Neural Network Ground States of Lattice Gauge Theories

An introduction and overview

Thomas Spriggs, Lattice 2026, Maryland, USA

Quick disclaimer:

- References are a guide for further reading
- Incomplete

Neural Quantum States (NQS)

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

- Simulate lattice gauge theories in **real**, **continuous** time

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

- Simulate lattice gauge theories in **real**, **continuous** time
  - Hamiltonian framework

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

- Simulate lattice gauge theories in **real**, **continuous** time
  - Hamiltonian framework
- Don't sample from the action

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

- Simulate lattice gauge theories in **real**, **continuous** time
  - Hamiltonian framework
- Don’t sample from the action
  - Learn the wavefunction

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

- Simulate lattice gauge theories in **real**, **continuous** time
    - Hamiltonian framework
- Don’t sample from the action
    - Learn the wavefunction
    - No sign problem*

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

- Simulate lattice gauge theories in **real**, **continuous** time
  - Hamiltonian framework
- Don’t sample from the action
  - Learn the wavefunction
  - No sign problem*
- Already done for: $\mathbf{Z}_2, U(1), SU(2)$ lattice gauge theories

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

Neural Quantum States (NQS)

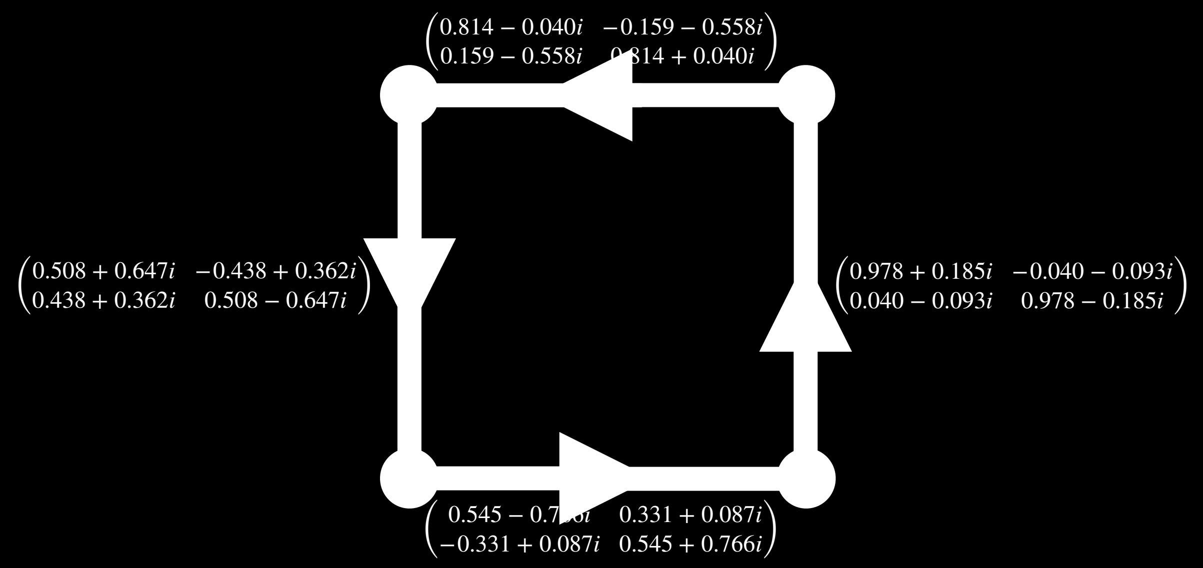
- Simulate lattice gauge theories in **real**, **continuous** time
  - Hamiltonian framework
- Don't sample from the action
  - Learn the wavefunction
  - No sign problem*
- Already done for: $\mathbf{Z}_2, U(1), SU(2)$ lattice gauge theories
  - Commonplace in condensed matter physics and quantum chemistry

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. *Science*, *355*(6325), 602-606.

NQS workflow for LGTs

NN

$|\Psi_\theta\rangle$



NQS workflow for LGTs

$$|\langle U_0|\Psi_\theta\rangle|^2$$

NN

$|\Psi_\theta\rangle$

NQS workflow for LGTs

$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$

NN
$|\Psi_\theta\rangle$

NQS workflow for LGTs

Accept/reject

$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$

NN

$|\Psi_\theta\rangle$

NQS workflow for LGTs

Accept/reject

$$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$$

NN

$|\Psi_\theta\rangle$

$$\{U_i\} \sim |\Psi_\theta\rangle$$

NQS workflow for LGTs

Accept/reject

$$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$$

NN

$|\Psi_\theta\rangle$

$$\{U_i\} \sim |\Psi_\theta\rangle$$

$$E \approx \frac{1}{N} \sum_{U_i \sim |\langle\Psi_\theta|U\rangle|^2} \mathscr{H}_L(U_i)$$

NQS workflow for LGTs

NN

$|\Psi_\theta\rangle$

Accept/reject

$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$

$\{U_i\} \sim |\Psi_\theta\rangle$

$$E \approx \frac{1}{N} \sum_{U_i \sim |\langle \Psi_\theta | U \rangle|^2} \mathscr{H}_L(U_i)$$

