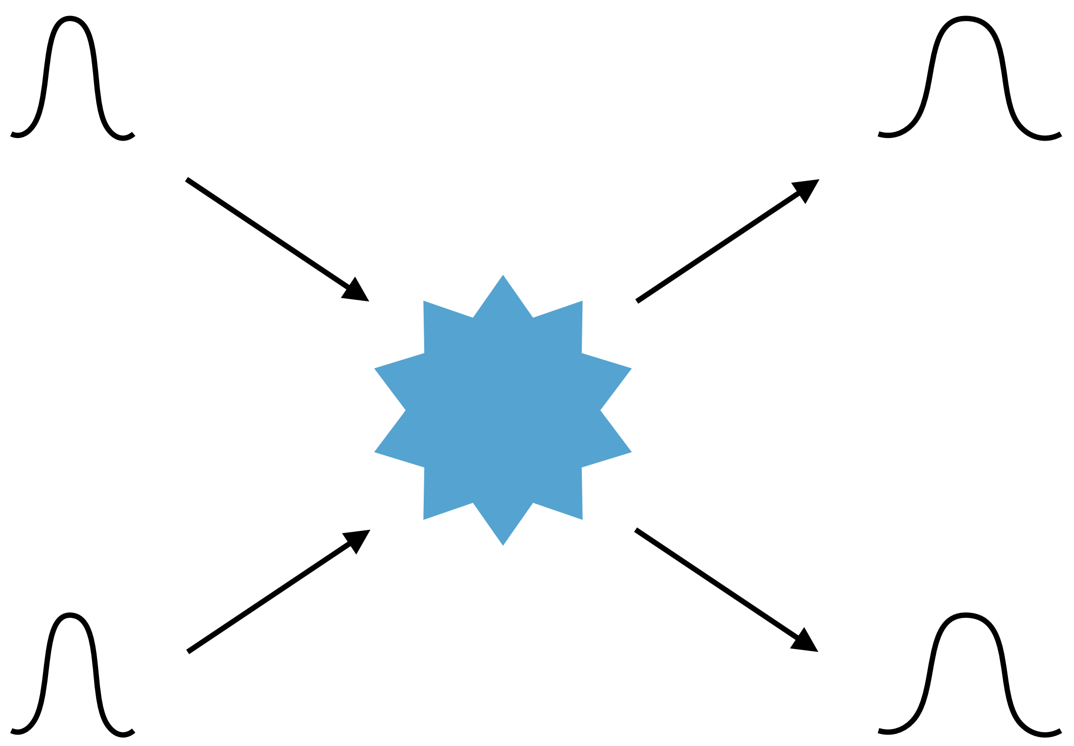
Quantum simulation in HEP

- Hamiltonian simulation using a quantum device
- applications to HEP [e.g. Jordan, Lee, Preskill, Science 336, 1130-1133 (2012)]
  - (ground/excited) state preparation
  - real-time dynamics (ex. scattering)
- how to extract physical information from final states?



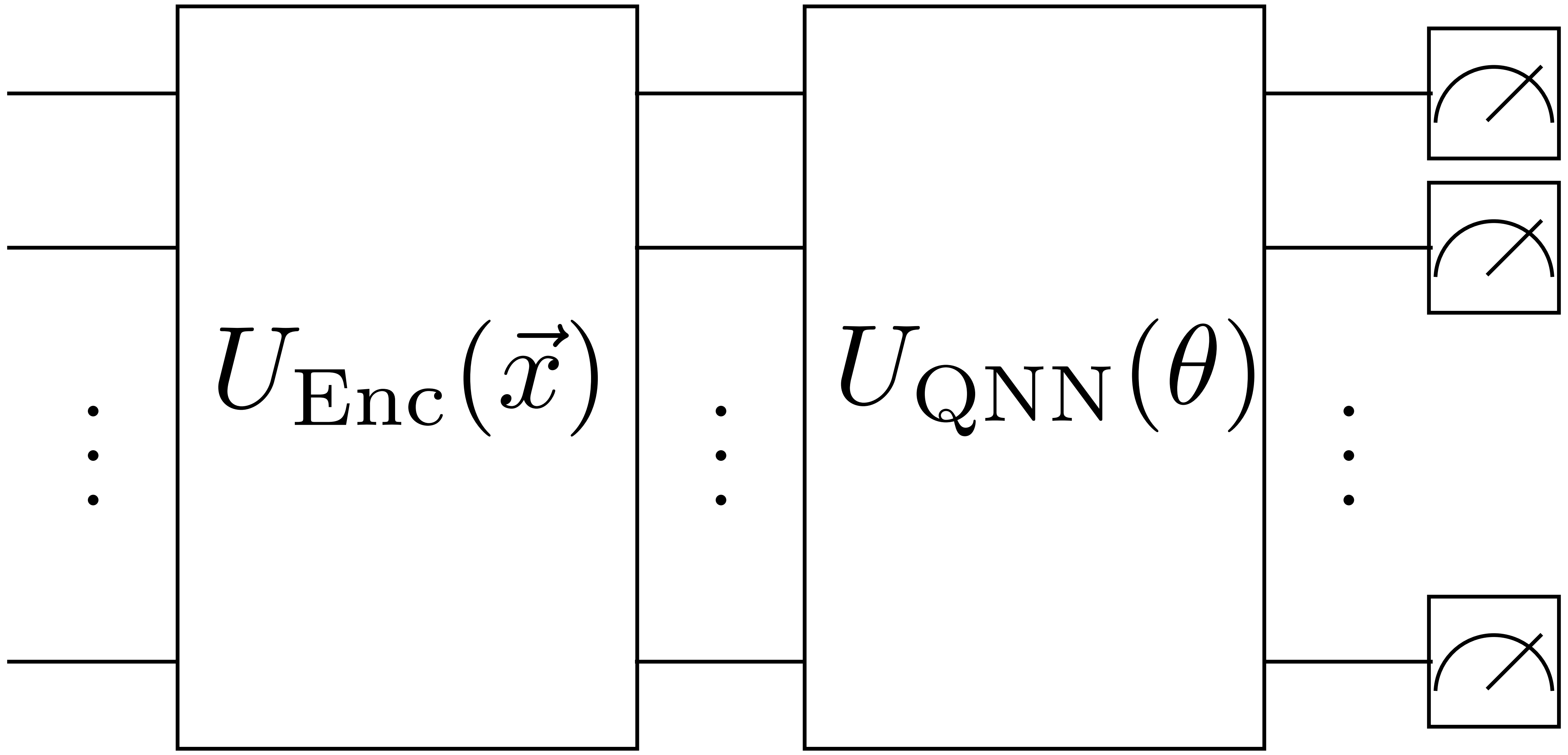
$$e^{-iHt} = \left(e^{-iH\Delta t}\right)^s$$



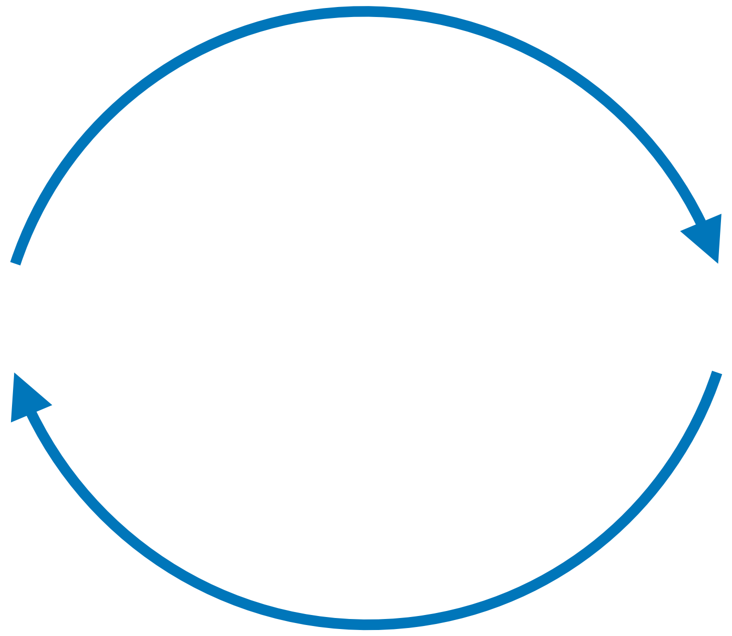


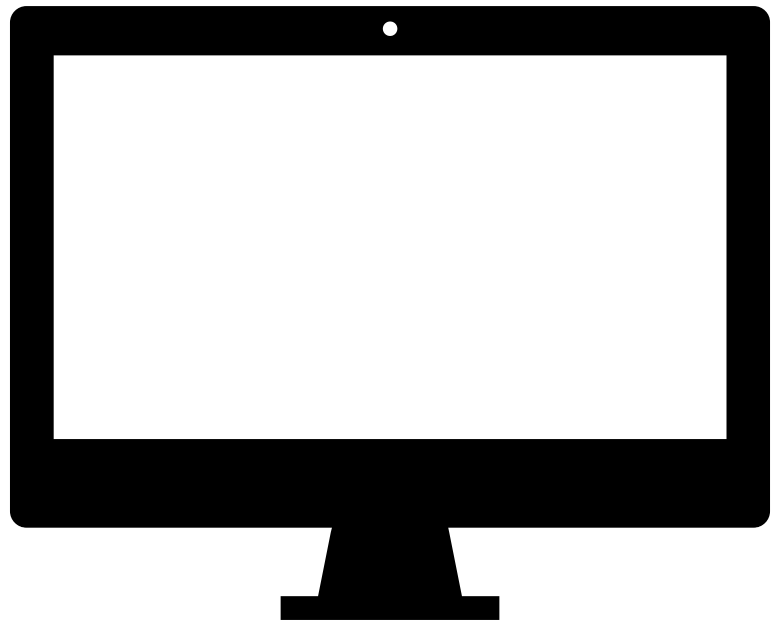
Quantum machine learning

- classical data learning:
  - training data $\{\vec{x}_i, y_i\}$ → predict a label y_i' associated with $\vec{x}_i'$
  - encoding data: $|\psi(\vec{x})\rangle = U_{\mathrm{Enc}}(\vec{x})|0\rangle$



$$\hat{y}(\theta) = \langle\psi|U^{\dagger}_{\mathrm{QNN}}(\theta) O U_{\mathrm{QNN}}(\theta)|\psi\rangle$$



