Title: Self-Correcting Quantum Many-Body Control using Reinforcement Learning with Tensor Networks

Speakers: Friederike Metz

Collection: New Frontiers in Machine Learning and Quantum

Date: November 22, 2022 - 2:00 PM

URL: https://pirsa.org/22110087

Abstract: Self-Correcting Quantum Many-Body Control using Reinforcement Learning with Tensor Networks

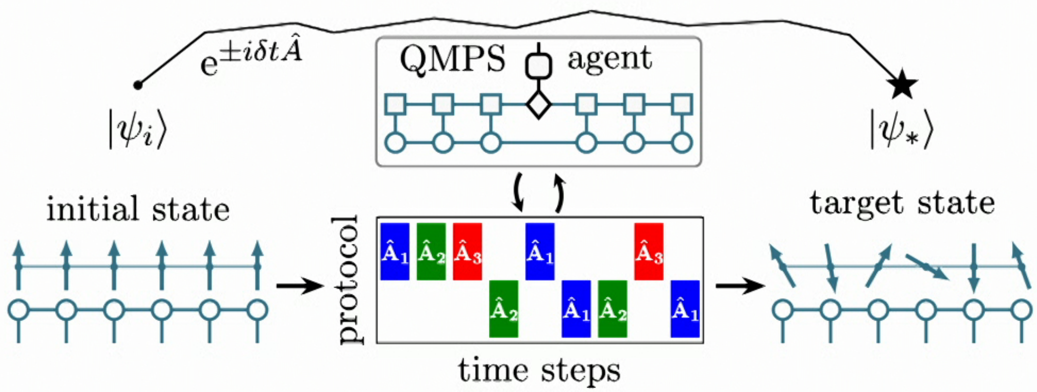
OUTLINE

- Background
  - Matrix product states
  - Reinforcement learning & Q-learning
- Matrix product state ansatz for Q-learning (QMPS)
- Control studies
- NISQ device integration

QUANTUM MANY-BODY CONTROL

Essential for most quantum technologies (computing, simulation, metrology)

e.g. state preparation



Problem: Hilbert space dimension grows exponentially with system size

Our control framework: Deep reinforcement learning (RL) with matrix product states (MPS)

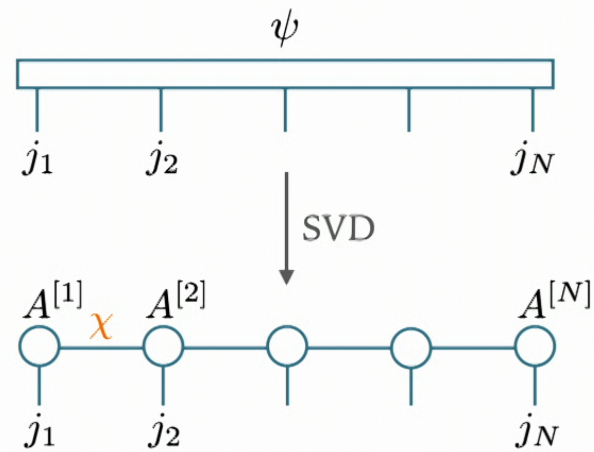
Compressing the quantum many-body state

Trainable machine learning architecture for the RL agent

MPS FOR QUANTUM MANY-BODY SYSTEMS

Problem: Hilbert space dimension grows exponentially with system size: 2^N

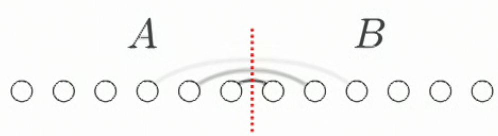
$$|\psi\rangle = \sum_{j_1,j_2,\ldots,j_N} \psi_{j_1,j_2,\ldots,j_N} |j_1\rangle |j_2\rangle \ldots |j_N\rangle \,, \; j_n = 0,1$$



→ Decompose wave function into product of matrices: $2^N \rightarrow 2N\chi^2$

$$\psi_{j_1,j_2,\ldots,j_N} = \sum_{\alpha_1,\ldots,\alpha_{N-1}} A^{[1]j_1}_{\alpha_1} A^{[2]j_2}_{\alpha_1\alpha_2} A^{[3]j_3}_{\alpha_2\alpha_3} \ldots A^{[N]j_N}_{\alpha_{N-1}}$$

Bond dimension χ related to amount of entanglement:



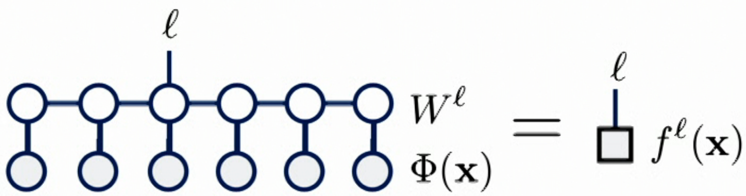
MPS are efficient for 1d area-law entangled states: $S\,(L_A) = -\,\mathrm{Tr}\,\rho_A \log \rho_A \leq \log(\chi) = \mathrm{const.}$

J. Hauschild, F. Pollmann, SciPost Phys. Lect. Notes 5 (2018)

MPS FOR MACHINE LEARNING

E. M. Stoudenmire, D. J. Schwab, Adv. Neural Inf. Process. Syst. 29, 4799 (2016)

$$f^{\ell}(\mathbf{x}) = W^{\ell} \cdot \Phi(\mathbf{x})$$



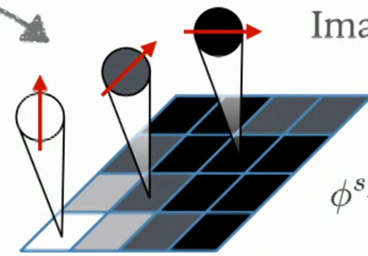
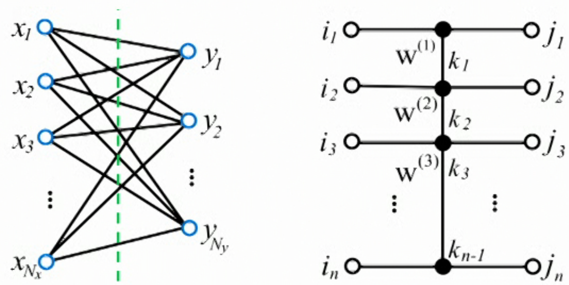


Image pixels are mapped to “spins”

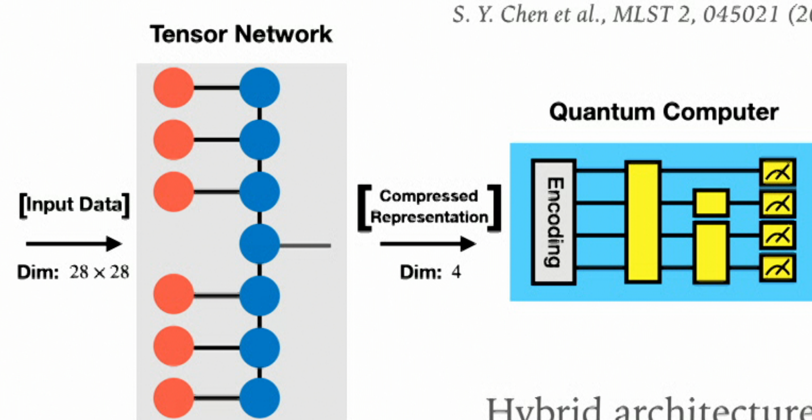
$$\phi^{s_j}(x_j) = \left[\cos\left(\frac{\pi}{2}x_j\right), \sin\left(\frac{\pi}{2}x_j\right)\right]$$