NQS workflow for LGTs

Accept/reject

$$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$$

NN

$|\Psi_\theta\rangle$

$$\{U_i\} \sim |\Psi_\theta\rangle$$

$$E_\theta, \partial_\theta E_\theta$$

$$E \approx \frac{1}{N} \sum_{U_i \sim |\langle \Psi_\theta | U\rangle|^2} \mathscr{H}_L(U_i)$$

NQS workflow for LGTs

Accept/reject

$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$

NN

$|\Psi_\theta\rangle$

$\{U_i\} \sim |\Psi_\theta\rangle$

$E_\theta, \partial_\theta E_\theta$

$$E \approx \frac{1}{N} \sum_{U_i \sim |\langle \Psi_\theta | U\rangle|^2} \mathscr{H}_L(U_i)$$

$$E_{GS} = \mathrm{min}_\theta \left(\frac{\langle \Psi_\theta | \mathscr{H} | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} \right)$$

NQS workflow for LGTs

Accept/reject

$$|\langle U_0|\Psi_\theta\rangle|^2 \quad |\langle U_1|\Psi_\theta\rangle|^2$$

NN

$|\Psi_\theta\rangle$

$\{U_i\} \sim |\Psi_\theta\rangle$

Task: find function that returns ground state wavefunction amplitudes

$E_\theta, \partial_\theta E_\theta$

$$E \approx \frac{1}{N} \sum_{U_i \sim |\langle \Psi_\theta | U\rangle|^2} \mathcal{H}_L(U_i)$$

$$E_{GS} = \mathrm{min}_\theta \left(\frac{\langle \Psi_\theta | \mathcal{H} | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle} \right)$$

Lagrangians and Hamiltonians

Lagrangian vs Hamiltonian formalism

From an NQS perspective

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
    - Observables found through the path integral
- Discretise space: regularising theory

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral
- Discretise space: regularising theory
- Discretise time and Wick rotate ($t \to i\tau$)

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral
- Discretise space: regularising theory
- Discretise time and Wick rotate ($t \to i\tau$)

$$\int dx O(x) p(x) \approx \frac{1}{N} \sum_{x \sim p(x)} O(x)$$

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
    - Observables found through the path integral
- Discretise space: regularising theory
- Discretise time and Wick rotate ($t \to i\tau$)

$$\int dx O(x) p(x) \approx \frac{1}{N} \sum_{x \sim p(x)} O(x)$$

$$\updownarrow$$

$$\int D[U] O[U] e^{S_E[U]} \approx \frac{1}{N} \sum_{U \sim e^{S_E[U]}} O[U]$$

Lagrangian vs Hamiltonian formalism

From an NQS perspective

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral
- Discretise space: regularising theory

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
    - Observables found through the path integral
- Discretise space: regularising theory
- ~~Discretise time and Wick rotate ($t \to i\tau$)~~

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral
- Discretise space: regularising theory
- ~~Discretise time and Wick rotate ($t \to i\tau$)~~

$$\int dx O(x) p(x) \approx \frac{1}{N} \sum_{x \sim p(x)} O(x)$$

Lagrangian vs Hamiltonian formalism

From an NQS perspective

- Continuum Dirac equation
  - Observables found through the path integral
- Discretise space: regularising theory
- ~~Discretise time and Wick rotate ($t \rightarrow i\tau$)~~

$$\int dx O(x) p(x) \approx \frac{1}{N} \sum_{x \sim p(x)} O(x)$$

Lagrangian vs Hamiltonian formalism

Trade off

Lose access to real time dynamics

Get a known expression for the probability distribution of gauge fields

Complex actions ruin this: e^{S_E} is no longer a probability

Lagrangian vs Hamiltonian formalism

Trade off

Keep access to real time dynamics

Have to learn the probability distribution (using a neural network)

Probability distribution is always real

Kogut, J., & Susskind, L. (1975). Hamiltonian formulation of Wilson's lattice gauge theories. *Physical Review D*, *11*(2), 395.

Neural (network) wavefunctions

A quick aside into spin systems (easier to picture)

Some definitions: wavefunction

Some definitions: wavefunction

- Let $|\Psi\rangle = \sum_i a_i |x_i\rangle$

Some definitions: wavefunction

- Let $|\Psi\rangle = \sum_i a_i |x_i\rangle$
    - 2 qubits, spin basis: $|\Psi\rangle = a_0|00\rangle + a_1|01\rangle + a_2|10\rangle + a_3|11\rangle$

Some definitions: wavefunction

- Let $|\Psi\rangle = \sum_i a_i |x_i\rangle$
  - 2 qubits, spin basis: $|\Psi\rangle = a_0|00\rangle + a_1|01\rangle + a_2|10\rangle + a_3|11\rangle$
  - “Finding ground state” == finding four* numbers