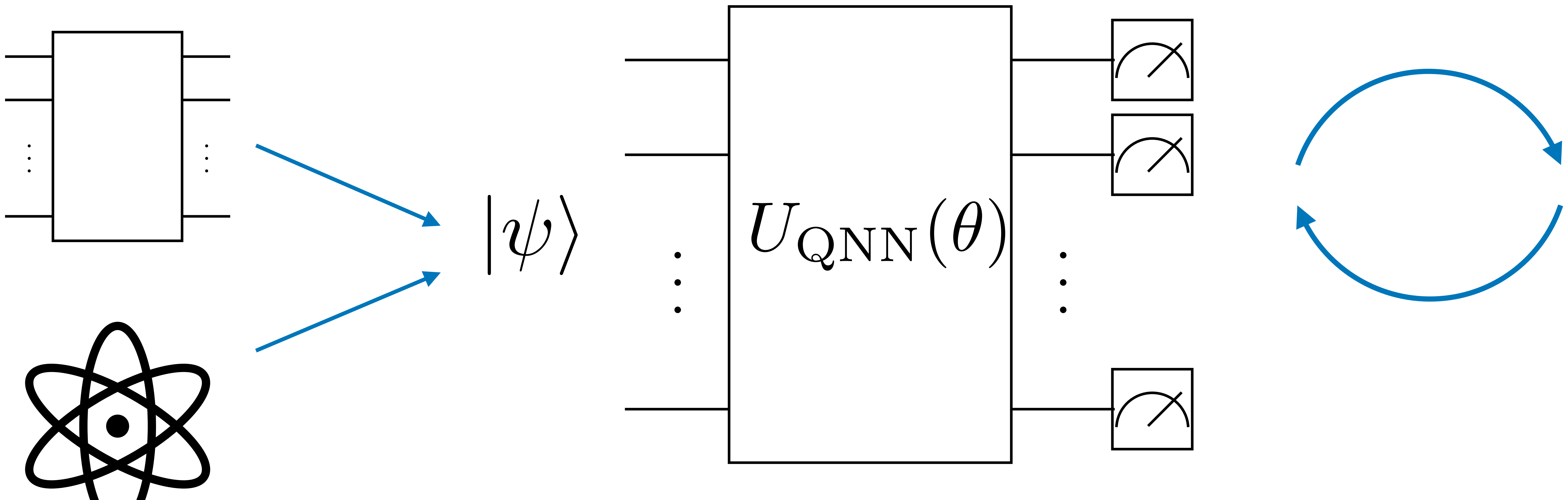


$$L(\theta) = \frac{1}{N_{\mathrm{data}}}\sum_i (y_i - \hat{y}_i)^2$$

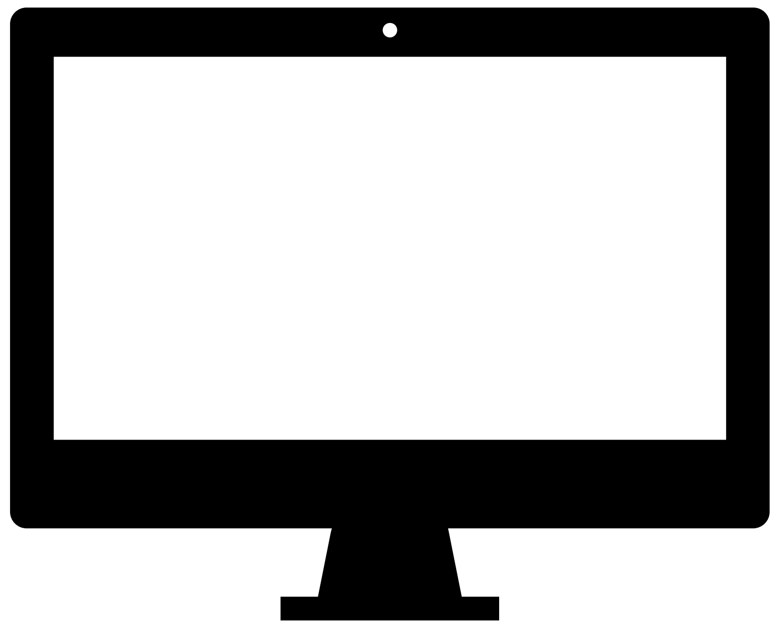
$$\theta_{\mathrm{opt}} = \arg\min_{\theta} L(\theta)$$

Quantum machine learning

- quantum data learning:
  - training data $\{|\psi_i\rangle, y_i\}$ → predict a label y_i' associated with $|\psi_i'\rangle$
  - data may come from
    - simulation using quantum circuit
    - another quantum devices
    - quantum sensor (?)



$$\hat{y}(\theta) = \langle\psi|U_{\mathrm{QNN}}^{\dagger}(\theta) O U_{\mathrm{QNN}}(\theta)|\psi\rangle$$



$$L(\theta) = \frac{1}{N_{\mathrm{data}}} \sum_i (y_i - \hat{y}_i)^2$$

$$\theta_{\mathrm{opt}} = \arg\min_{\theta} L(\theta)$$

Quantum convolutional neural network

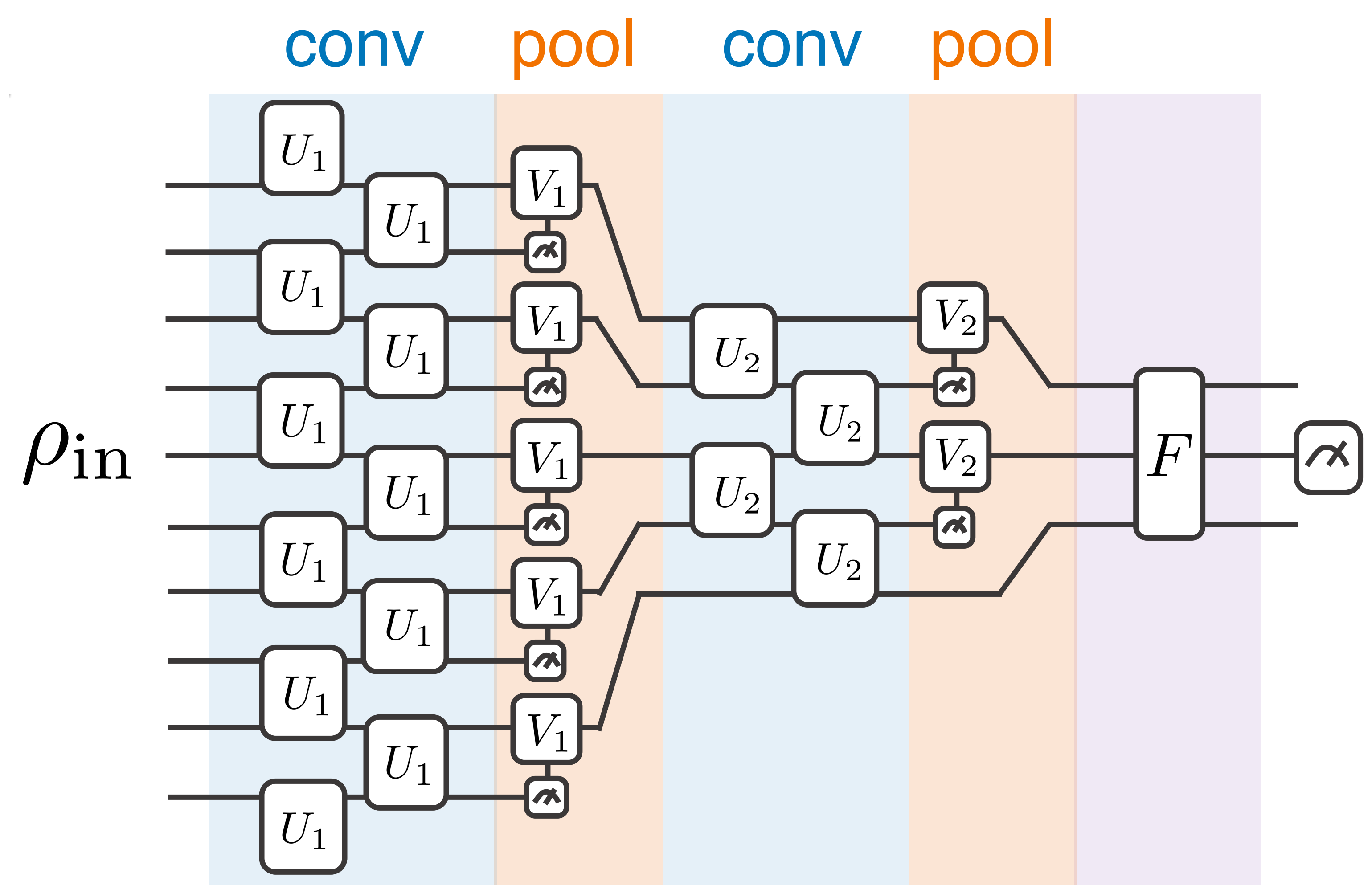
[Chong, Choi, Lukin, *Nat. Phys.* **15**, 1273–1278 (2019)]

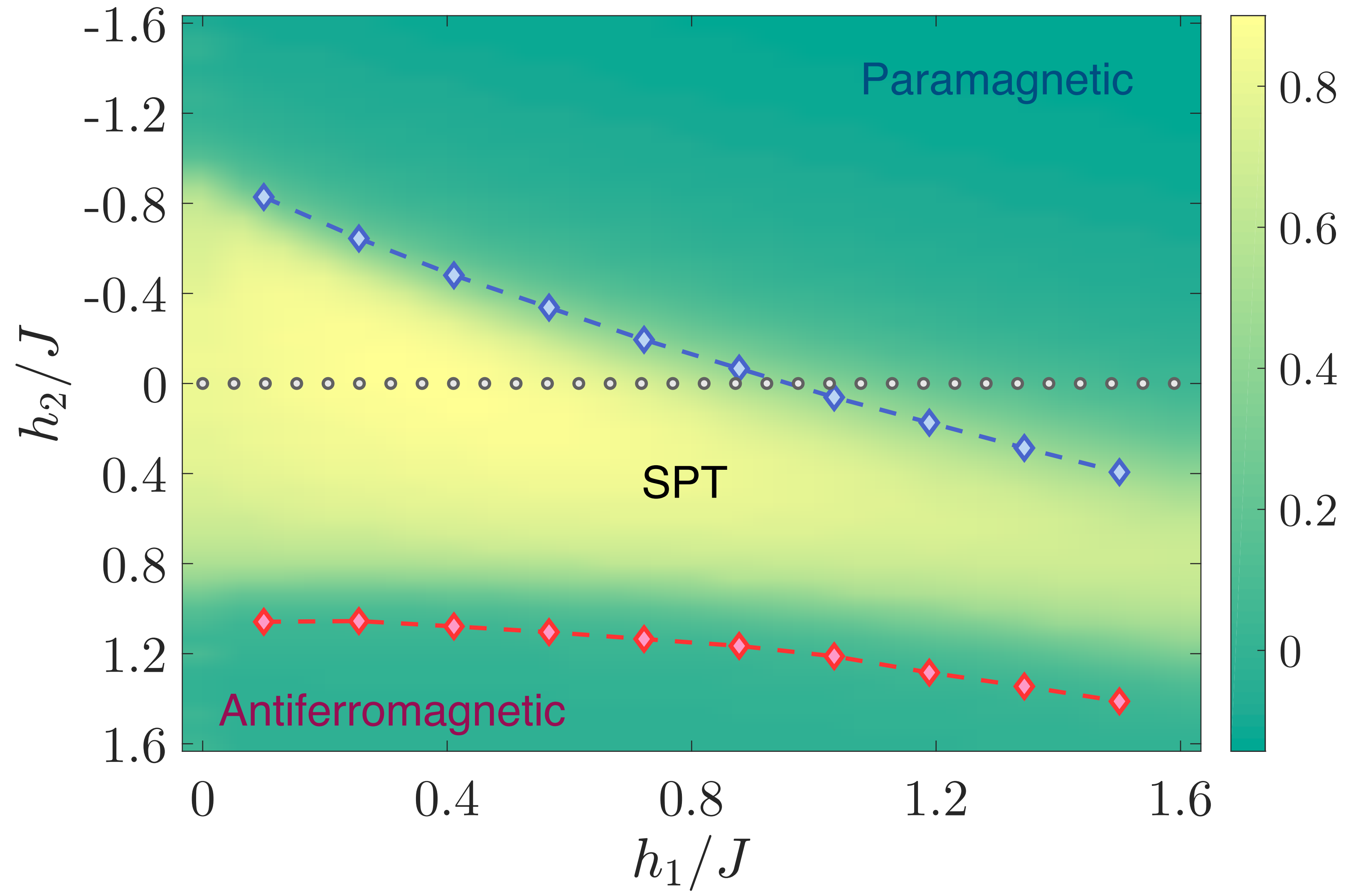
- Quantum convolutional neural network (QCNN)
  - convolution layer: e.g., general SU(4)
  - pooling layer: reduce the number of qubits
- useful application: phase recognition