Z. Gao et al., Phys. Rev. Research 2, 023300 (2020)



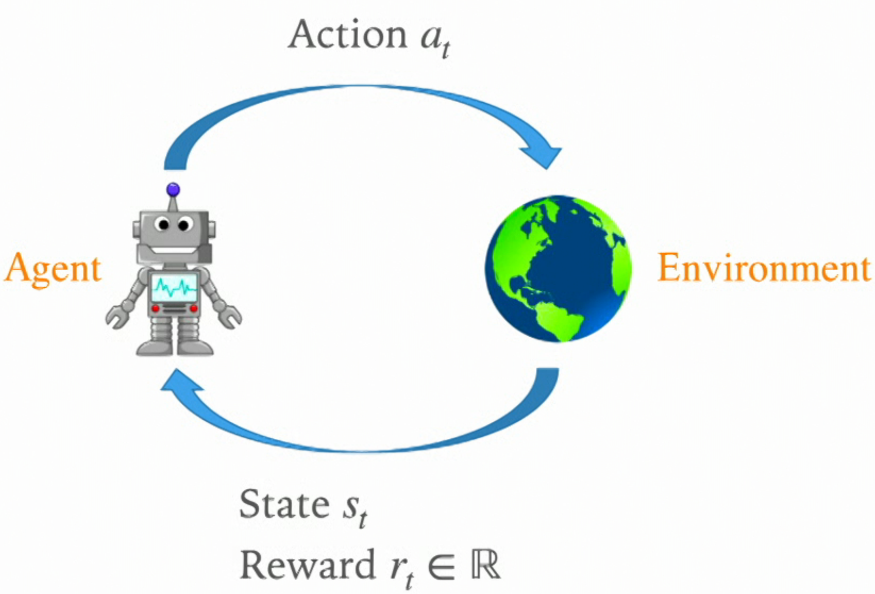
Compress neural network weight matrices

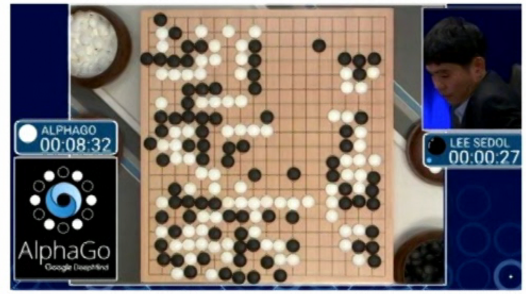
S. Y. Chen et al., MLST 2, 045021 (2021)



Hybrid architectures

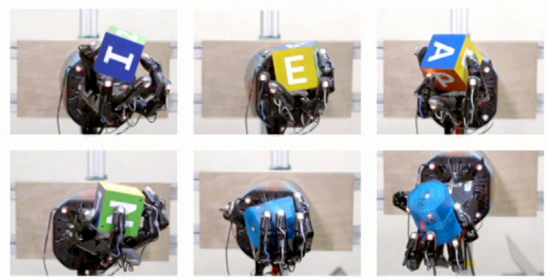
REINFORCEMENT LEARNING





D. Silver et al. Nature **529**, *484 (2016)*

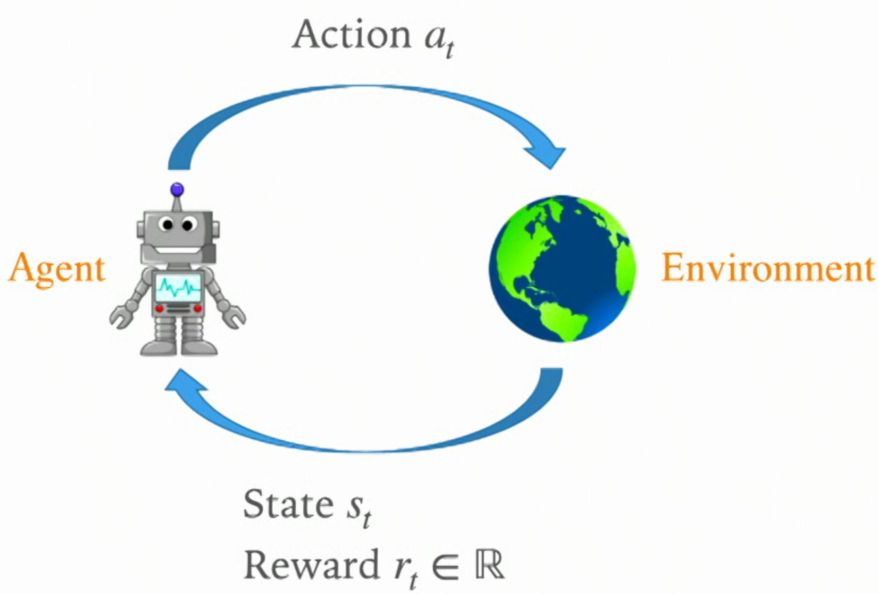
O. Vinyals et al. Nature **575**, *350 (2019)*



OpenAI et al. arXiv:1808.00177 (2019)

Mnih et al., Nature 518, 529–533 (2015)
R. S. Sutton, A. G. Barto, Reinforcement learning: An introduction. MIT press

REINFORCEMENT LEARNING

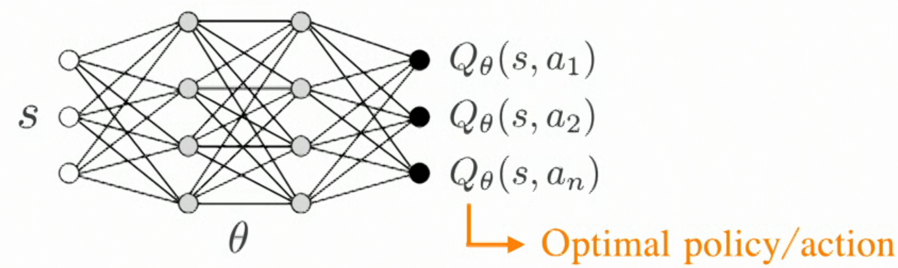


Q-learning: Learn optimal Q values $Q^*(s, a)$

Maximum expected sum of future rewards if you start in state s and take action a

$\rightarrow$ optimal action: $a_t = \arg\max_a Q^*(s_t, a)$

DQN (Deep Q-Network):

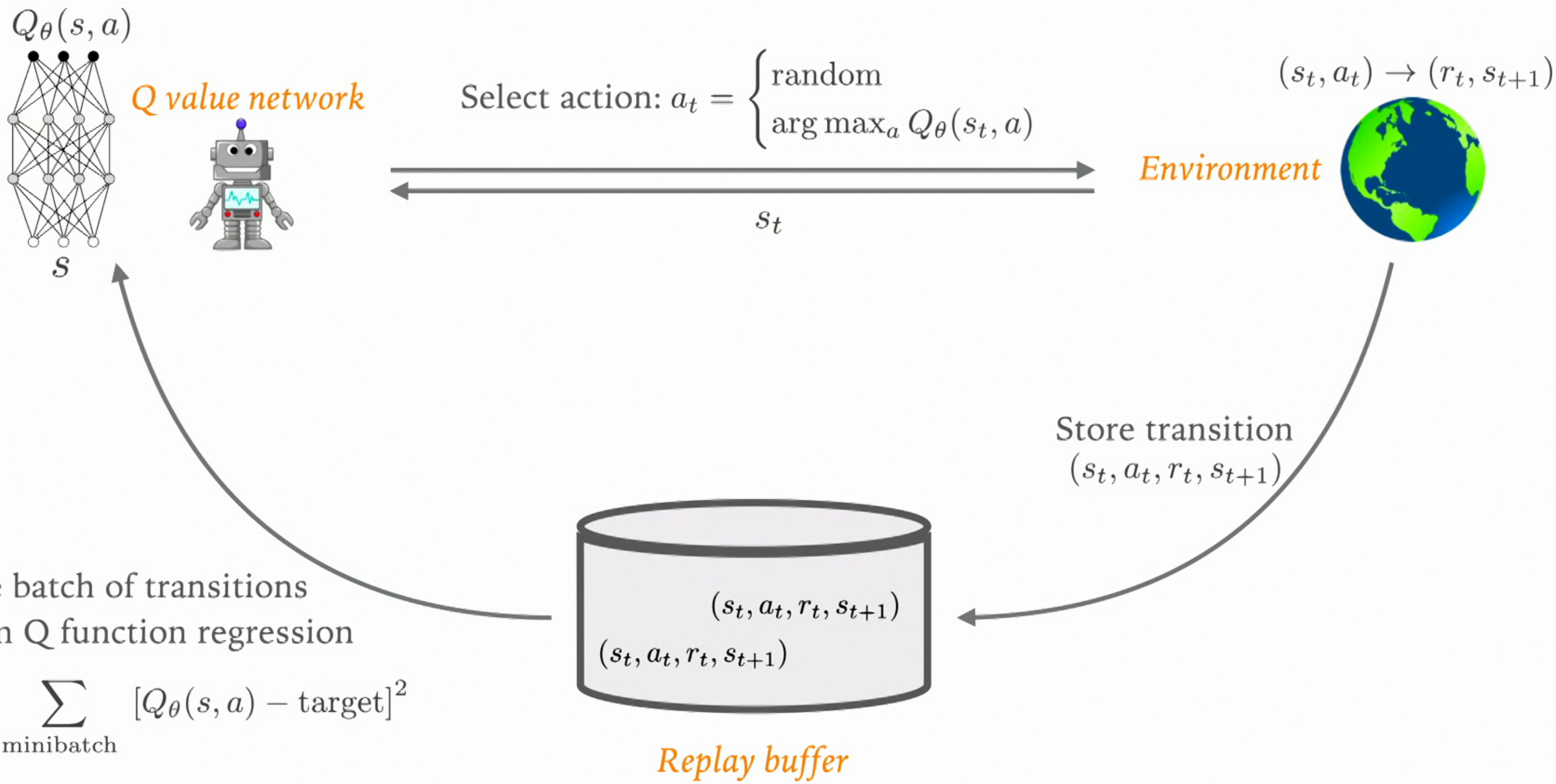


Optimize by minimizing:

$$L(\theta) = \mathbb{E}\left[\left(Q_\theta(s,a) - \left[r(s,a) + \max_{a'} Q_{\bar{\theta}}(s', a')\right]\right)^2\right]$$