Some definitions: variational wavefunction

- Let $|\Psi\rangle = \sum_i a_i |x_i\rangle$
  - 2 qubits, spin basis: $|\Psi\rangle = a_0|00\rangle + a_1|01\rangle + a_2|10\rangle + a_3|11\rangle$
  - "Finding ground state" == finding a function that gives four* numbers

$$|\Psi_\theta\rangle = \sum_i f_\theta(x_i)|x_i\rangle$$

Some definitions: variational wavefunction

$$|\Psi_\theta\rangle = \sum_i f_\theta(x_i)\,|x_i\rangle$$

Some definitions: variational wavefunction

$$|\Psi_\theta\rangle = \sum_i f_\theta(x_i)\,|x_i\rangle$$

$$\langle x_0|\Psi_\theta\rangle = \sum_i \langle x_0|f_\theta(x_i)|x_i\rangle = f_\theta(x_0)$$

Some definitions: variational wavefunction

$$|\Psi_\theta\rangle = \sum_i f_\theta(x_i)\,|x_i\rangle$$

$$\langle x_0|\Psi_\theta\rangle = \sum_i \langle x_0|f_\theta(x_i)|x_i\rangle = f_\theta(x_0)$$

"Configuration in, amplitude out"

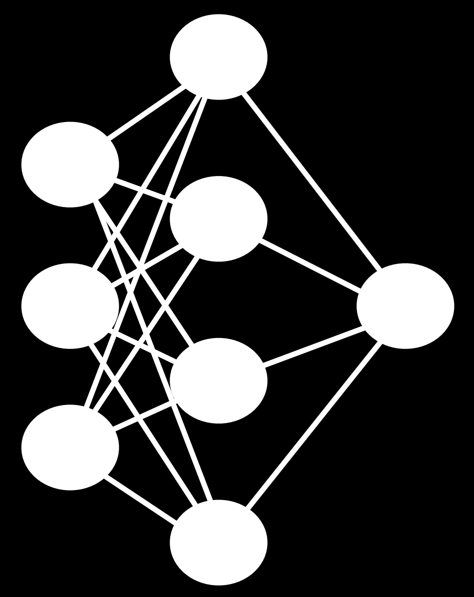
Some definitions: variational wavefunction

$$|\Psi_\theta\rangle = \sum_i f_\theta(x_i)\,|x_i\rangle$$

$$\langle x_0|\Psi_\theta\rangle = \sum_i \langle x_0|f_\theta(x_i)|x_i\rangle = f_\theta(x_0)$$

“Configuration in, amplitude out”

$\langle x_0|$ $\longrightarrow$



$\longrightarrow$ $f_\theta(x_0)$

Wavefunctions in gauge theories

Finding the probability measure

Take an un-normalised variational state, $|\Psi_\theta\rangle$, a Hilbert space, $\{|\vec{x}\rangle\}$, and a Hamiltonian, $\mathscr{H}$

$$E = \langle \mathscr{H} \rangle_\theta = \frac{\langle \Psi_\theta | \mathscr{H} | \Psi_\theta \rangle}{\langle \Psi_\theta | \Psi_\theta \rangle}$$

Introduce a complete set of states $\sum_{\vec{x}} |\vec{x}\rangle\langle\vec{x}|$

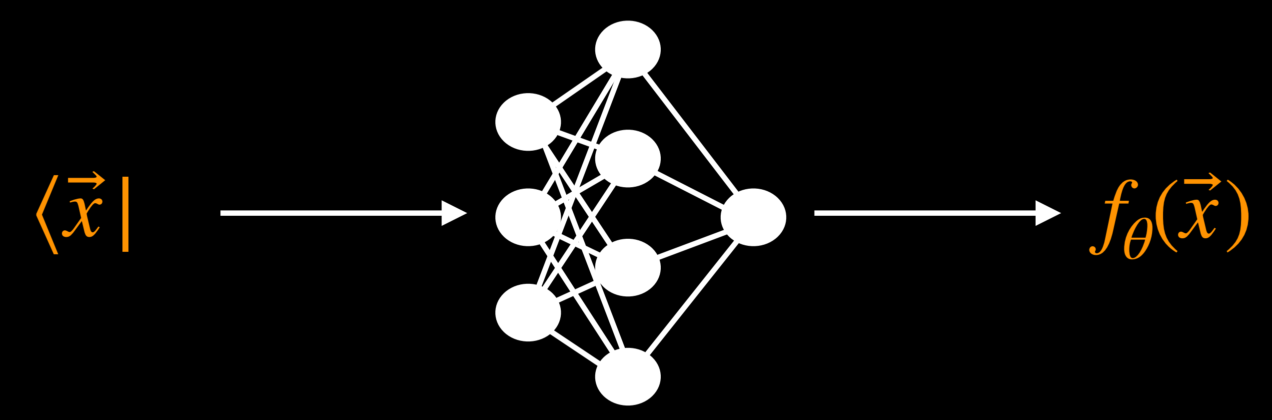
$$E = \frac{\sum_{\vec{x}} \langle \Psi_\theta | \vec{x} \rangle \langle \vec{x} | \mathscr{H} | \Psi_\theta \rangle}{\sum_{\vec{x}'} \langle \Psi_\theta | \vec{x}' \rangle \langle \vec{x}' | \Psi_\theta \rangle}$$

Multiply top and bottom by a factor of $\langle \vec{x} | \Psi_\theta \rangle$

$$E = \frac{\sum_{\vec{x}} |\langle \Psi_\theta | \vec{x} \rangle|^2 \langle \vec{x} | \mathscr{H} | \Psi_\theta \rangle}{\sum_{\vec{x}'} |\langle \Psi_\theta | \vec{x}' \rangle|^2 \langle \vec{x} | \Psi_\theta \rangle} = \frac{\sum_{\vec{x}} p(\vec{x}) \mathscr{H}_L(\vec{x})}{\sum_{\vec{x}'} p(\vec{x}')} \approx \frac{1}{N} \sum_{\vec{x} \sim |\langle \Psi_\theta | \vec{x} \rangle|^2} \mathscr{H}_L(x)$$

Medvidović, M., & Moreno, J. R. (2024). Neural-network quantum states for many-body physics. *The European Physical Journal Plus*, *139*(7), 1-26.