$$H = -J\sum_i Z_i X_{i+1} Z_{i+2} - h_1 \sum_i X_i - h_2 \sum_i X_i X_{i+1}$$

  - data: ground states $|\mathrm{GS}(h_1, h_2)\rangle$
  - label: phases of states $y(h_1, h_2) = \pm 1$
  - output:

$$\hat{y}(h_1, h_2; \theta) = \langle \mathrm{GS}| U_{\mathrm{QCNN}}^{\dagger}(\theta) O U_{\mathrm{QCNN}}(\theta) |\mathrm{GS}\rangle$$



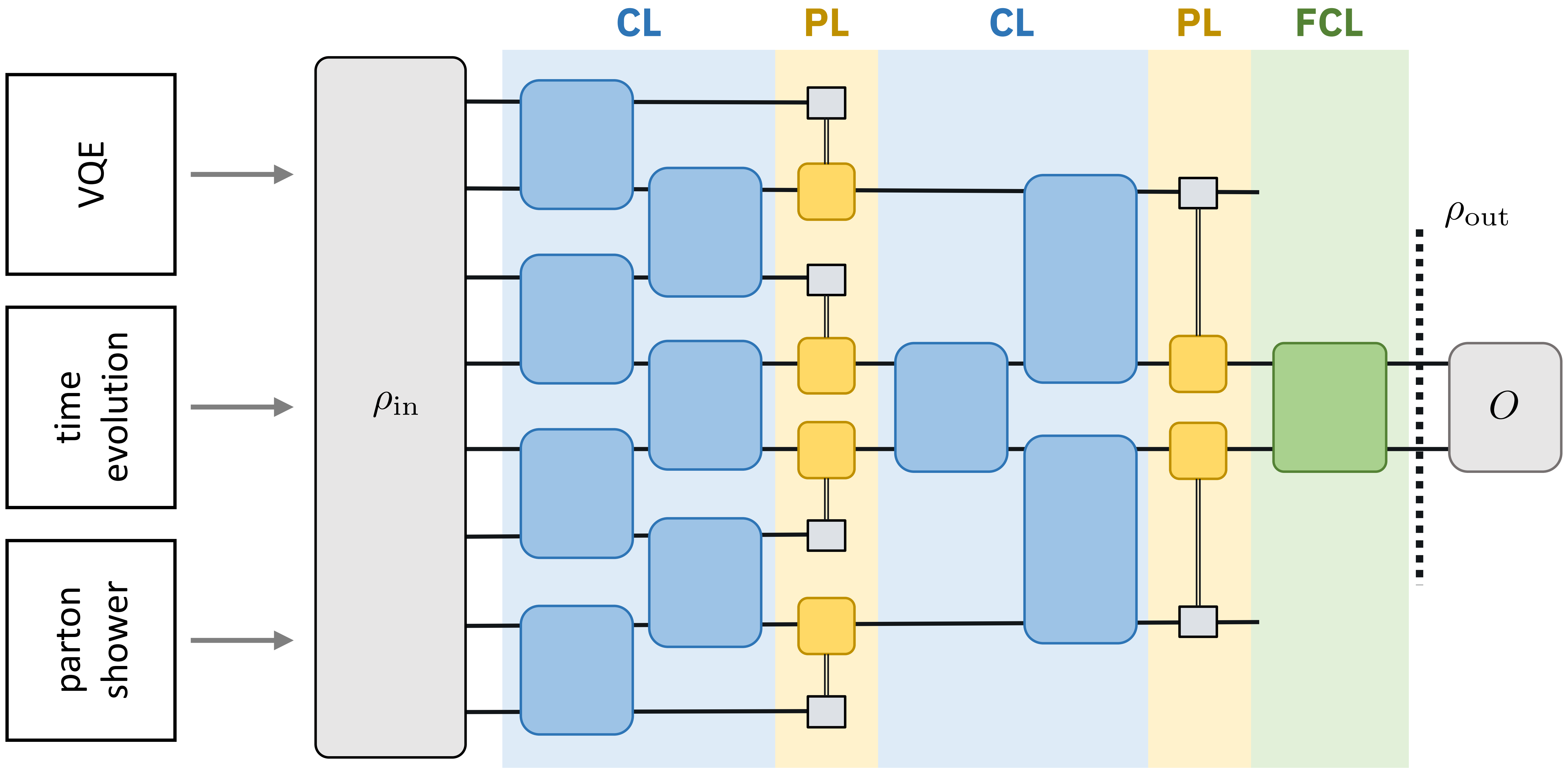


Quantum convolutional neural network

- nontrivial prediction ability for a specific spin model
- the advantage over the conventional method (direct measurement of order parameter) in terms of sampling complexity
- can avoid the barren plateau problem [A. Peshar, et.al., Phys. Rev. X **11**, 041011]
- open questions
  - application to other models?
  - quantum data beyond ground states?
  - practical advantage on near-term devices?

Summary of our results

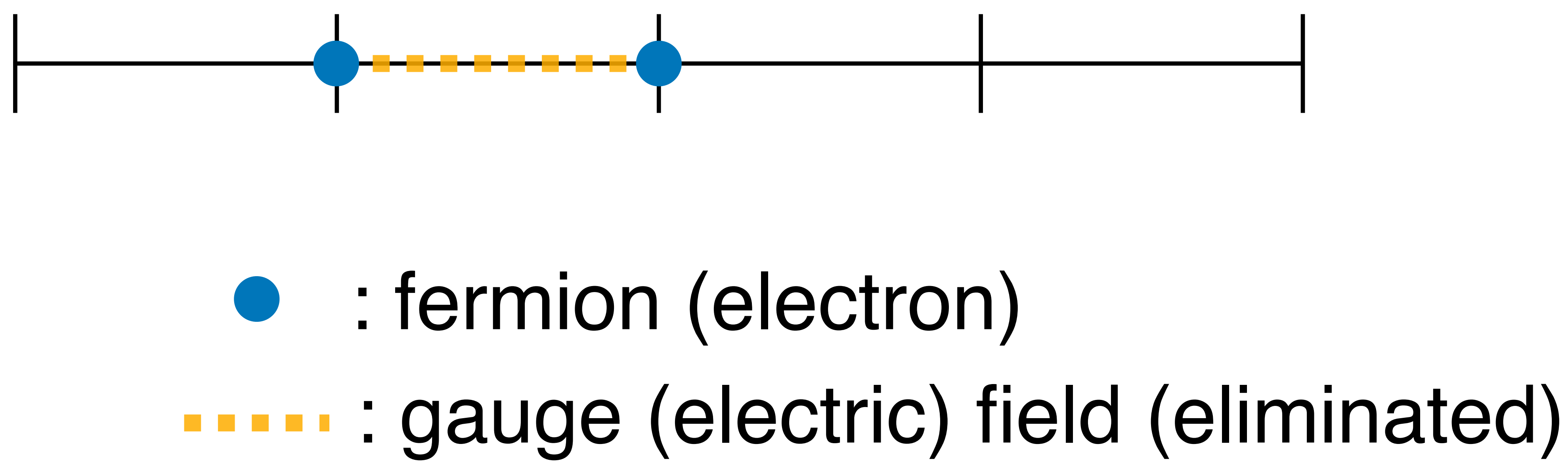
- investigate 3 applications of QCNN to HEP
  - 1d lattice gauge theory
  - phenomenological model
- obtain data from the quantum circuit
- single-qubit measurement in the end
- perform **statevector simulation**
  →showed nontrivial prediction ability