Mnih et al., Nature 518, 529–533 (2015)
R. S. Sutton, A. G. Barto, Reinforcement learning: An introduction. MIT press

DEEP Q LEARNING

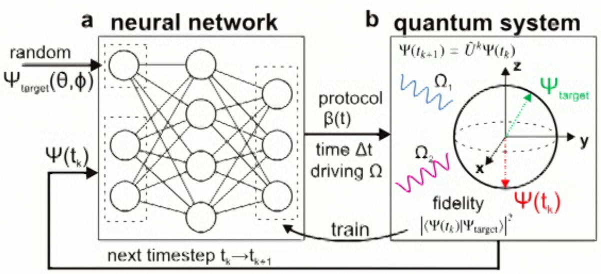


Sample batch of transitions
Perform Q function regression

$$L(\theta) = \sum_{\text{minibatch}} [Q_\theta(s,a) - \text{target}]^2$$

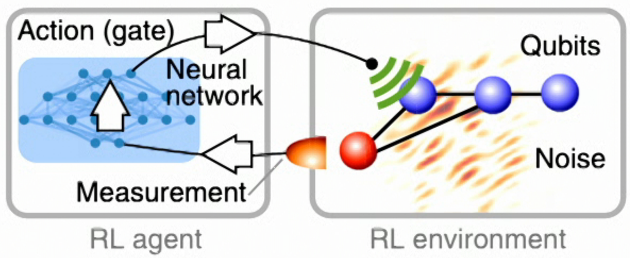
Mnih et al., Nature 518, 529–533 (2015)

RL FOR QUANTUM PHYSICS



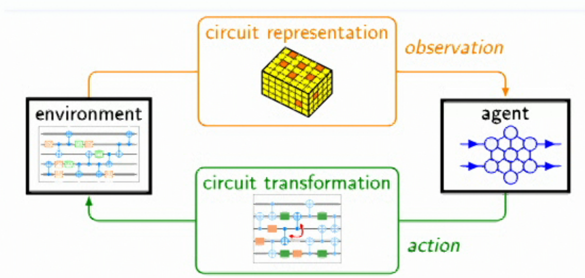
Quantum state preparation

T. Haug et al. MLST **2** 01LT02 (2021)
M. Bukov et al. Phys. Rev. X **8**, 031086 (2018)
A. Bolens, M. Heyl, Phys. Rev. Lett. **127**, 110502 (2021)
...



Quantum error correction

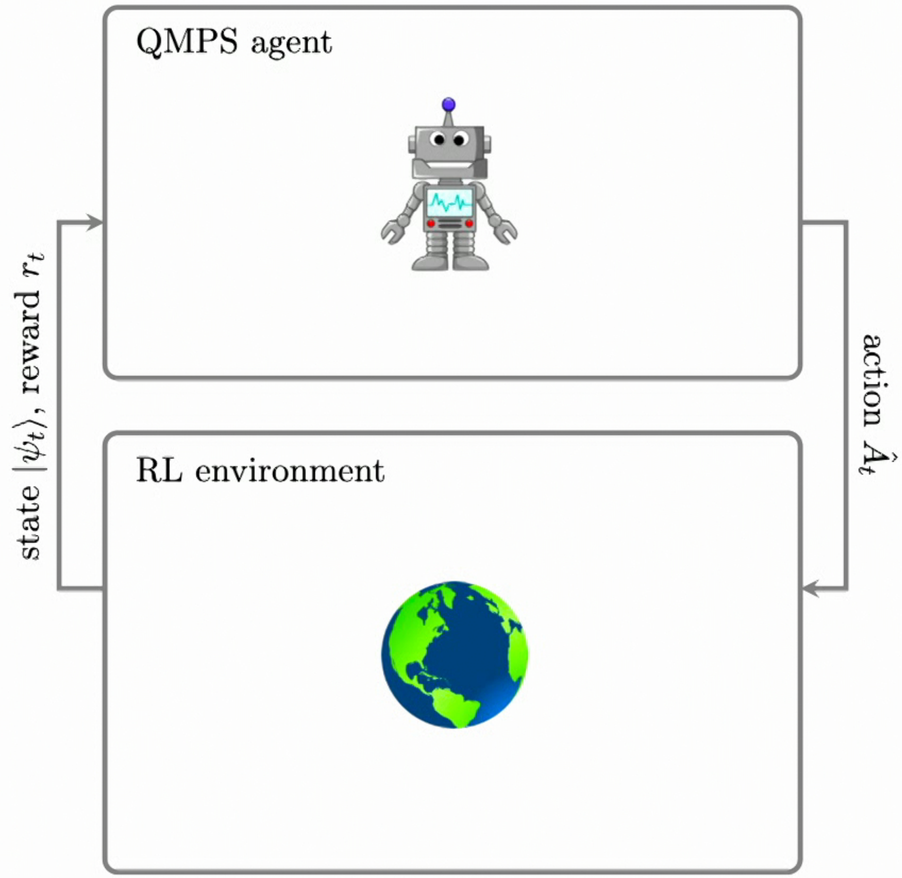
T. Fösel et al. Phys. Rev. X **8**, 031084 (2018)
H. P. Nautrup et al. Quantum **3**, 215 (2019)
R. Sweke et al. MLST **2** 025005 (2021)
...



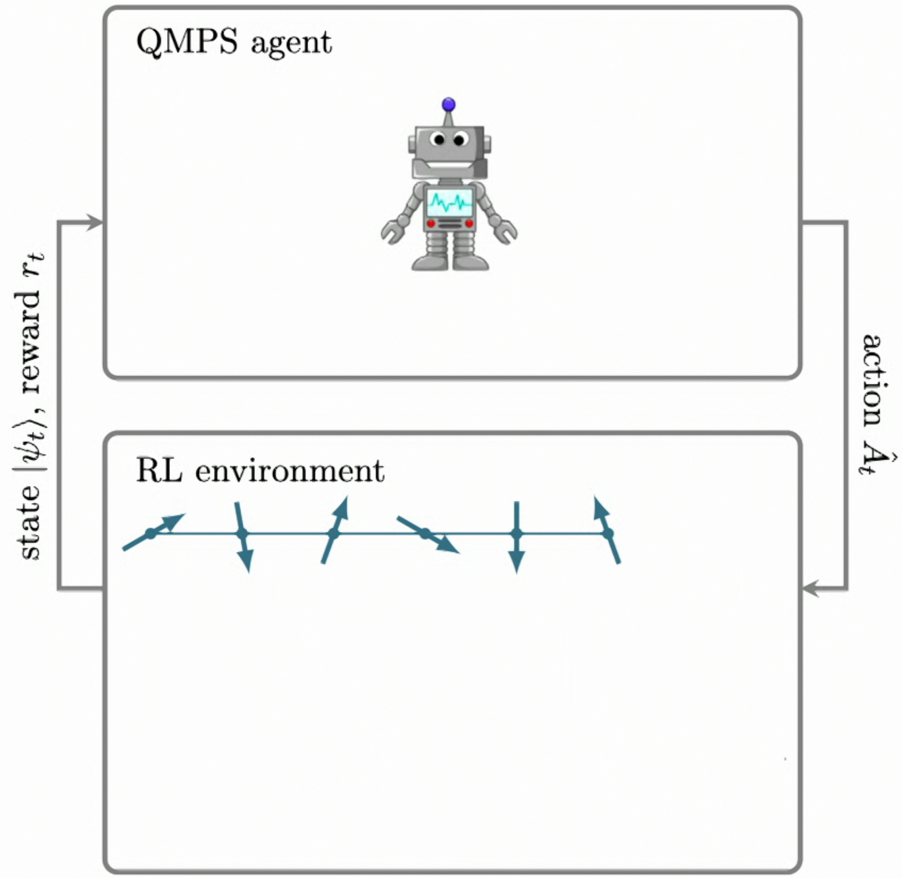
Quantum circuit compiling

T. Fösel et al. arXiv:2103.07585 (2021)
Y.-H. Zhang et al. Phys. Rev. Lett. **125**, 170501 (2020)
L. Moro et al. Commun. Phys. **4**, 178 (2021)
...