Finding the probability measure



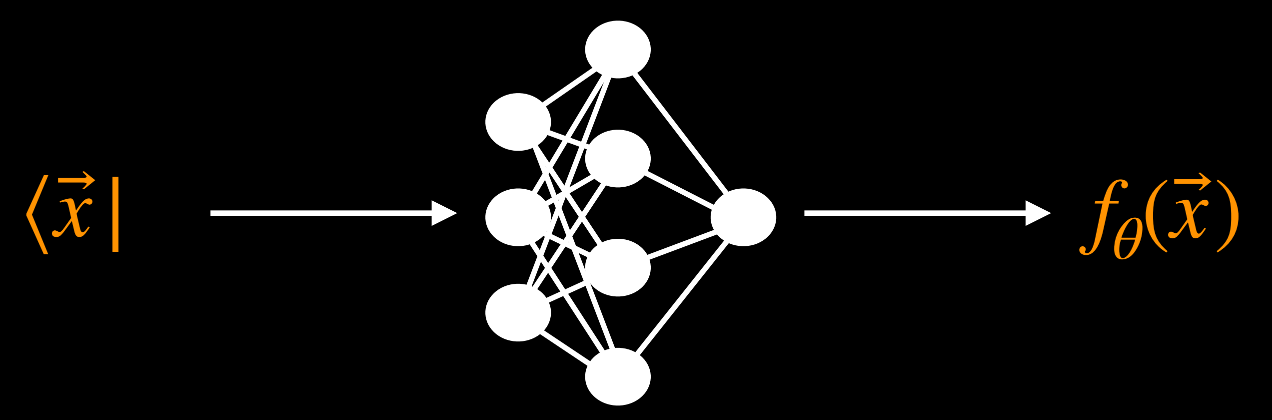
- We can use importance sampling
  - Draw samples according to

    $$| \langle \Psi_\theta | \vec{x} \rangle |^2$$

  - Measure **local** value of the observable on **each sample**

$$\mathscr{H}_L = \frac{\langle \vec{x} | \mathscr{H} | \Psi_\theta \rangle}{\langle \vec{x} | \Psi_\theta \rangle}$$

Finding the probability measure



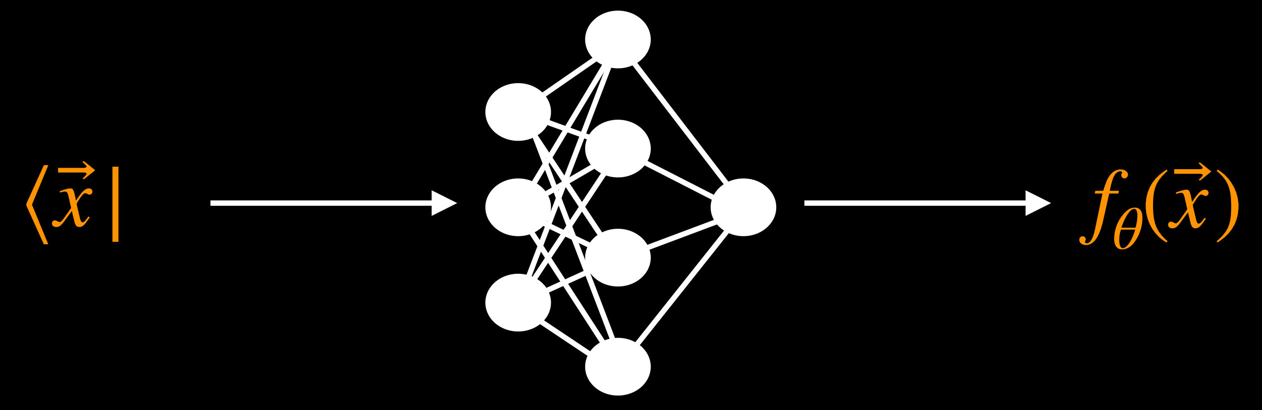
- We can use importance sampling
  - Draw samples according to

    $|\langle \Psi_\theta | \vec{x} \rangle|^2$ **Always real; no sign problem**

  - Measure **local** value of the observable on **each sample**

$$\mathscr{H}_L = \frac{\langle \vec{x} | \mathscr{H} | \Psi_\theta \rangle}{\langle \vec{x} | \Psi_\theta \rangle}$$