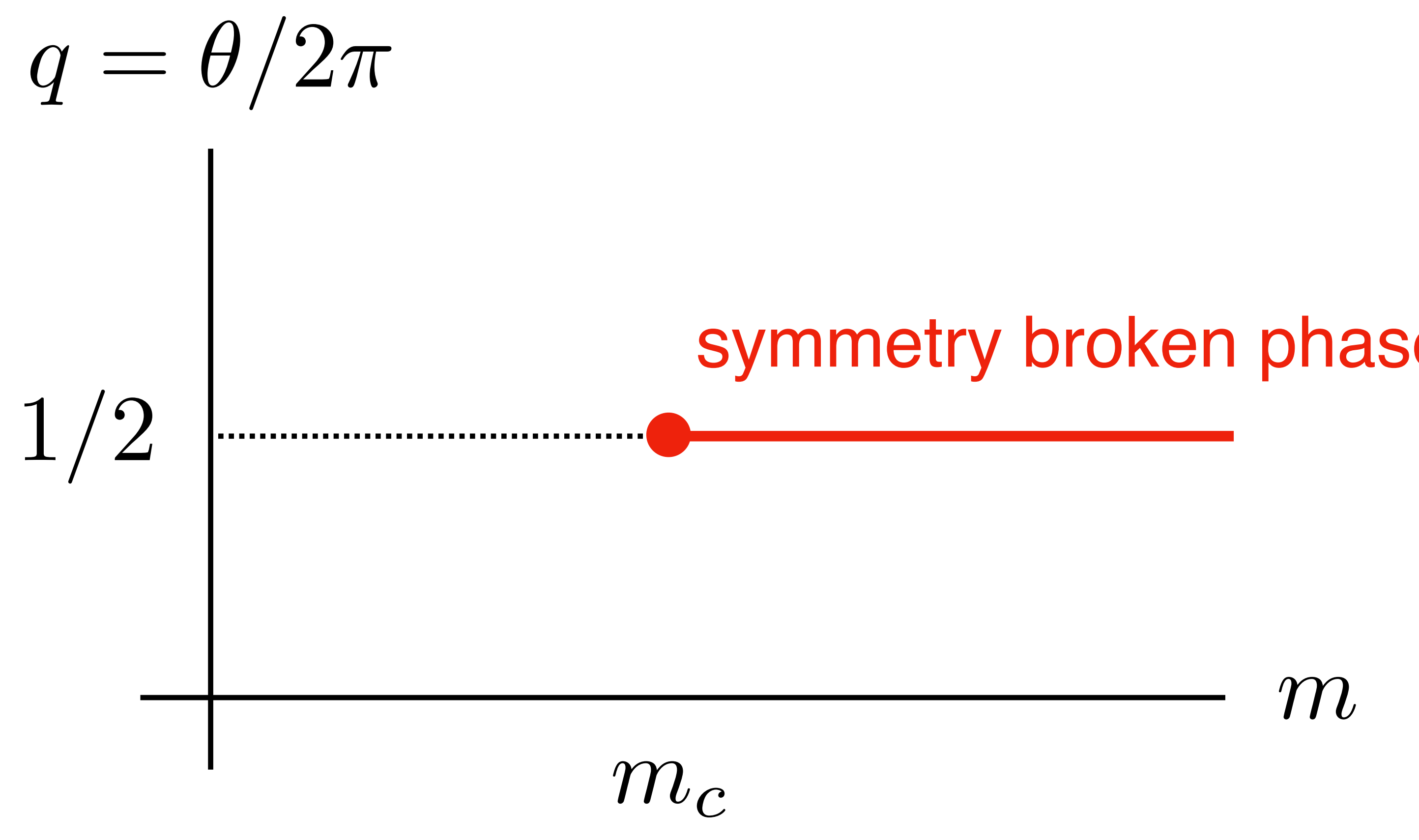


Schwinger model (mapped to spin model)

$$H_{\text{spin}} = J \sum_{n=0}^{N-2} \left(\sum_{k=0}^{n} \frac{Z_k + (-)^k}{2} + q \right)^2 + \frac{w}{2} \sum_{n=0}^{N-2} (X_n X_{n+1} + Y_n Y_{n+1}) + \frac{m}{2} \sum_{n=0}^{N-1} (-)^n Z_n$$

- simple gauge theory: 1+1d quantum electrodynamics: Schwinger model [Schwinger, Phys. Rev. 128, 2425, (1962)]
- still nontrivial: confinement, topological term (cannot be treated by MC method)
- phase transition at $q = 1/2$ and $m = m_c$





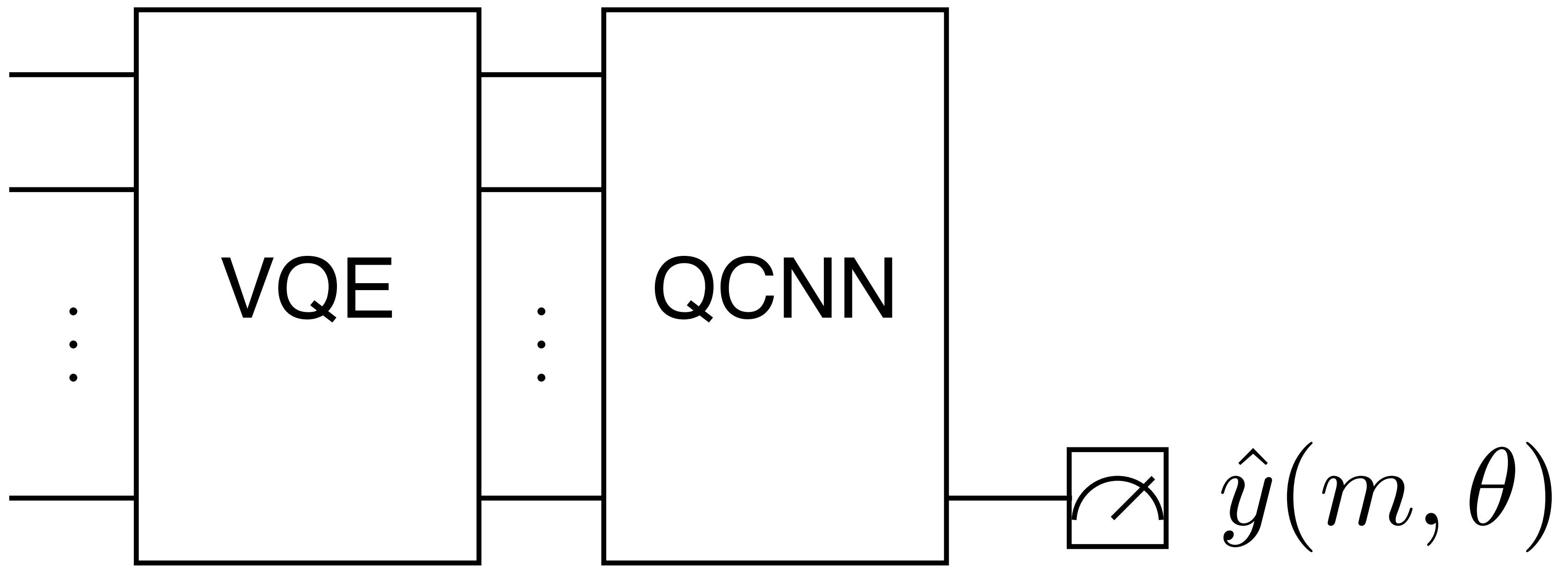
Classification of ground states

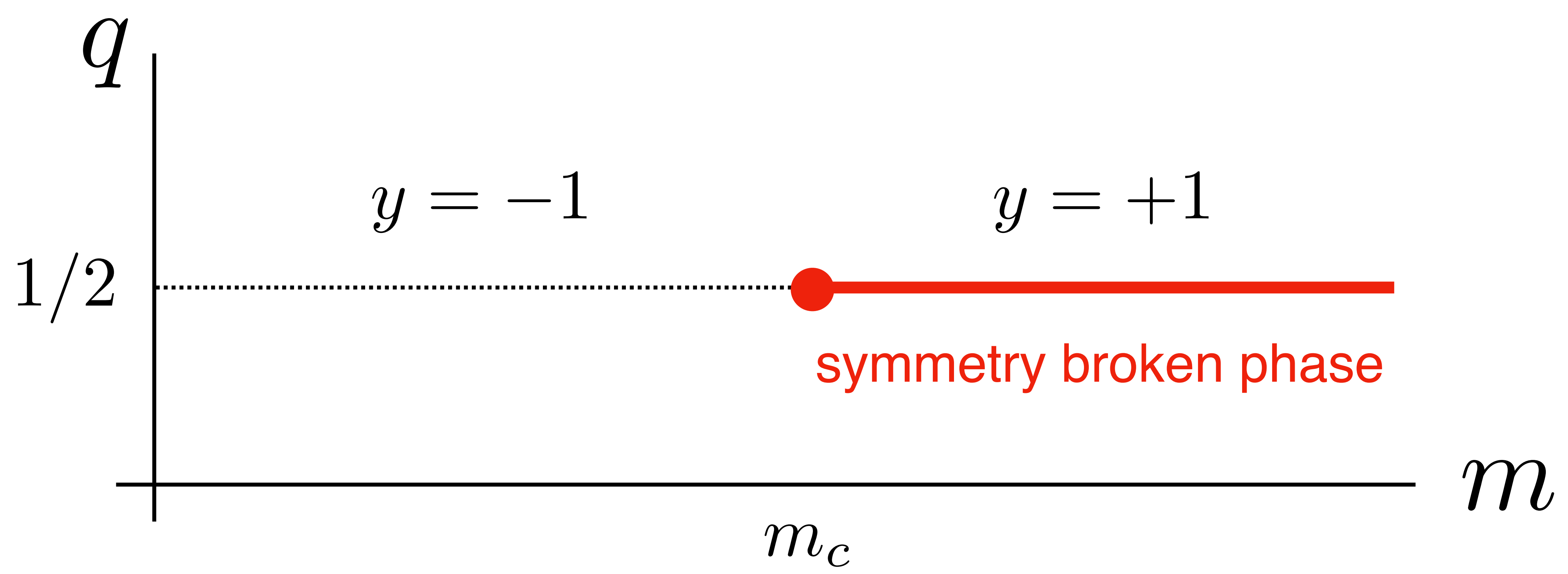
- physical parameters: $N = 8, q = 1/2$
- data: ground states $|\psi_{\mathrm{GS}}(m)\rangle$ from VQE
- label:

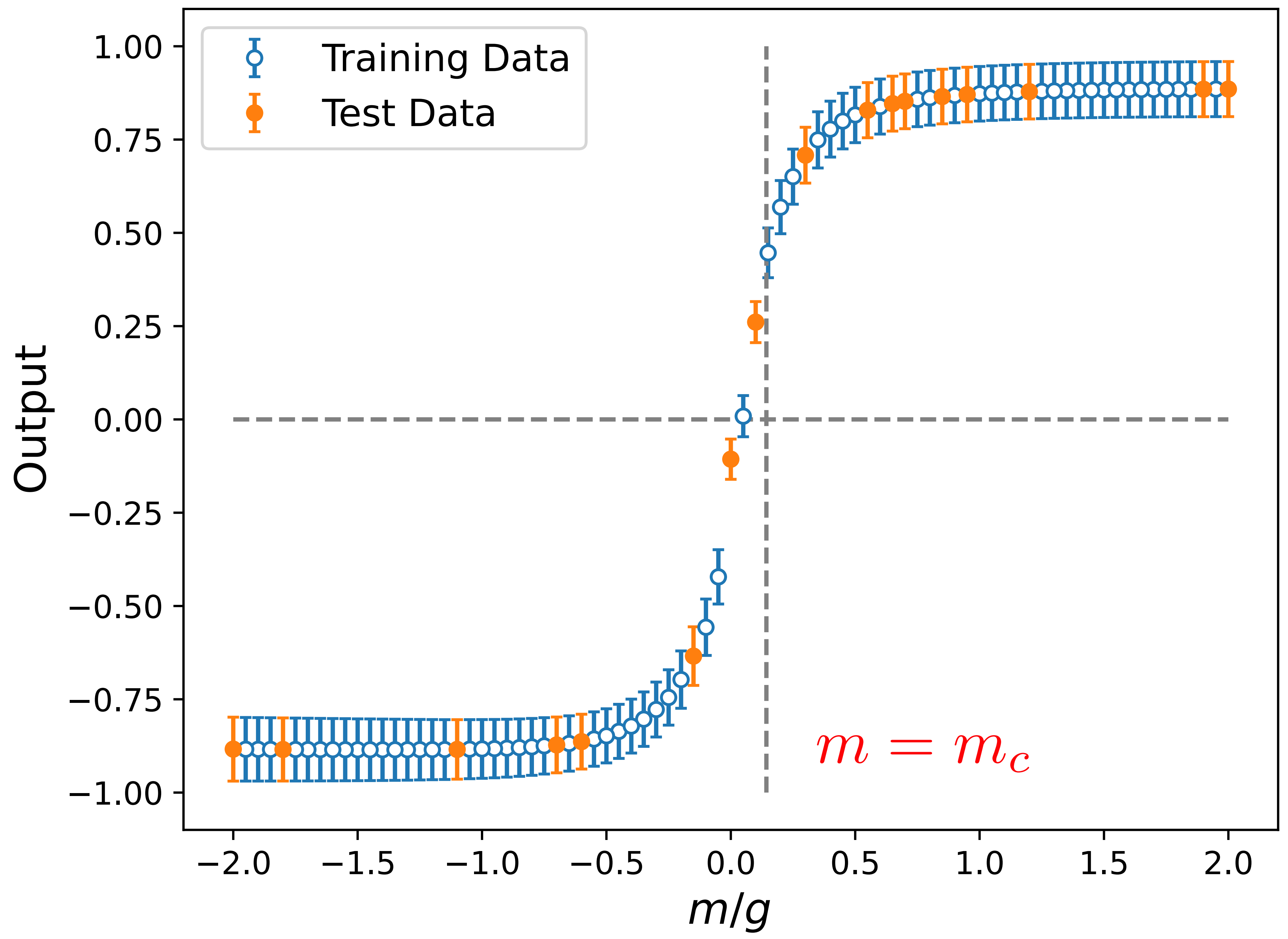
$$y(m) = \begin{cases} +1 & m > m_c \\ -1 & m < m_c \end{cases}$$