MATRIX PRODUCT STATE ANSATZ FOR Q-LEARNING (QMPS)

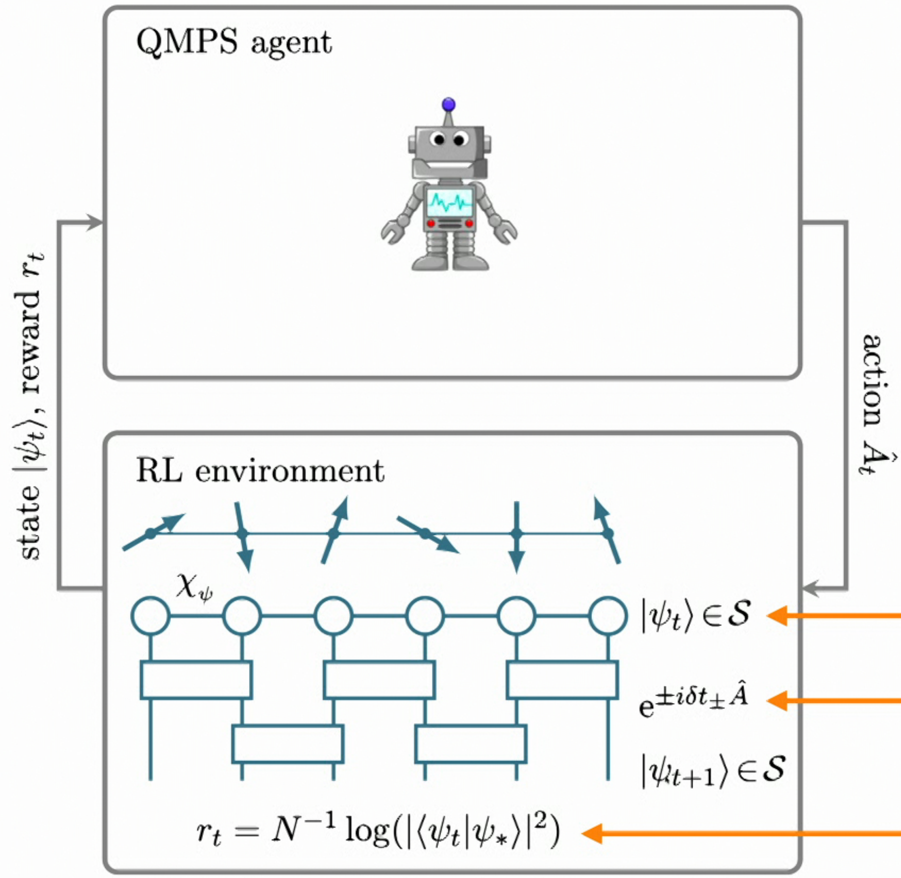


MATRIX PRODUCT STATE ANSATZ FOR Q-LEARNING (QMPS)



RL environment

MATRIX PRODUCT STATE ANSATZ FOR Q-LEARNING (QMPS)



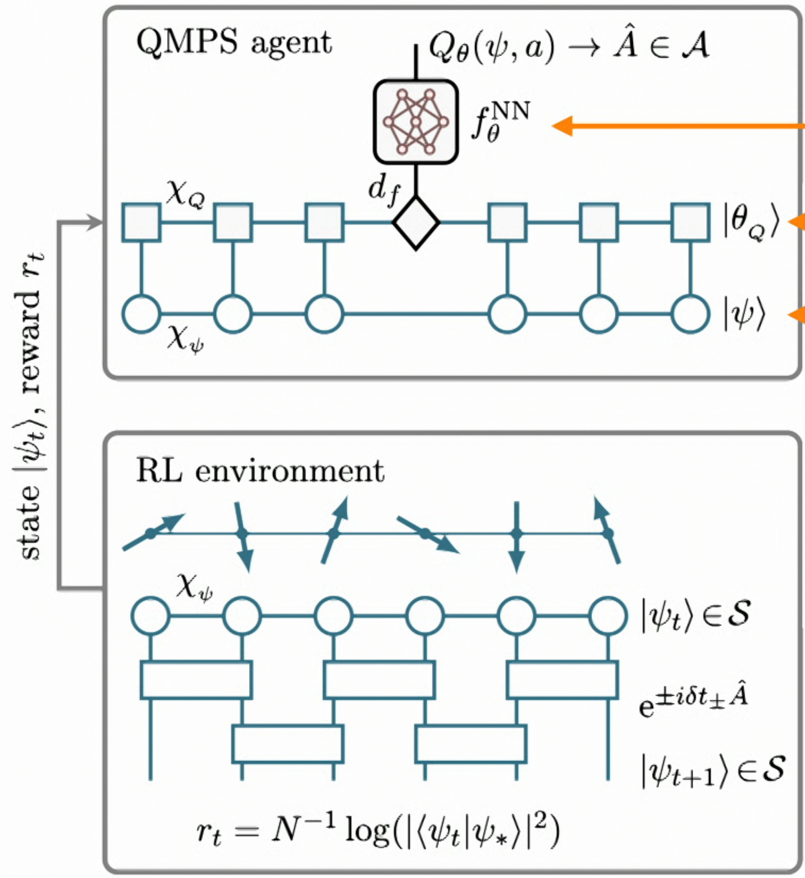
RL environment

State

Action

Reward (Goal: Prepare target state $|\psi_*\rangle$)

MATRIX PRODUCT STATE ANSATZ FOR Q-LEARNING (QMPS)



QMPS agent

Neural network (NN)

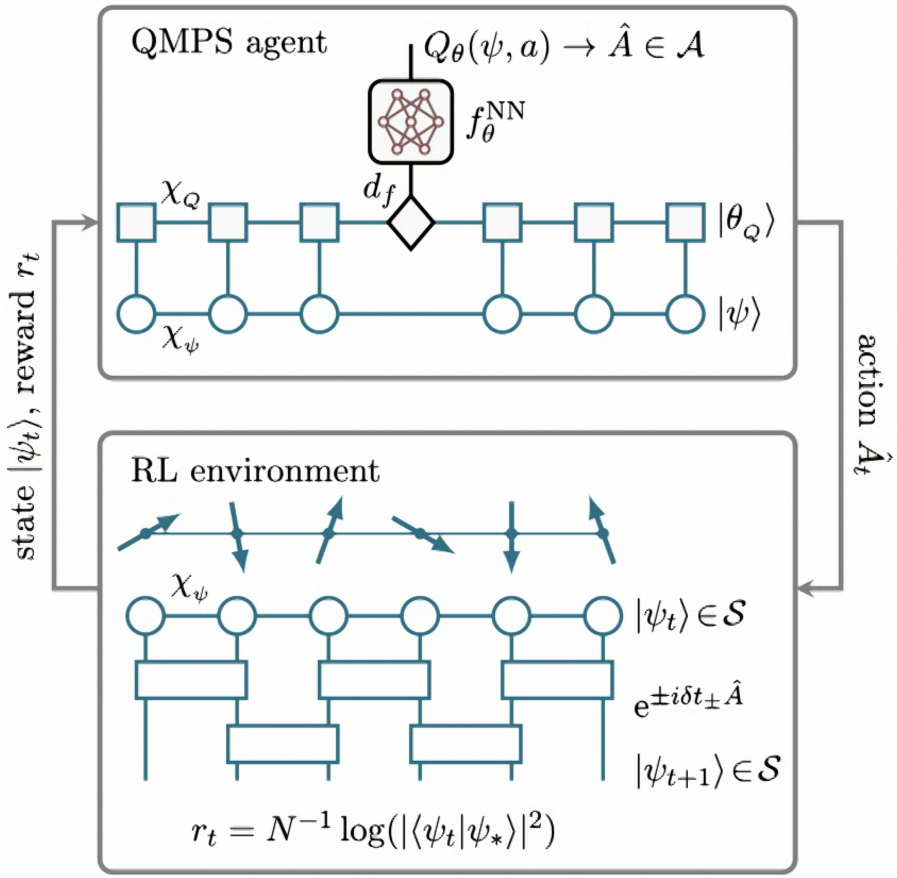
Trainable parameters (MPS)

Quantum state input (MPS)

$$Q_\theta(\psi, a) = f_\theta\left(N^{-1}\log\left(|\langle\theta_Q|\psi\rangle|^2\right)\right)$$