Finding the probability measure



- We can use importance sampling
  - Draw samples according to

$$|\langle \Psi_\theta | \vec{x} \rangle|^2$$

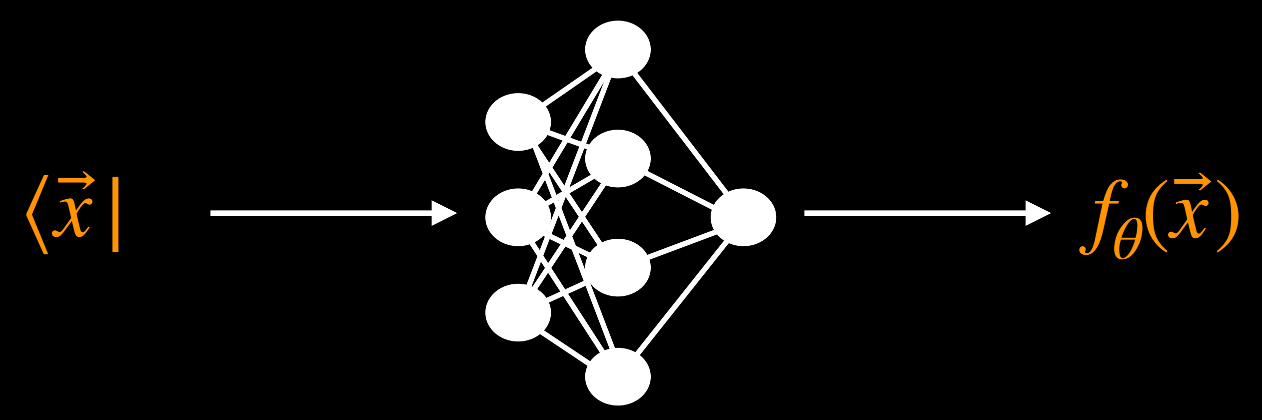
Always real; no sign problem

We do not know this distribution *a priori*

  - Measure **local** value of the observable on **each sample**

$$\mathcal{H}_L = \frac{\langle \vec{x} | \mathcal{H} | \Psi_\theta \rangle}{\langle \vec{x} | \Psi_\theta \rangle}$$

Finding the probability measure



- We can use importance sampling
  - Draw samples according to

    $$| \langle \Psi_\theta | \vec{x} \rangle |^2$$

    Always real; no sign problem

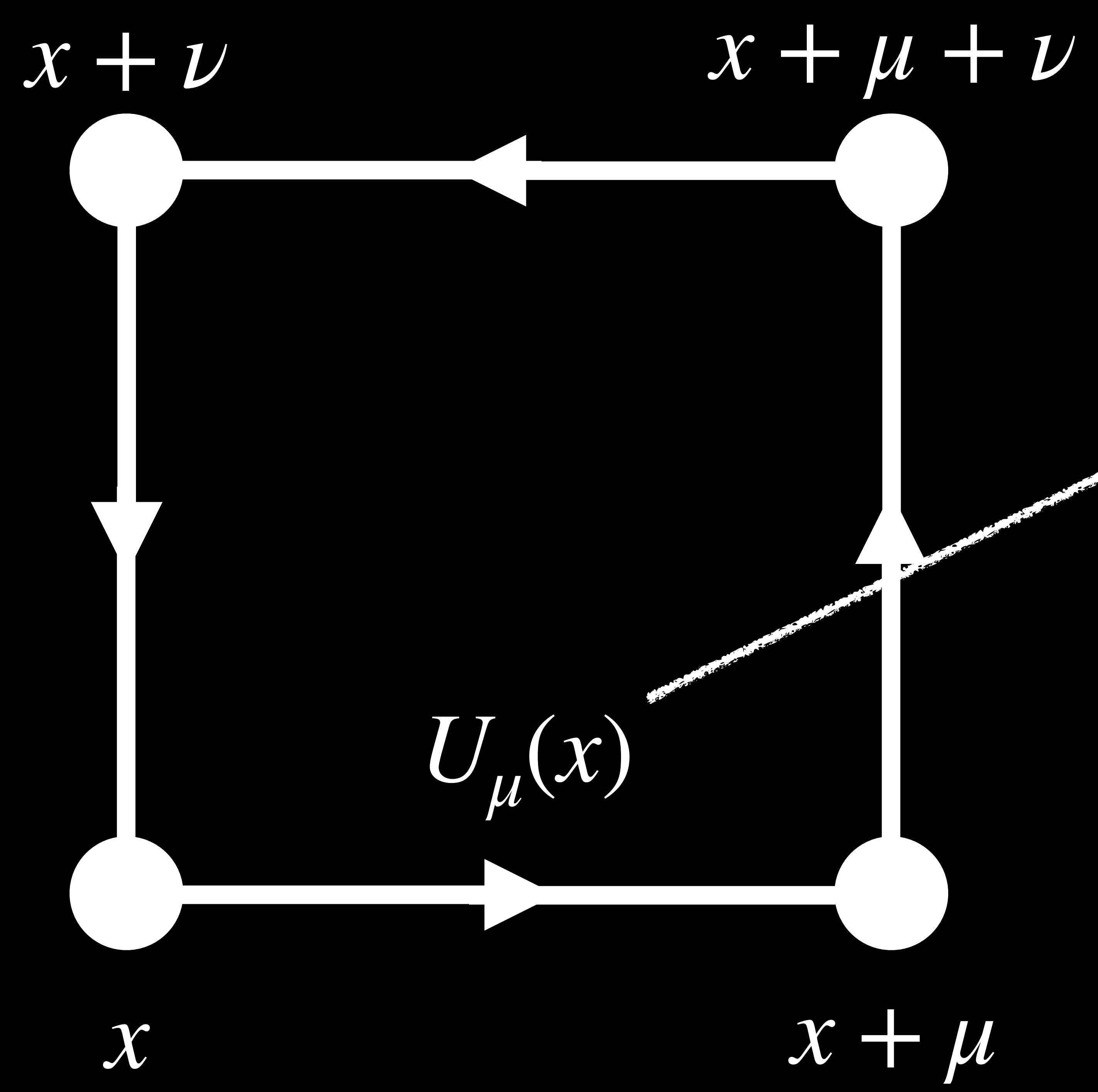
    We do not know this distribution *a priori*
  - Measure **local** value of the observable on **each sample**