- minimize the mean squared error

$$L(\theta) = \frac{1}{N_{\mathrm{data}}} \sum_i \left(y(m) - \hat{y}(m, \theta)\right)^2$$



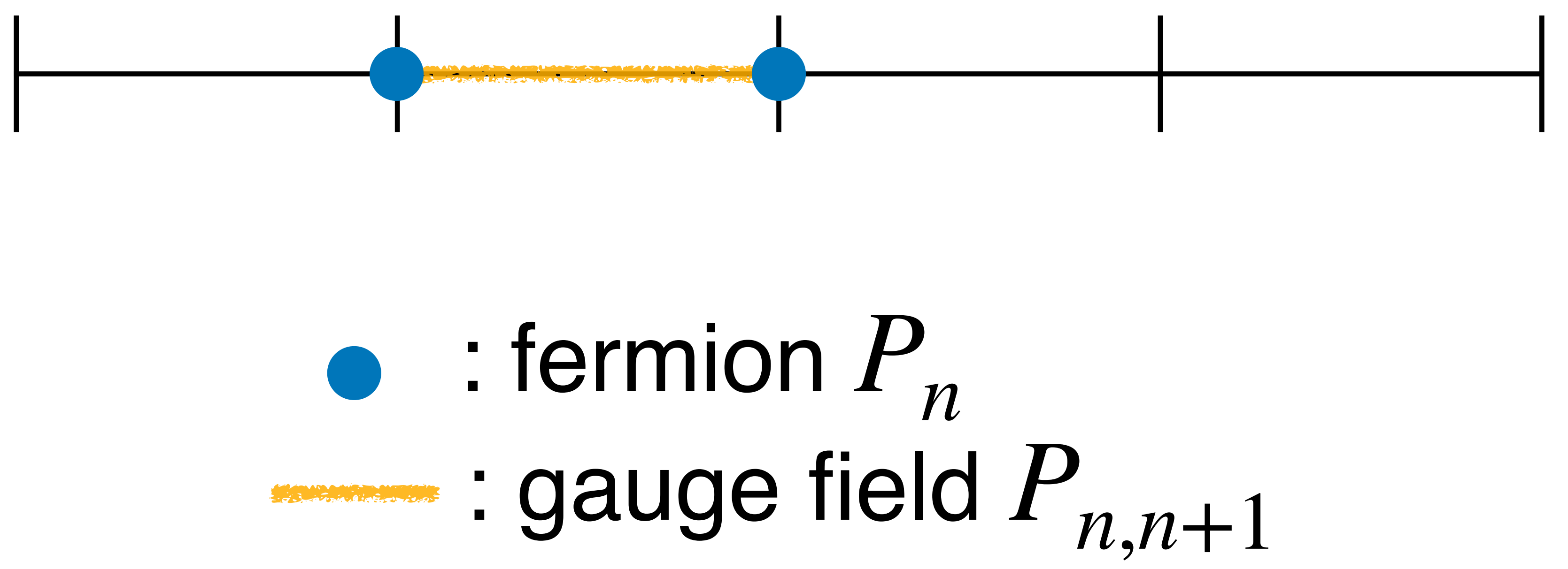




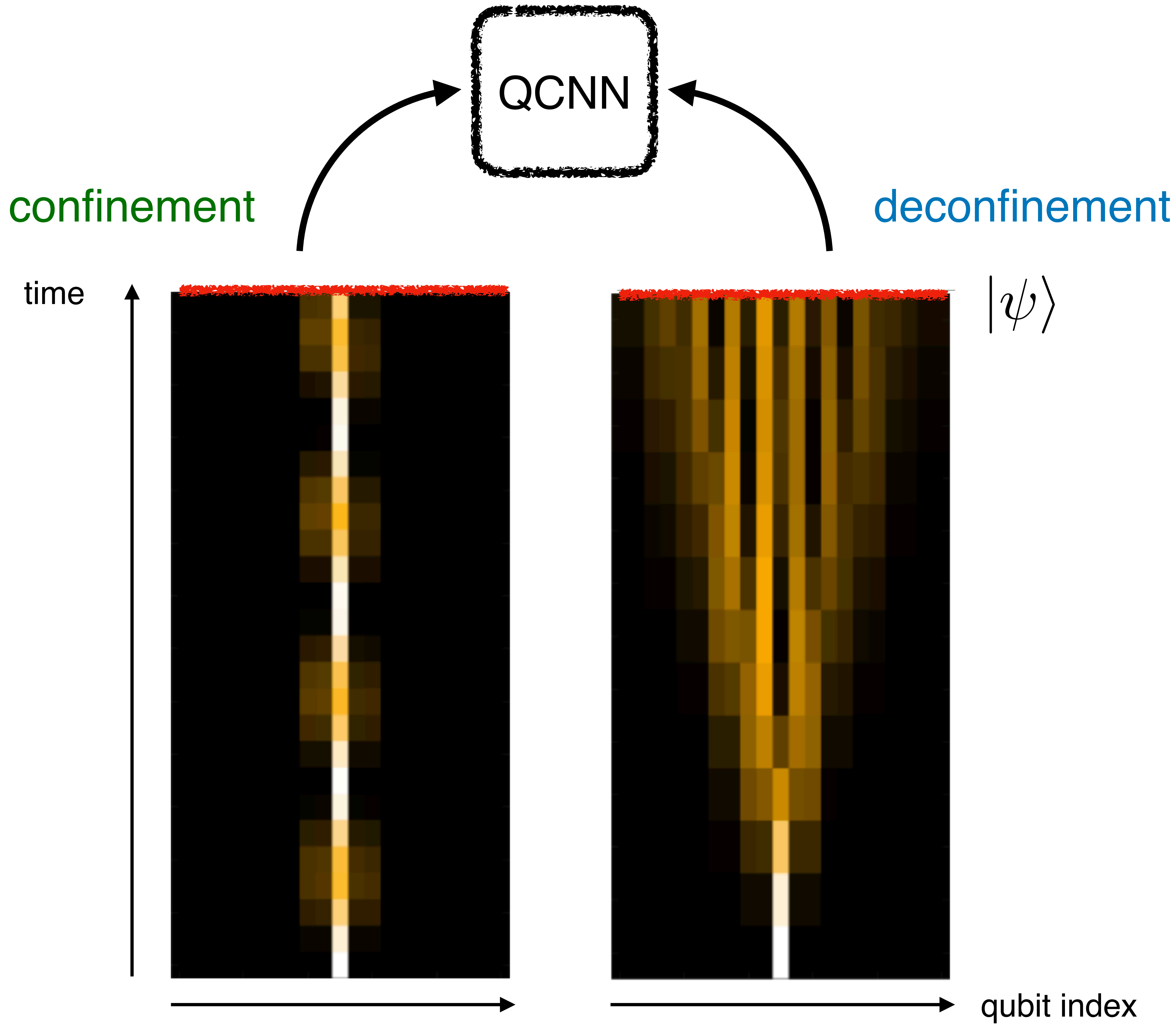
(1+1) d $\mathbb{Z}_2$ gauge theory

$$H = -\frac{J}{2}\sum_{n=0}^{N_s-1}(X_n Z_{n,n+1} X_{n+1} + Y_n Z_{n,n+1} Y_{n+1}) - f\sum_{n=0}^{N_s-1} X_{n,n+1} + \frac{m}{2}\sum_{n=0}^{N_s-1}(-)^n Z_n$$

- qualitatively different behavior depending on f
  - confinement for $f \neq 0$
  - deconfinement for $f = 0$
- Can QCNN distinguish these two after time evolution?



[the left figure; from Mildenberger et al., arXiv:2203.08905, (2022)]



Classification of time-evolved states

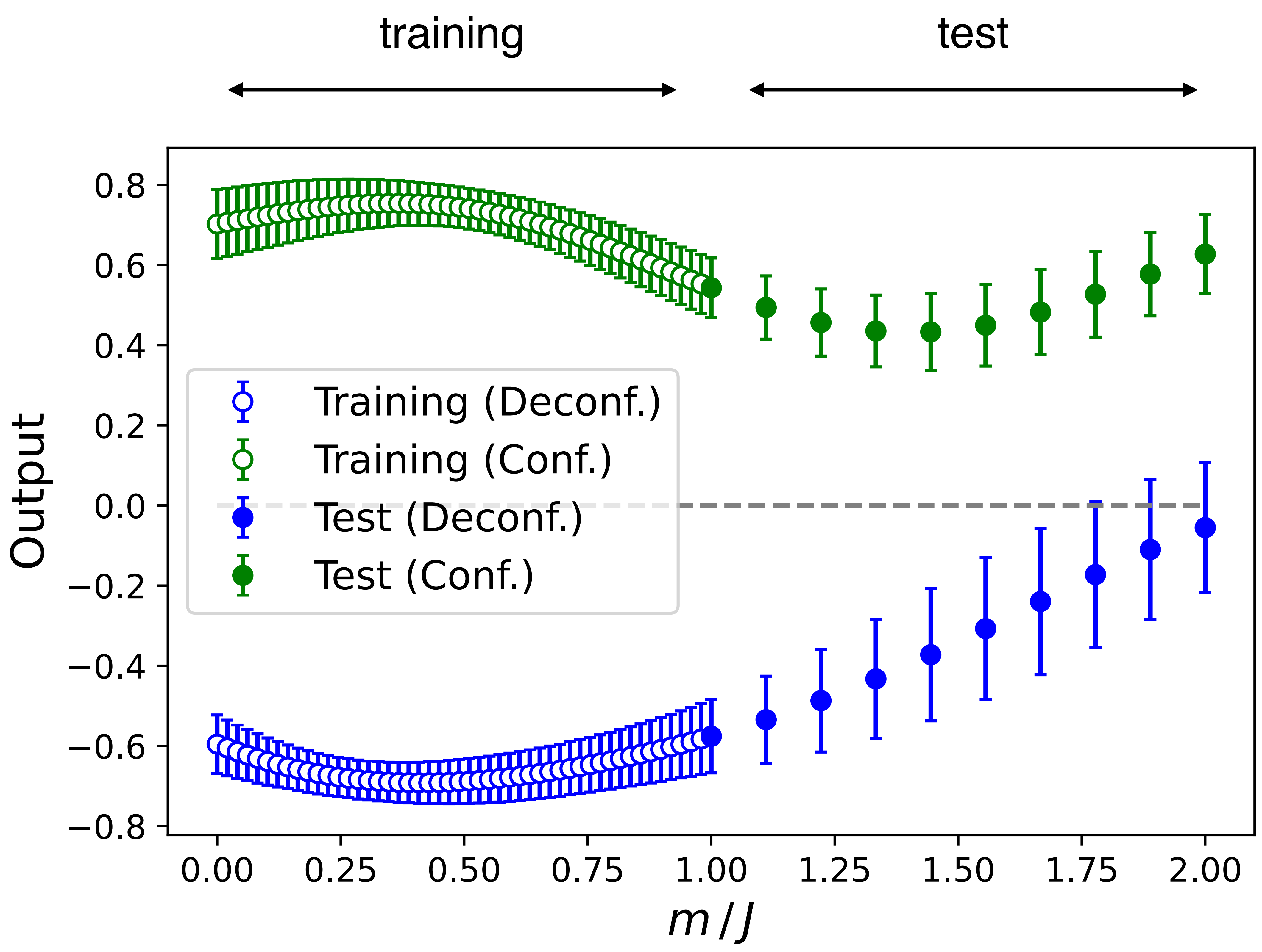
[Our results]