→ QMPS: Resources scale only linearly in system size N

MATRIX PRODUCT STATE ANSATZ FOR Q-LEARNING (QMPS)



QMPS framework

Start each episode from a (random) initial state

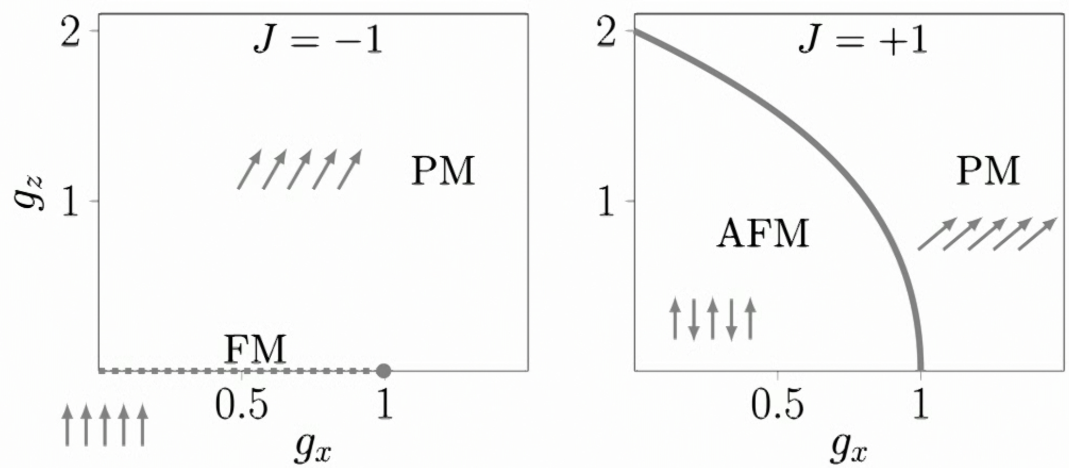
Terminate episode when fidelity threshold is reached, e.g. $F = 0.99$

OUTLINE

- Background
  - Matrix product states
  - Reinforcement learning & Q-learning
- Matrix product state ansatz for Q-learning (QMPS)
- Control studies
- NISQ device integration

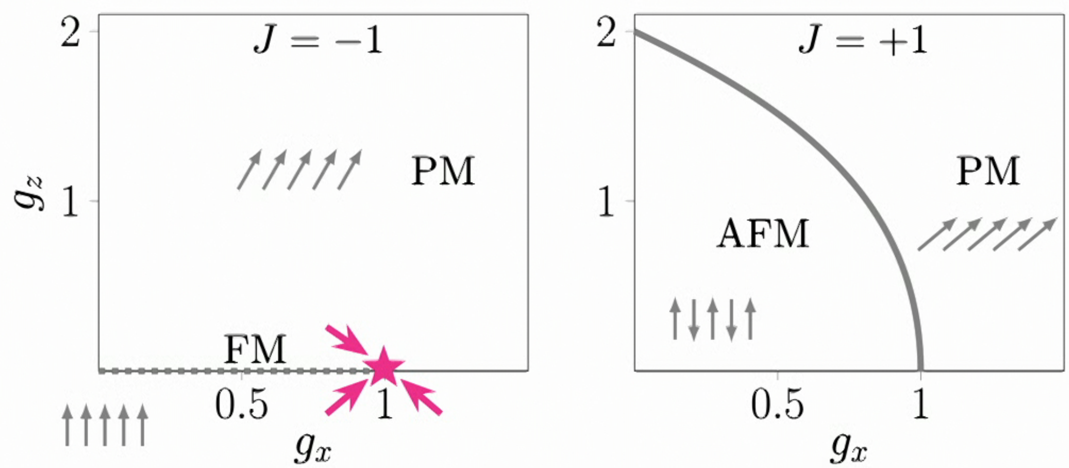
CONTROL STUDIES

Mixed-field Ising model: $\hat{H}_{\text{Ising}} = J \sum_i \hat{Z}_i \hat{Z}_{i+1} - g_x \sum_i \hat{X}_i - g_z \sum_i \hat{Z}_i$



CONTROL STUDIES

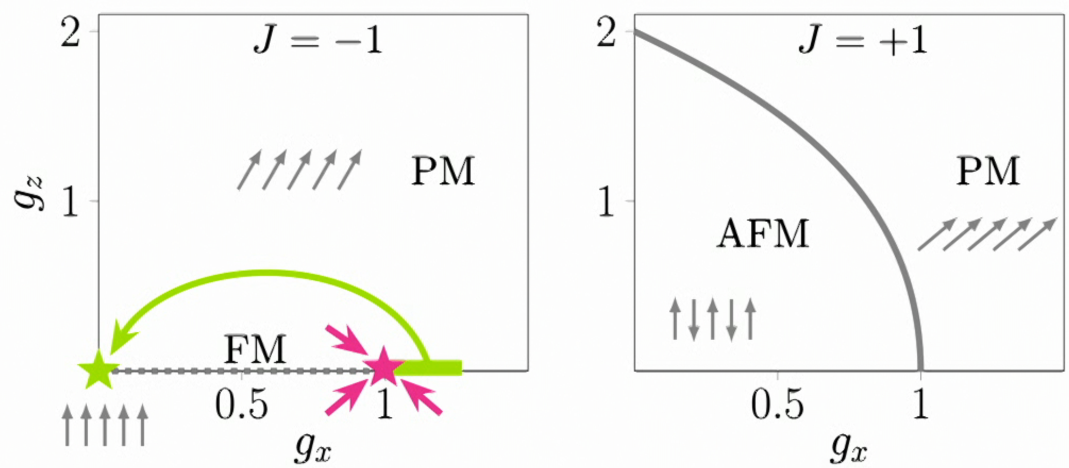
Mixed-field Ising model: $\hat{H}_{\text{Ising}} = J \sum_i \hat{Z}_i \hat{Z}_{i+1} - g_x \sum_i \hat{X}_i - g_z \sum_i \hat{Z}_i$



Control Study A:
critical state from
arbitrary states

CONTROL STUDIES

Mixed-field Ising model: $\hat{H}_{\text{Ising}} = J \sum_i \hat{Z}_i \hat{Z}_{i+1} - g_x \sum_i \hat{X}_i - g_z \sum_i \hat{Z}_i$

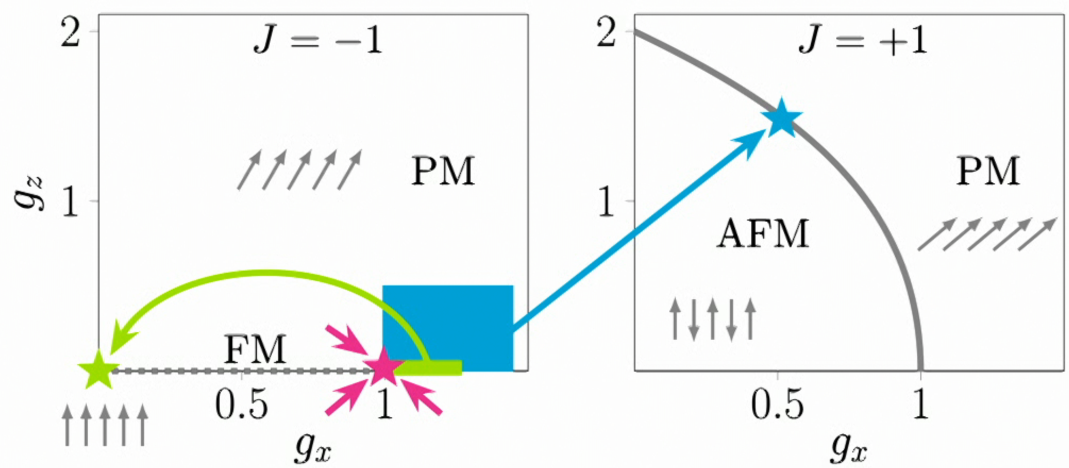


Control Study A:
critical state from
arbitrary states

Control Study B:
z-polarized state from
PM ground states

CONTROL STUDIES

Mixed-field Ising model: $\hat{H}_{\text{Ising}} = J \sum_i \hat{Z}_i \hat{Z}_{i+1} - g_x \sum_i \hat{X}_i - g_z \sum_i \hat{Z}_i$



Control Study A: critical state from arbitrary states

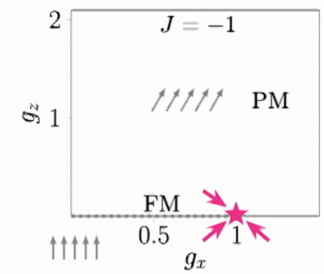
Control Study B: z-polarized state from PM ground states