$$\mathcal{H}_L = \frac{\langle \vec{x} | \mathcal{H} | \Psi_\theta \rangle}{\langle \vec{x} | \Psi_\theta \rangle}$$

$$\text{Energy} \approx \frac{1}{N} \sum_{\vec{x} \sim |\langle \Psi_\theta | \vec{x} \rangle|^2} \mathcal{H}_L(x)$$

Gauge-theory-aware neural networks

SU(2) example



$U_\mu(x) \in SU(2)$

e.g.

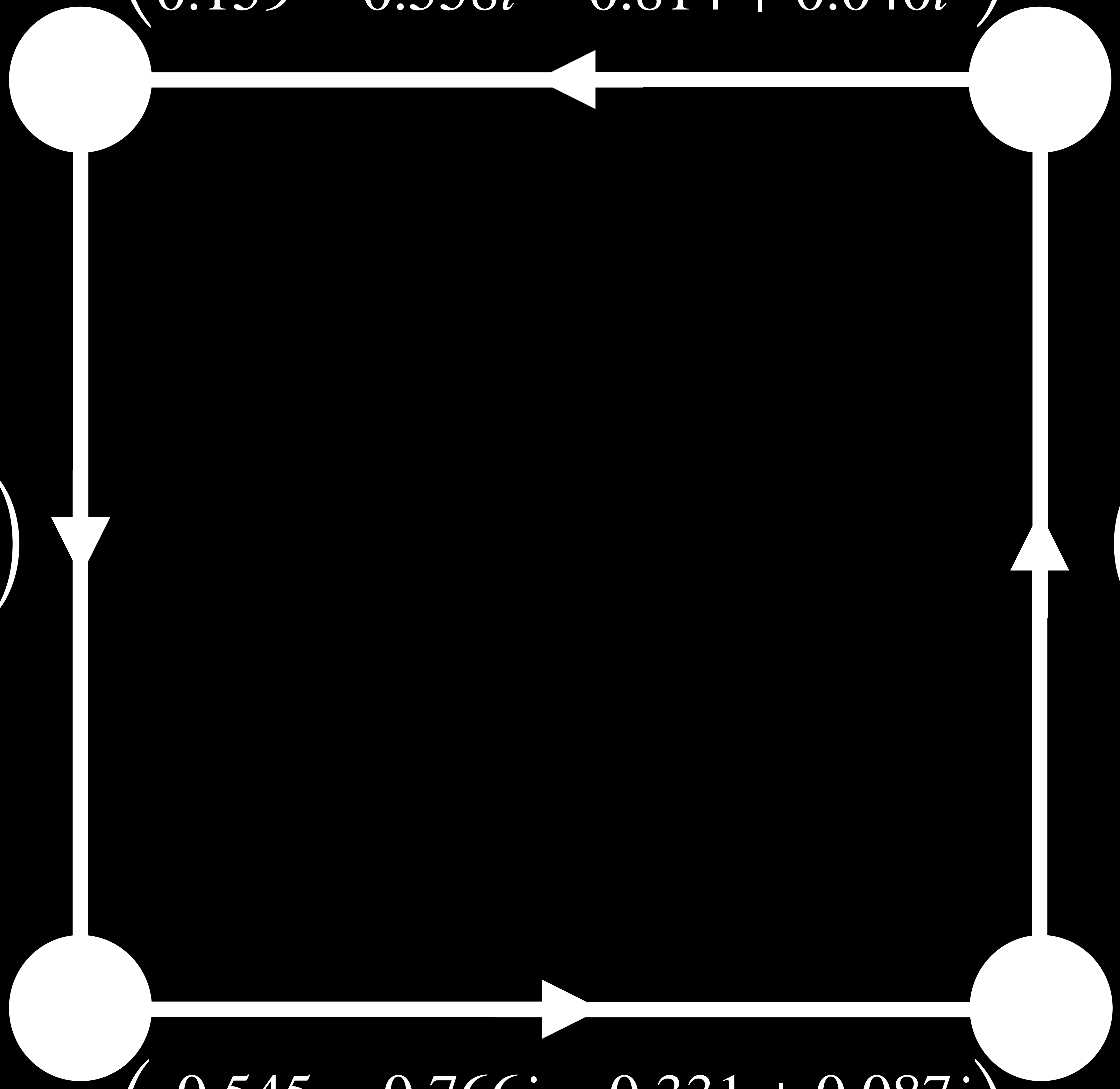
$$U_\mu(x) = \begin{pmatrix} 0.545 - 0.766i & 0.331 + 0.087i \\ -0.331 + 0.087i & 0.545 + 0.766i \end{pmatrix}$$