$$H = -\frac{J}{2}\sum_{n=0}^{N_s-1}(X_n Z_{n,n+1}X_{n+1} + Y_n Z_{n,n+1}Y_{n+1}) - f\sum_{n=0}^{N_s-1} X_{n,n+1} + \frac{m}{2}\sum_{n=0}^{N_s-1}(-)^n Z_n$$

- physical parameters: $N_s = 2$
- input data:
  - time evolved state $|\psi(m,f)\rangle = e^{-iH(m,f)T}|\psi_0\rangle$ obtained from Suzuki-Trotter decomposition
  - split into test/train data by the value of mass
- label:

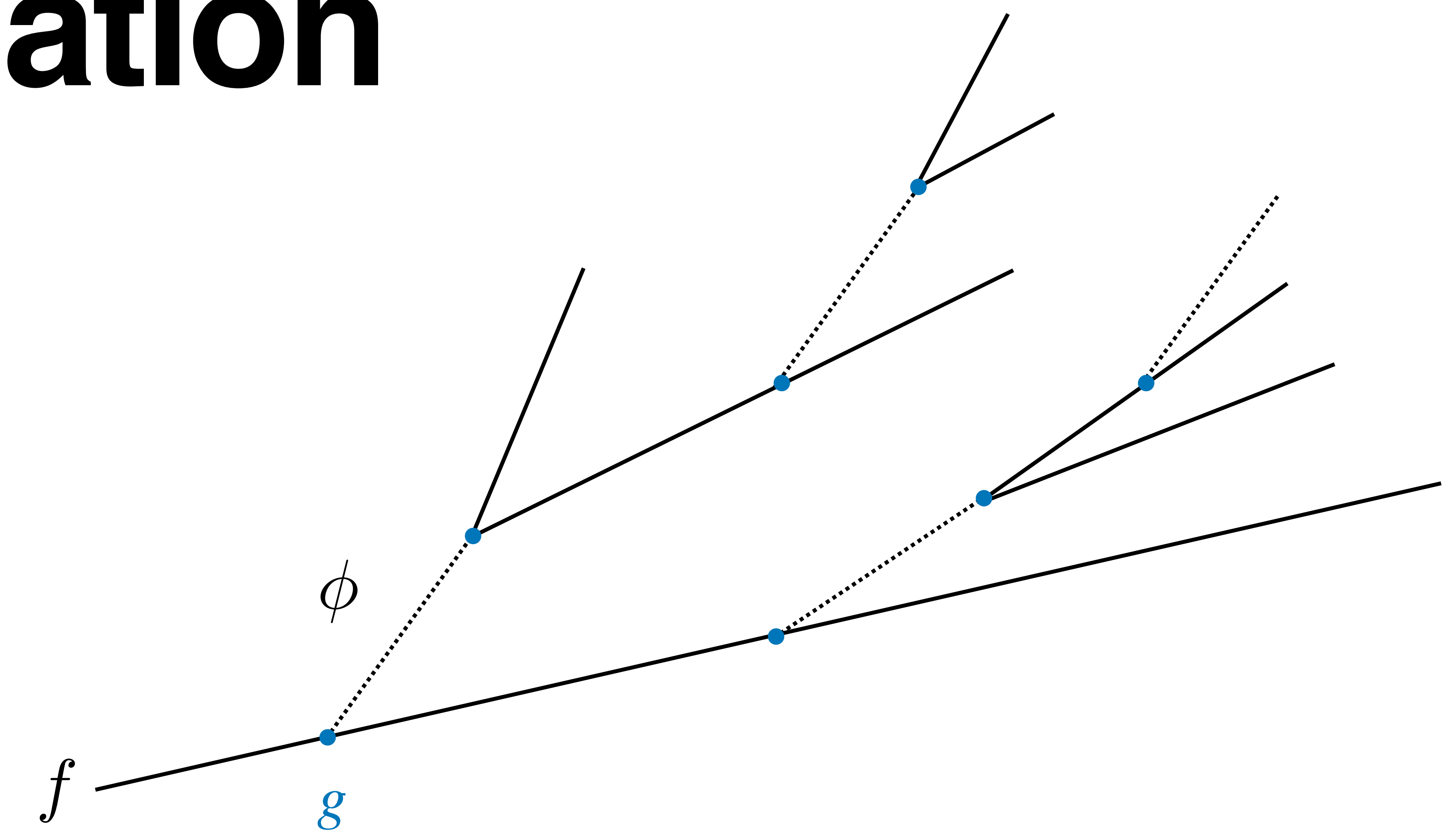
$$y(m,f) = \begin{cases} +1 & f \neq 0 \quad \text{confine} \\ -1 & f = 0 \quad \text{deconfine} \end{cases}$$

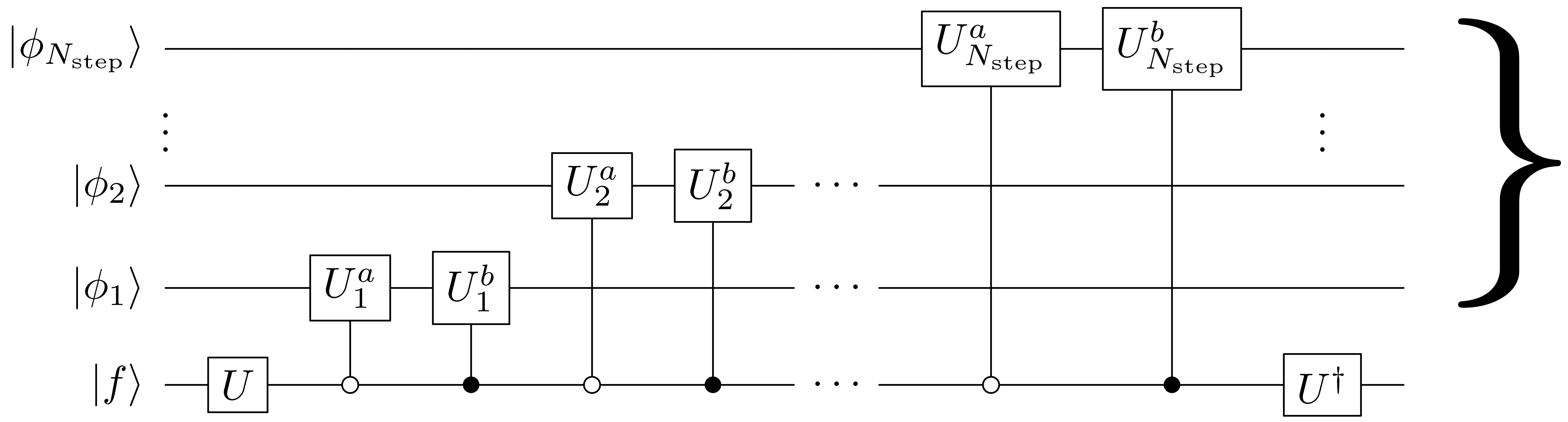
- correct prediction for $m \lesssim 1.75$



Parton shower (PS) simulation

- simple model: fermion f and boson ϕ
- [$\phi \to f\bar{f}$, or $f \to \phi f$, or nothing] at each step
- quantum circuit for simulation of a simplified model
  [Nachman et al., Phys. Rev. Lett. 126, 062001, (2021)]
  - simplest case: no $\phi \to f\bar{f}$
  - coupling: g
  - the number of steps: N_{step}(=maxmial number of ϕ-bosons at the final state)