Control Study C: critical state from PM ground states

UNIVERSAL STATE PREPARATION

Prepare ground state ($J = -1, g_x = 1, g_z = 0$) from randomly drawn initial states

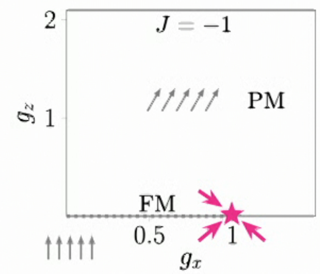
$\rightarrow N = 4$ spins

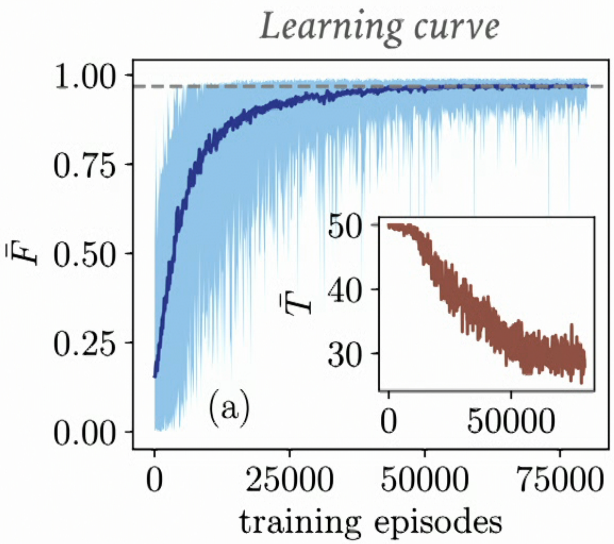


UNIVERSAL STATE PREPARATION

Prepare ground state ($J = -1, g_x = 1, g_z = 0$) from randomly drawn initial states

→ $N = 4$ spins





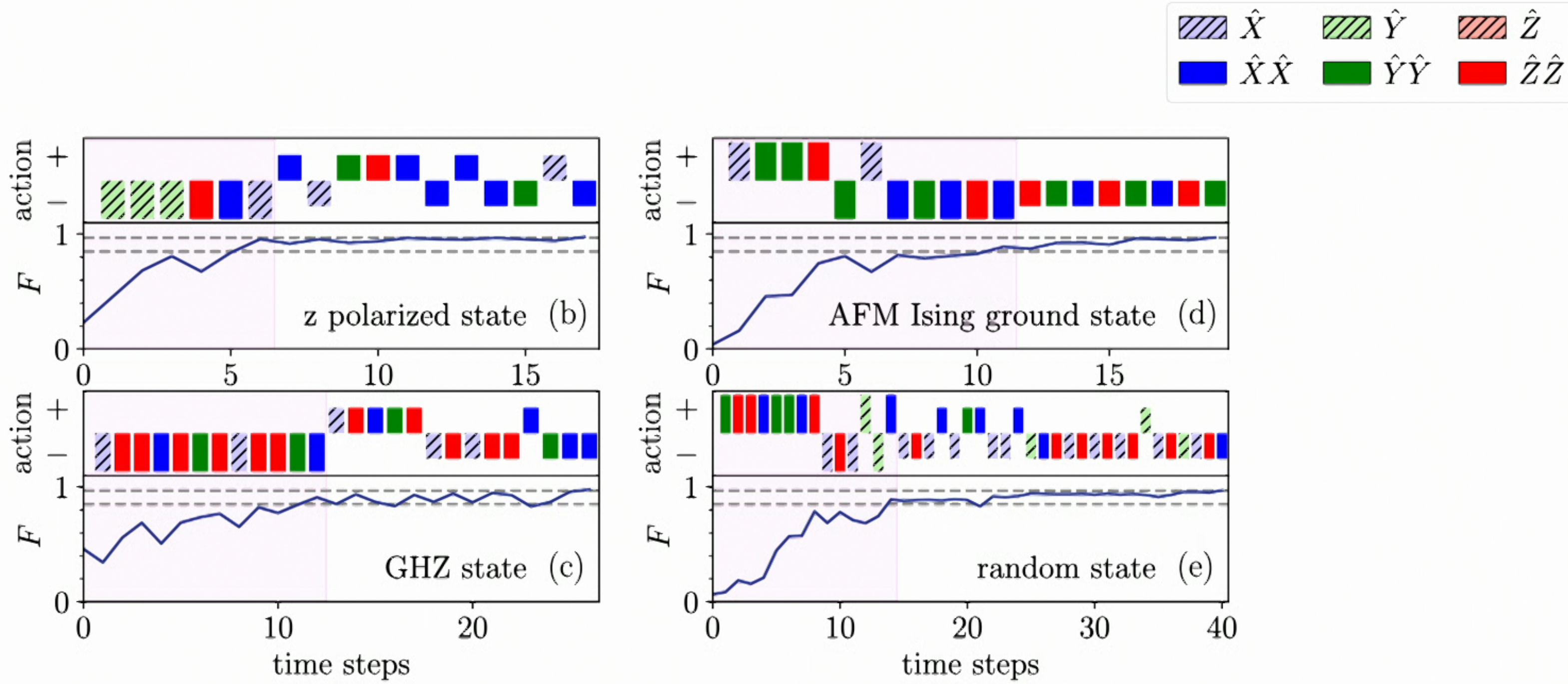
← Fidelity threshold $F^* = 0.97$

← Number of episode steps

UNIVERSAL STATE PREPARATION

Prepare ground state ($J = -1, g_x = 1, g_z = 0$) from randomly drawn initial states

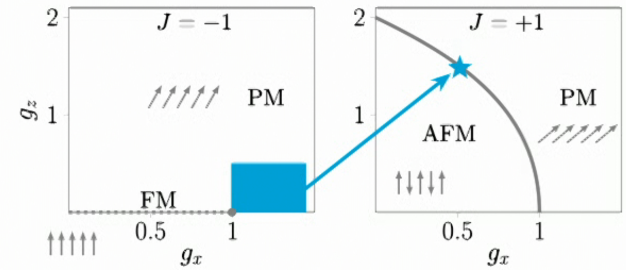
→ $N = 4$ spins

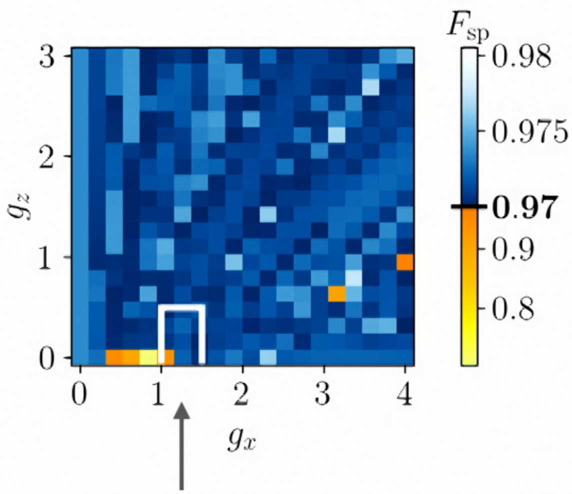


ROBUST CRITICAL STATE PREPARATION

Prepare critical state ($J = +1, g_x = 0.5, g_z = 1.5$) from PM ground states with opposite coupling

$\rightarrow N = 16$, single particle fidelity threshold $F^*_{\mathrm{sp}} = \sqrt[N]{F^*} = 0.97$



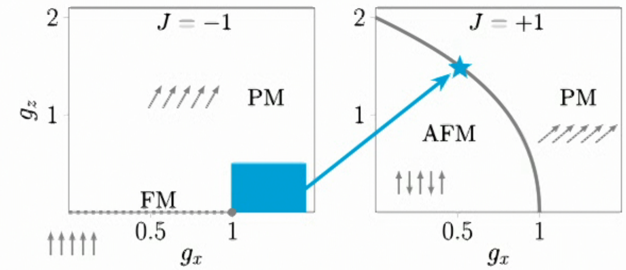


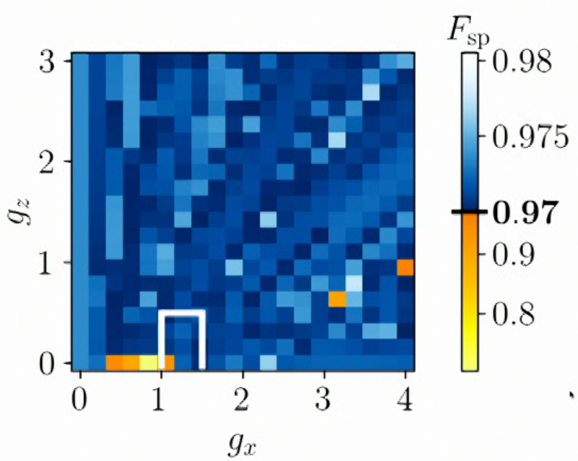
QMPS agent can extrapolate optimal protocols well beyond training region