SU(2) example

Gauge symmetry

$$\begin{pmatrix} 0.814 - 0.040i & -0.159 - 0.558i \\ 0.159 - 0.558i & 0.814 + 0.040i \end{pmatrix}$$

$$\begin{pmatrix} 0.508 + 0.647i & -0.438 + 0.362i \\ 0.438 + 0.362i & 0.508 - 0.647i \end{pmatrix}$$



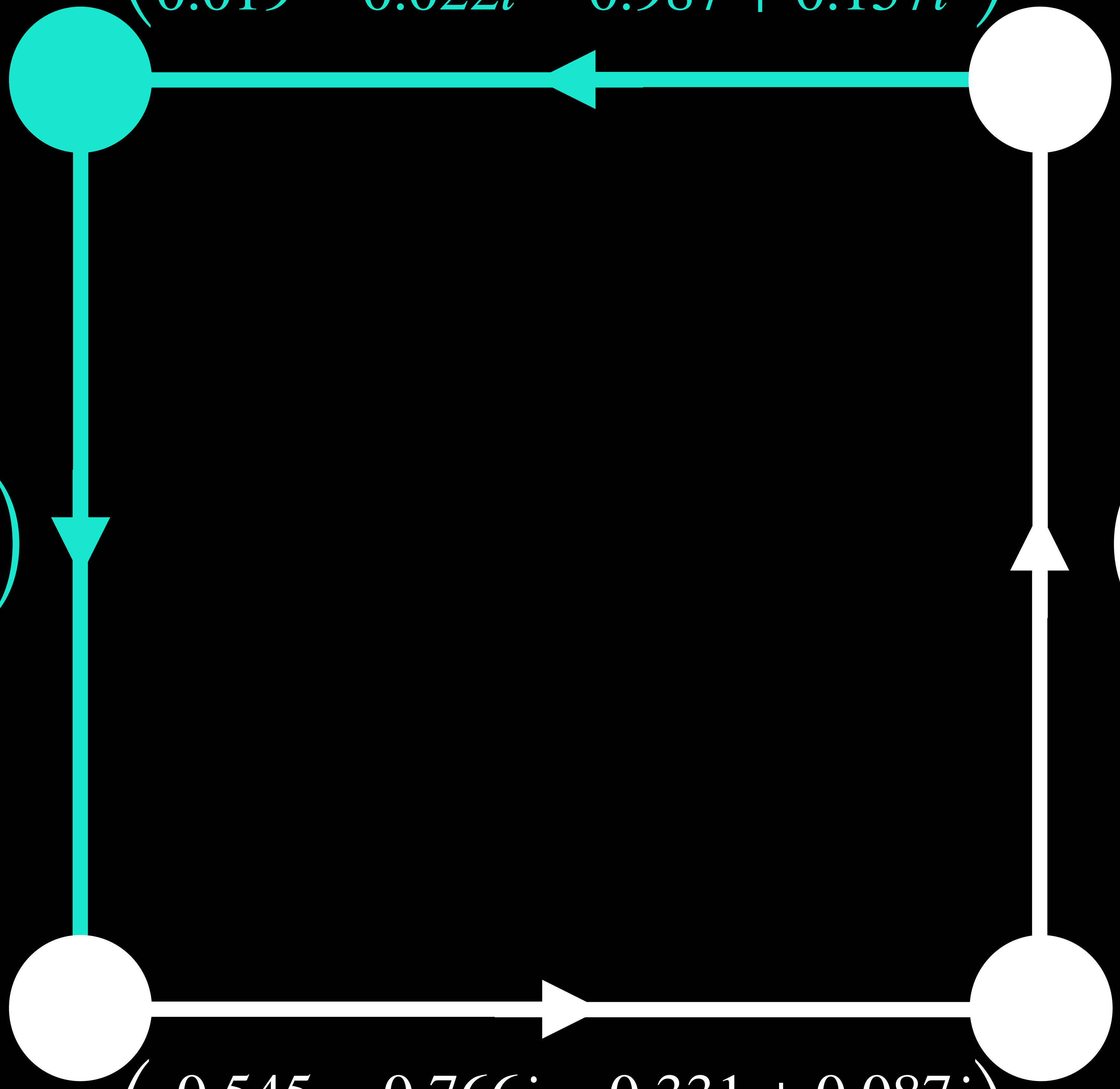
$$\begin{pmatrix} 0.978 + 0.185i & -0.040 - 0.093i \\ 0.040 - 0.093i & 0.978 - 0.185i \end{pmatrix}$$

$$\begin{pmatrix} 0.545 - 0.766i & 0.331 + 0.087i \\ -0.331 + 0.087i & 0.545 + 0.766i \end{pmatrix}$$