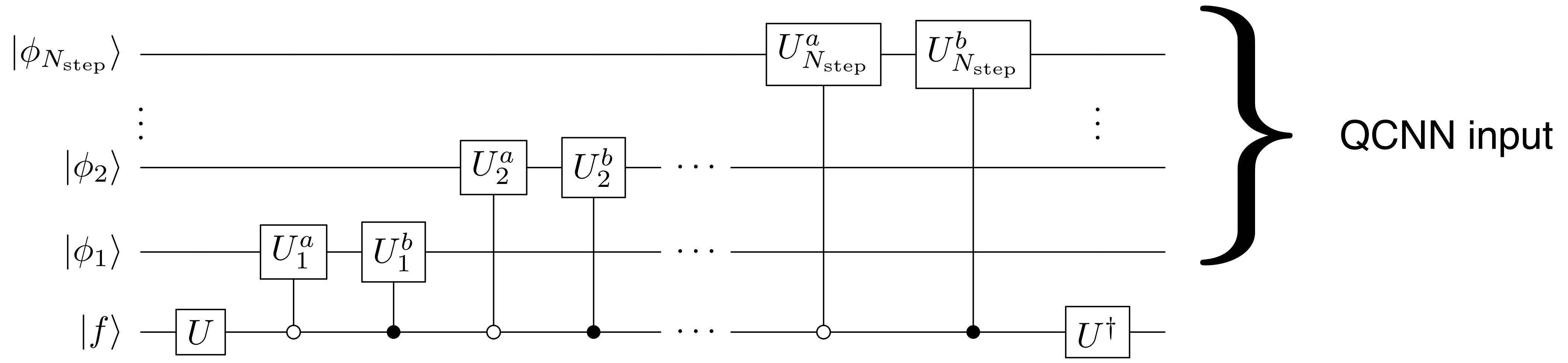


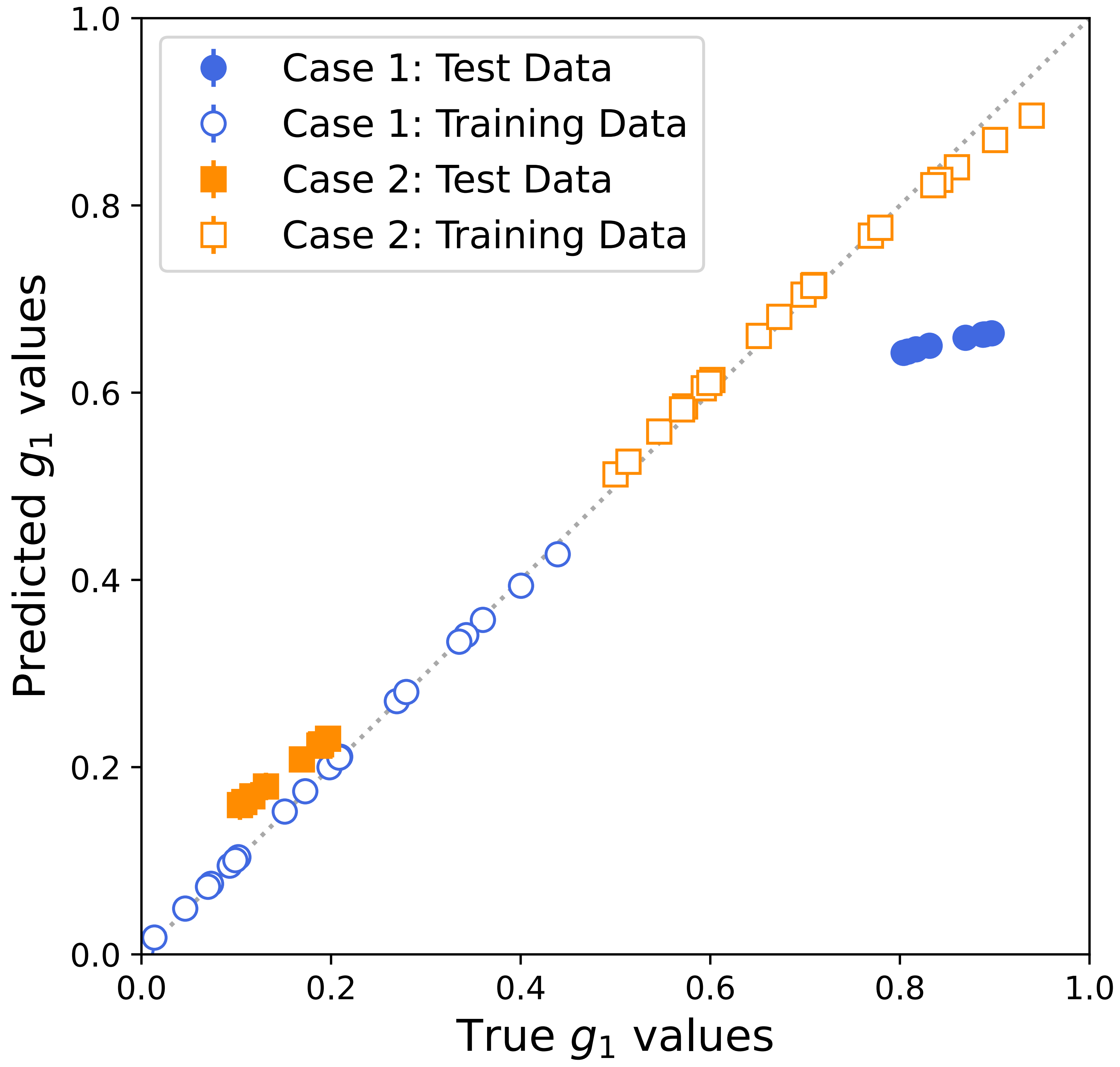


QCNN input
→prediction of coupling g

Coupling regression in PS simulation [Our results]

- the number of steps: $N_{\text{step}} = 8$
- case 1:
  - training data $g \in [0,0.5]$
  - test data $g \in [0.8,0.9]$
- case 2:
  - training data $g \in [0.5,1]$
  - test data $g \in [0.1,0.2]$
- prediction with good accuracy
- small offset for case 1





Summary and future direction

- we study QCNN applications to HEP
    - classification of ground states in the Schwinger model
    - classification of time-evolved states in 1+1d Z2 gauge theory
    - coupling regression in a parton shower simulation
- single qubit measurements after QCNN
- statevector simulation → nontrivial prediction with good accuracy
- Future direction
    - quantitative estimation of sampling complexity
    - the effect of quantum noise, implementation on real hardware
    - bounds for generalization errors?
    - extension to higher-dimensional gauge theory

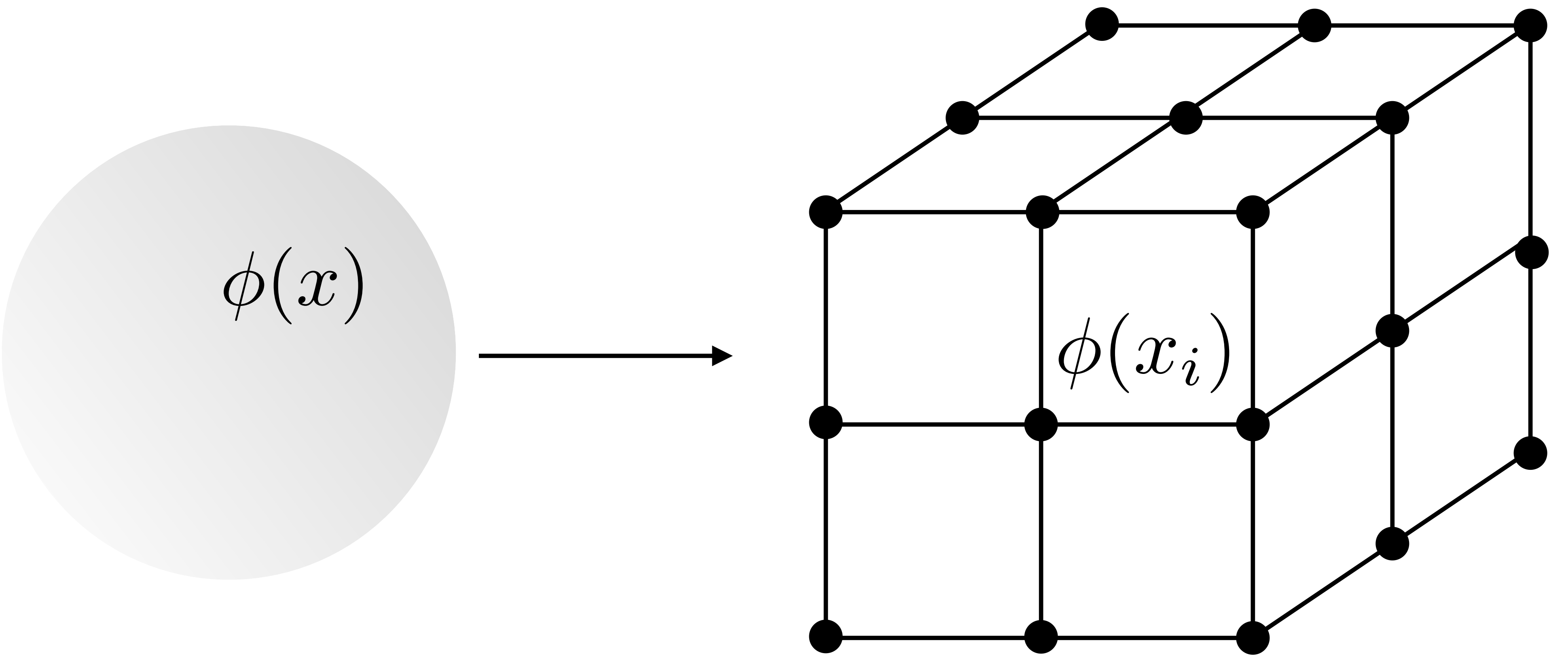
Backups

Lattice gauge theory

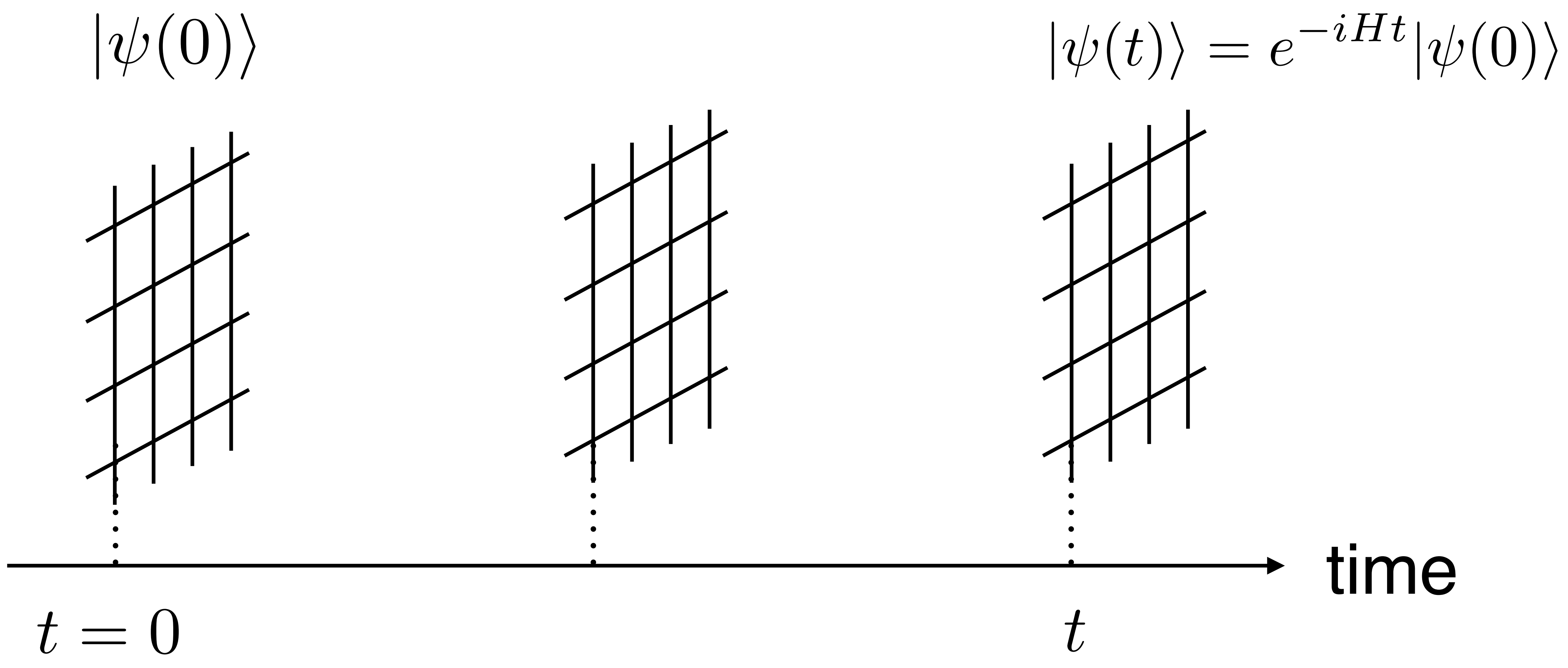
- (conventional) lattice gauge theory
  - discretize **spacetime**
    →using Monte Carlo method

$$Z = \int [d\phi] e^{-S[\phi]} \rightarrow \sum_{\{\phi_i\}} e^{-S(\phi_i)}$$

  - infamous sign problem
    - topological term
    - real-time dynamics, etc.