ROBUST CRITICAL STATE PREPARATION

Prepare critical state ($J = +1, g_x = 0.5, g_z = 1.5$) from PM ground states with opposite coupling

$\rightarrow N = 16$, single particle fidelity threshold $F^*_{\mathrm{sp}} = \sqrt[N]{F^*} = 0.97$



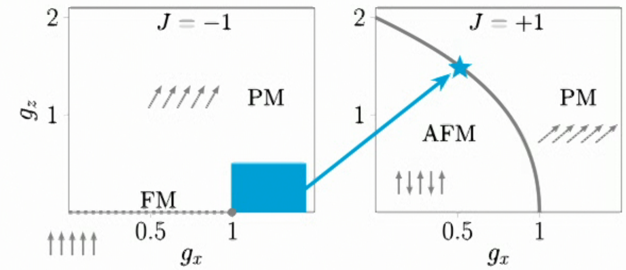


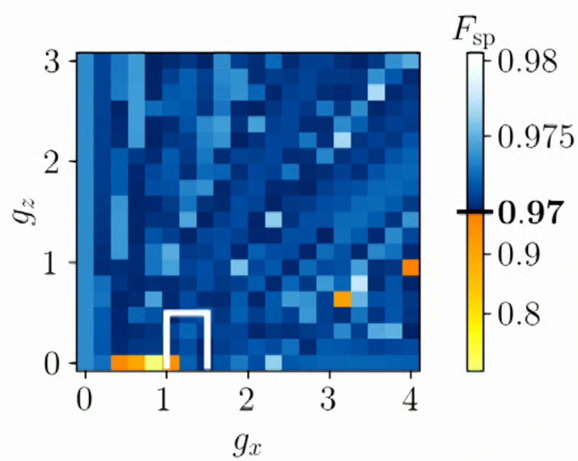
Noise: At each step, white Gaussian noise with std σ is added to step duration δt $\quad (e^{\pm i\delta t_{\pm}\hat{A}})$

ROBUST CRITICAL STATE PREPARATION

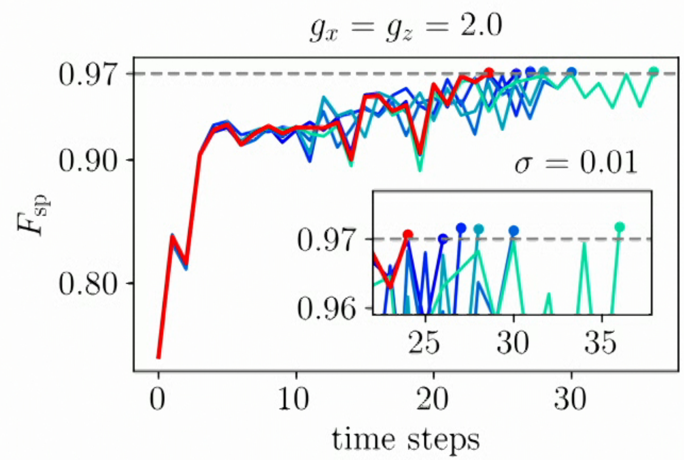
Prepare critical state ($J = +1, g_x = 0.5, g_z = 1.5$)
from PM ground states with opposite coupling

$\rightarrow N = 16$, single particle fidelity threshold $F^*_{\rm sp} = \sqrt[N]{F^*} = 0.97$





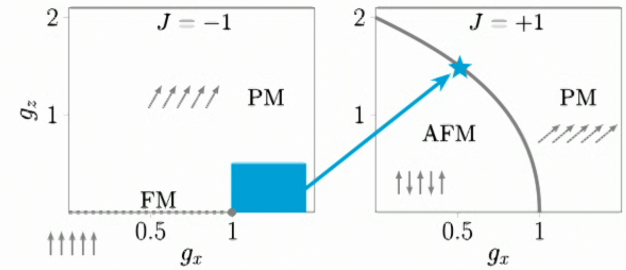
Noise: At each step, white Gaussian noise with std σ is added to step duration δt $\quad (e^{\pm i\delta t_\pm \hat{A}})$

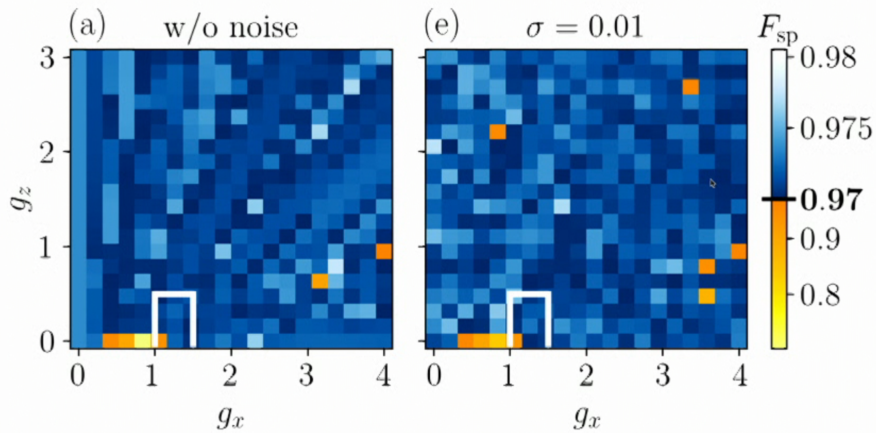


ROBUST CRITICAL STATE PREPARATION

Prepare critical state ($J = +1, g_x = 0.5, g_z = 1.5$)
from PM ground states with opposite coupling

$\rightarrow N = 16$, single particle fidelity threshold $F^*_{\mathrm{sp}} = \sqrt[N]{F^*} = 0.97$





Noise: At each step, white Gaussian noise with std σ is added to step duration δt $\quad (e^{\pm i\delta t_{\pm}\hat{A}})$

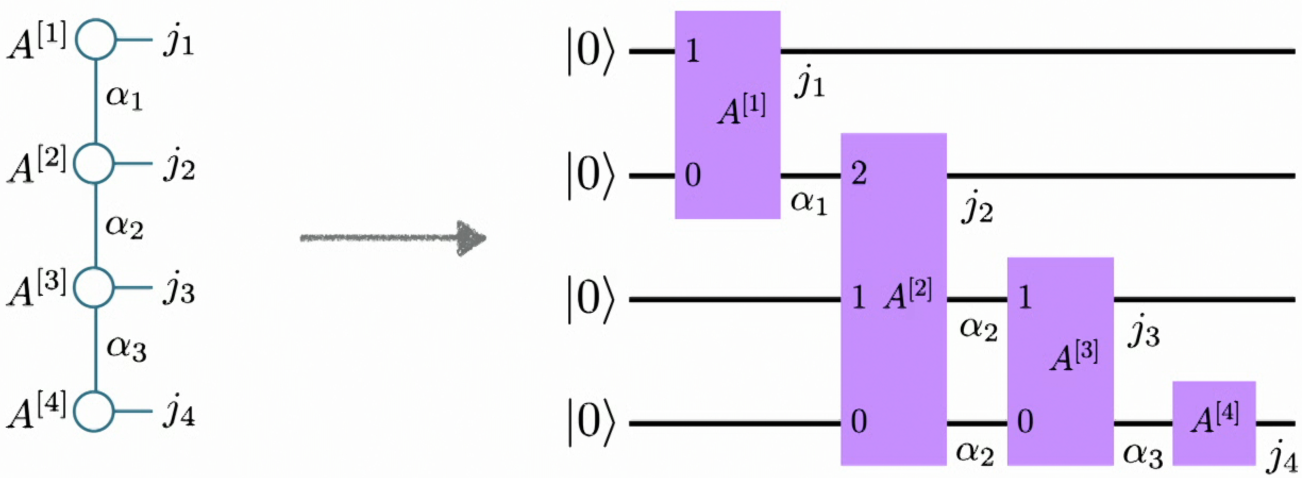
OUTLINE

- Background
  - Matrix product states
  - Reinforcement learning & Q-learning
- Matrix product state ansatz for Q-learning (QMPS)
- Control studies
- NISQ device integration

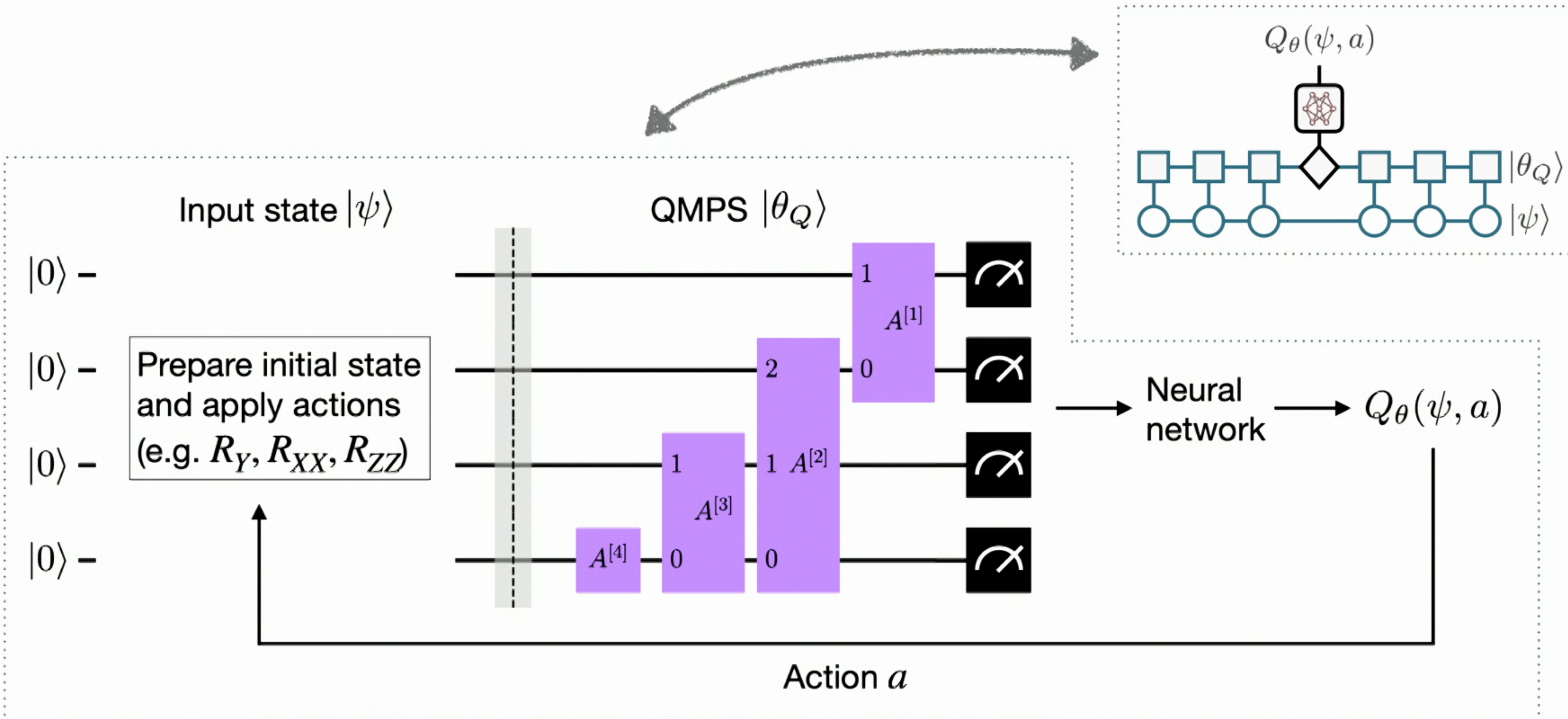
MPS TO CIRCUIT MAPPING

QMPS agent takes quantum state as input at each time step → requires full quantum state tomography

Instead: Map MPS to quantum circuit → integrate with NISQ device simulations



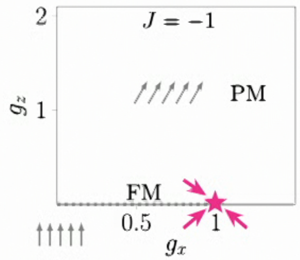
QMPS CIRCUIT FRAMEWORK

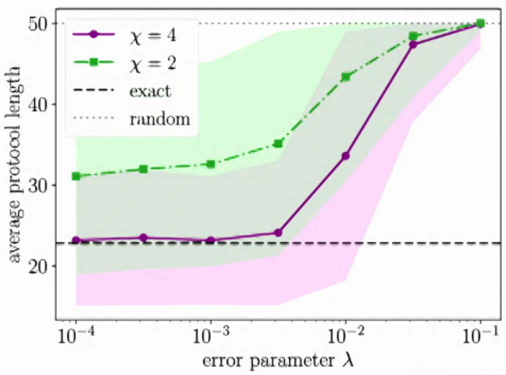


OUTLOOK

QMPS circuit

→ Universal control ($N = 4$)

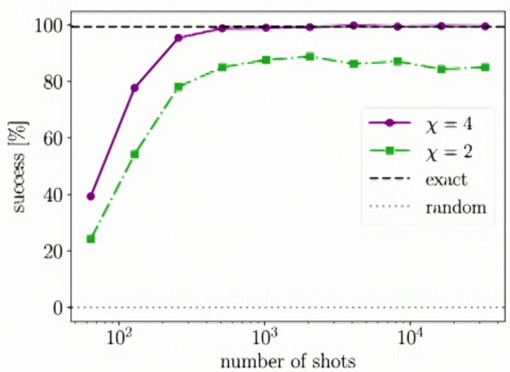




Adding depolarizing noise:

$$E(\rho) = (1 - \lambda)\rho + \lambda \frac{I}{2^N}$$

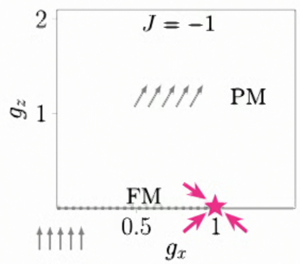
Number of shots for sampling fidelity

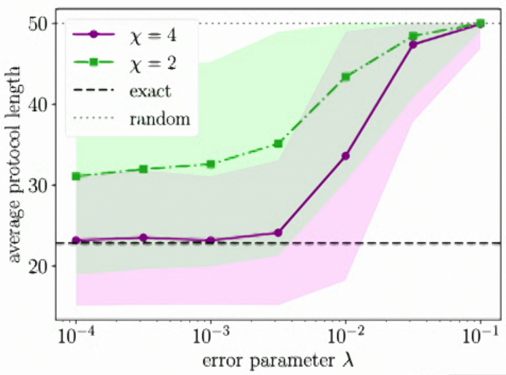


OUTLOOK

QMPS circuit

→ Universal control ($N = 4$)

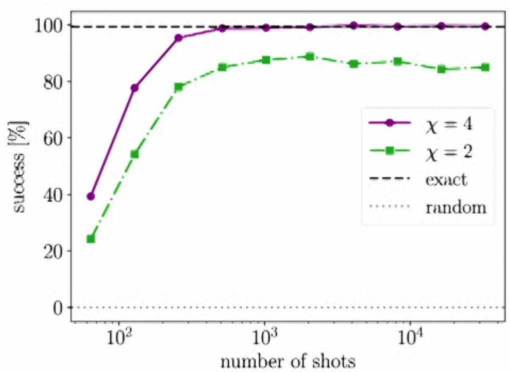




Adding depolarizing noise:

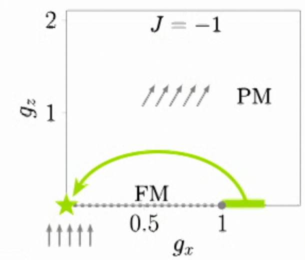
$$E(\rho) = (1 - \lambda)\rho + \lambda \frac{I}{2^N}$$

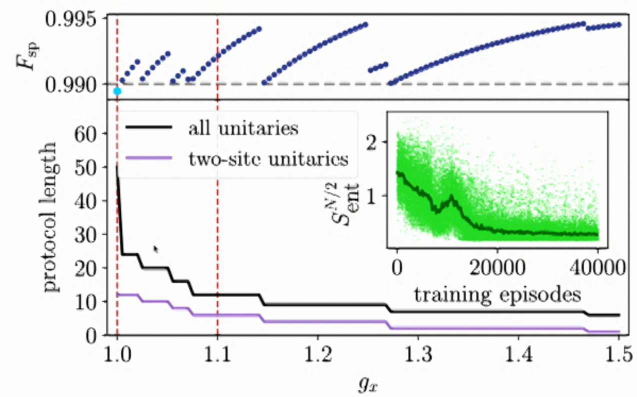
Number of shots for sampling fidelity



Many-body state preparation

→ $N = 32$ spins (beyond ED)





Marin Bukov