- Hamiltonian simulation
  - discretize **space**
  - no sign problem!
  - need exponential resources…
    - quantum simulation
    - tensor network, etc.



Schwinger model

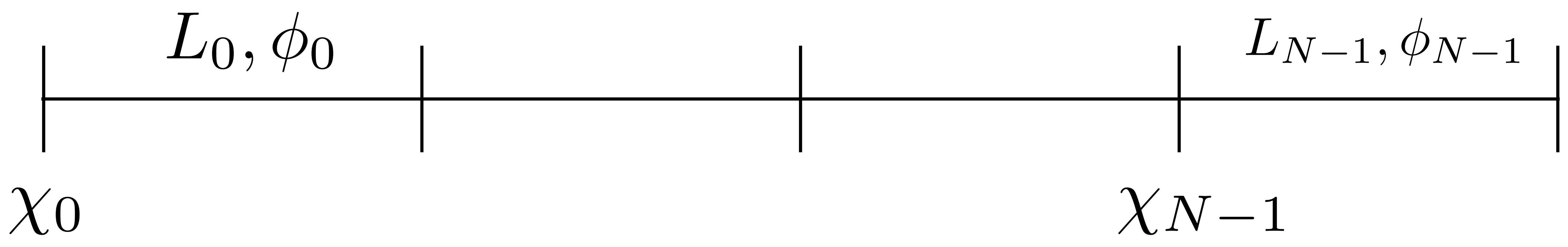
- ultimate goal: 3+1d non-Abelian gauge theory
- simple toy model: 1+1d U(1) gauge theory = **Schwinger model** [Schwinger]

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \boxed{\frac{g\theta}{4\pi}\epsilon^{\mu\nu}F_{\mu\nu}} + \mathrm{i}\bar{\psi}\gamma^{\mu}\left(\partial_{\mu} + \mathrm{i}gA_{\mu}\right)\psi - m\bar{\psi}\psi$$

- exactly solvable when $m = 0$
- simple but still non-trivial
  - screening/confinement transition
  - we can include topological term (cannot be treated in MC method)

Lattice Hamiltonian of Schwinger model

- χ_n: staggered fermion [Susskind, Kogut-Susskind]
- L_n, ϕ_n: link variables (gauge field)



$$H_{\text{lat}} = J \sum_{n=0}^{N-2} \left(L_n + \frac{\theta}{2\pi} \right)^2 - \mathrm{i}w \sum_{n=0}^{N-2} \left(\chi_n^\dagger e^{\mathrm{i}\phi_n} \chi_{n+1} - \text{c.c.} \right) + m \sum_{n=0}^{N-1} (-)^n \chi_n^\dagger \chi_n$$

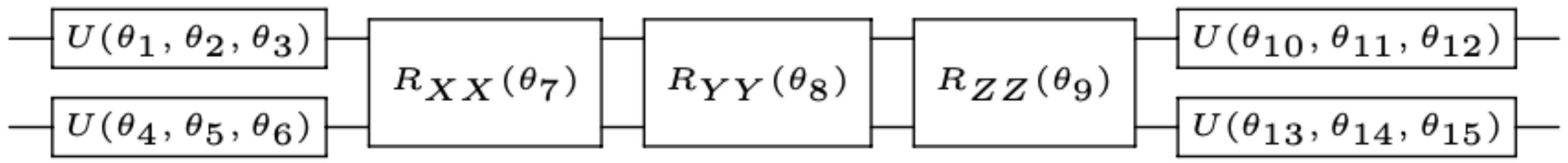
- gauge invariance: Gauss's law constraint

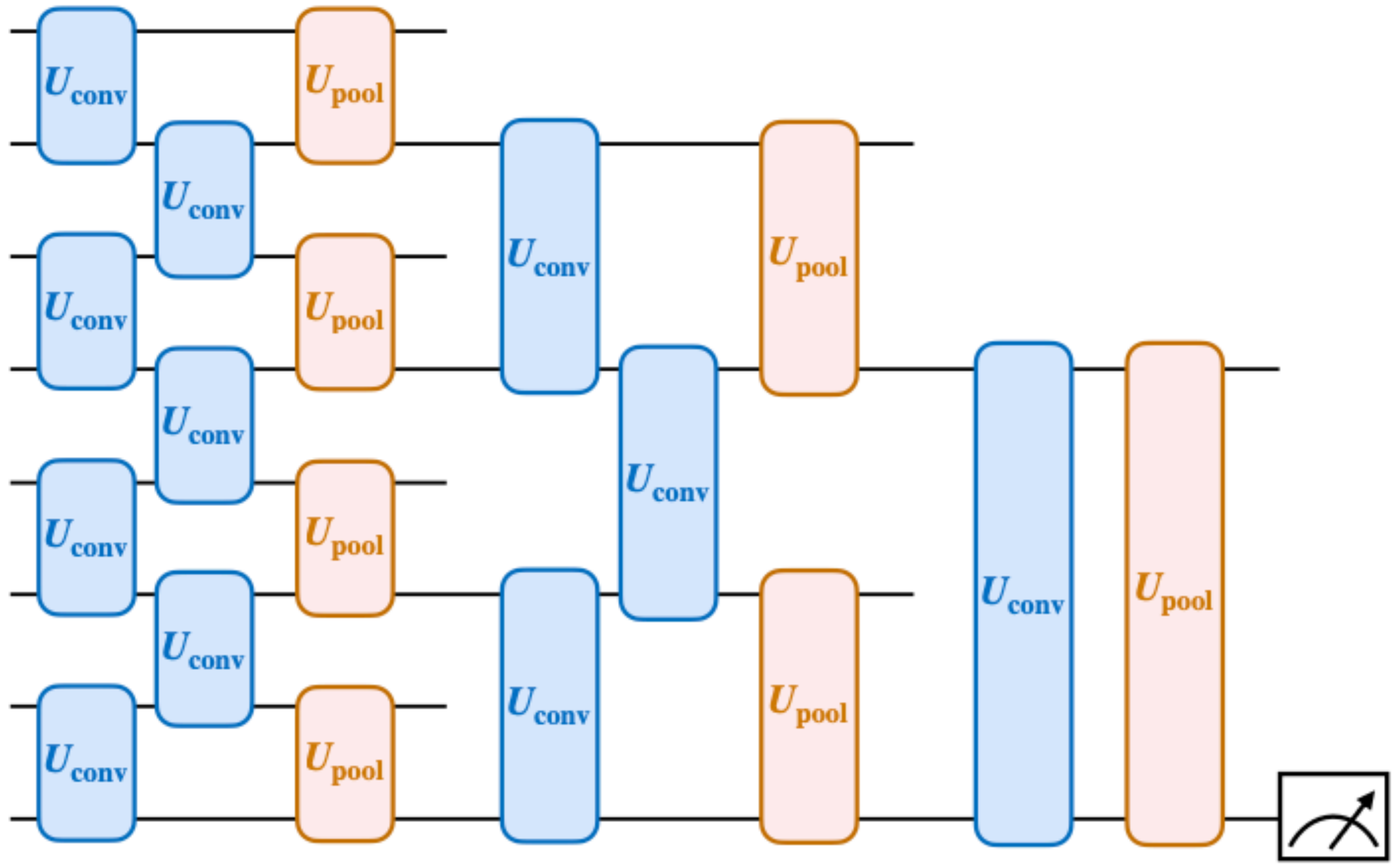
$$L_n - L_{n-1} = \chi_n^\dagger \chi_n - \frac{1 - (-)^n}{2}$$

- we can **eliminate** gauge fields!
  - automatically gauge invariant, no boson fields
  - cannot be used in higher dimension

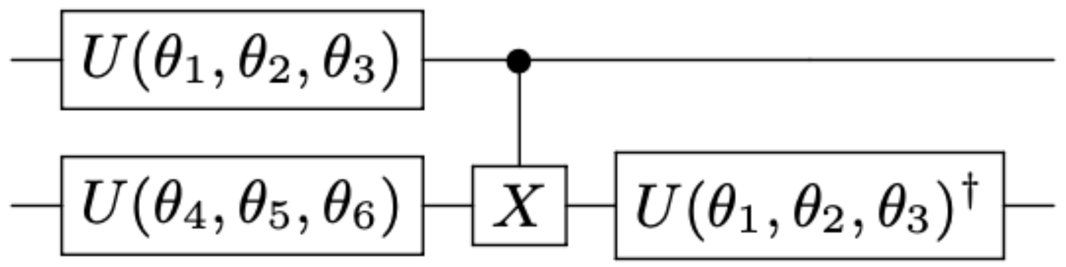
Details on QCNN circuit

- convolution layer
  - two-qubits gates
  - full SU(4) with 15 parameters





- pooling layer
  - two qubits -> single qubit
  - 6 parameters
  - replace classically-controlled gate with CX gate



Hamiltonian variational ansatz for the Schwinger model

$$U_{\rm HVA}(\boldsymbol{\lambda}) = \prod_{l=0}^{N'_L-1} \left[\exp\left(-{\rm i}\lambda_l^{(0)} H_Z\right) \exp\left(-{\rm i}\lambda_l^{(1)} H_{XY}^{(\rm odd)}\right) \exp\left(-{\rm i}\lambda_l^{(2)} H_{XY}^{(\rm even)}\right) \right],$$

$$H_Z = \sum_{n=0}^{N_s-2} \left[\sum_{i=0}^{n} \frac{Z_i + (-1)^i}{2} + \frac{\theta}{2\pi} \right]^2 + \frac{m}{2} \sum_{n=0}^{N_s-1} (-1)^n Z_n\,, \tag{B2}$$

$$H_{XY}^{(\rm odd)} = \sum_{m:{\rm odd}} \left[X_{2m-1}X_{2m} + Y_{2m-1}Y_{2m} \right], \tag{B3}$$

$$H_{XY}^{(\rm even)} = \sum_{m:{\rm even}} \left[X_{2m}X_{2m+1} + Y_{2m}Y_{2m+1} \right]. \tag{B4}$$

QCNN vs HEA

- task: PS simulation
- QCNNm2: simplified conv./pool. layer
- HEA: Hardware-efficient ansatz
- top: case1, bottom: case2