SU(2) example

Gauge symmetry

$$\begin{pmatrix} 0.987 - 0.157i & -0.019 - 0.022i \\ 0.019 - 0.022i & 0.987 + 0.157i \end{pmatrix}$$



$$\begin{pmatrix} 0.294 + 0.756i & 0.057 + 0.582i \\ -0.057 + 0.582i & 0.294 - 0.756i \end{pmatrix}$$

$$\begin{pmatrix} 0.978 + 0.185i & -0.040 - 0.093i \\ 0.040 - 0.093i & 0.978 - 0.185i \end{pmatrix}$$

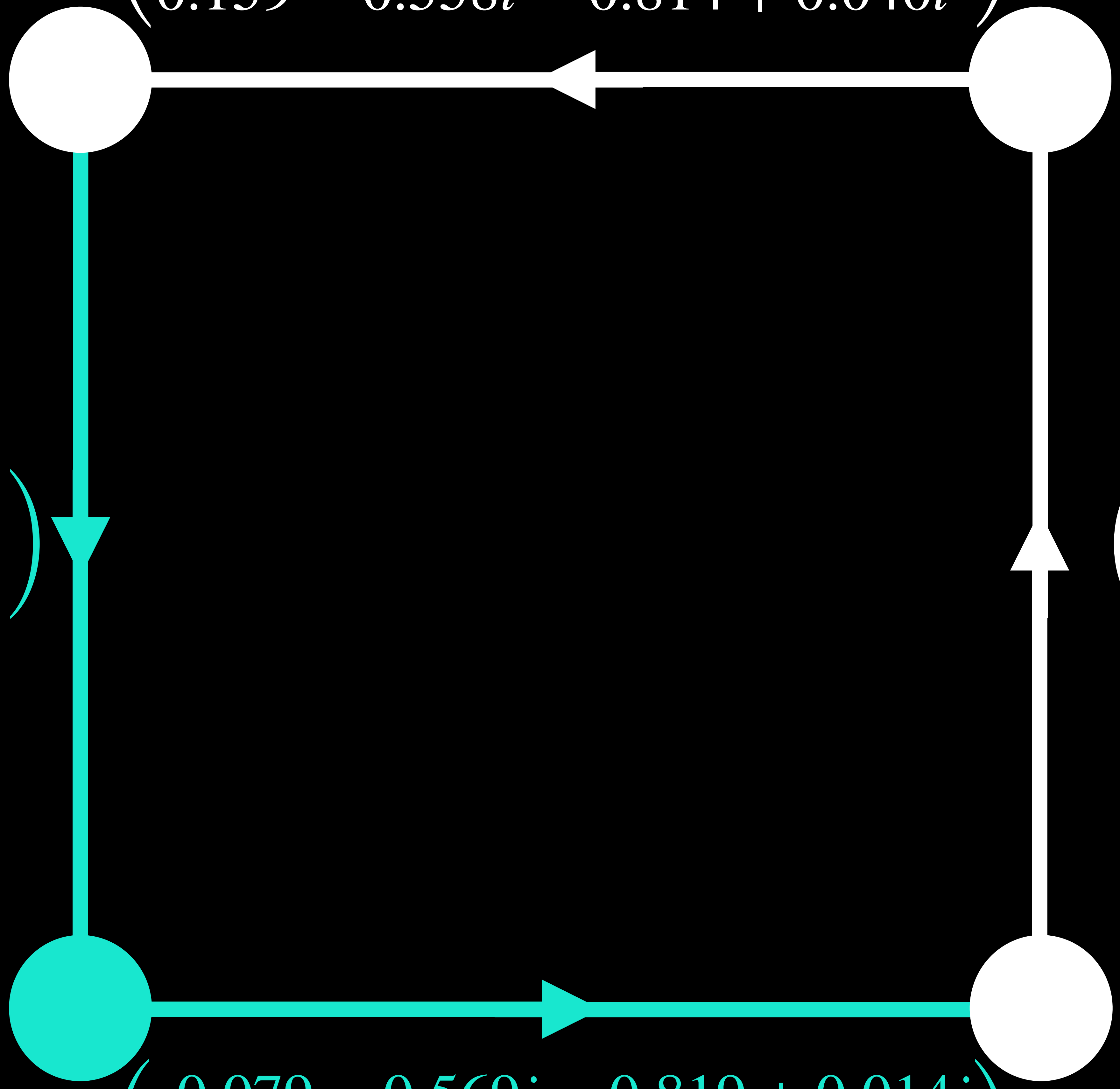
$$\begin{pmatrix} 0.545 - 0.766i & 0.331 + 0.087i \\ -0.331 + 0.087i & 0.545 + 0.766i \end{pmatrix}$$

SU(2) example

Gauge symmetry

$$\begin{pmatrix} 0.814 - 0.040i & -0.159 - 0.558i \\ 0.159 - 0.558i & 0.814 + 0.040i \end{pmatrix}$$

$$\begin{pmatrix} 0.911 + 0.379i & 0.035 + 0.157i \\ -0.035 + 0.157i & 0.911 - 0.379i \end{pmatrix}$$



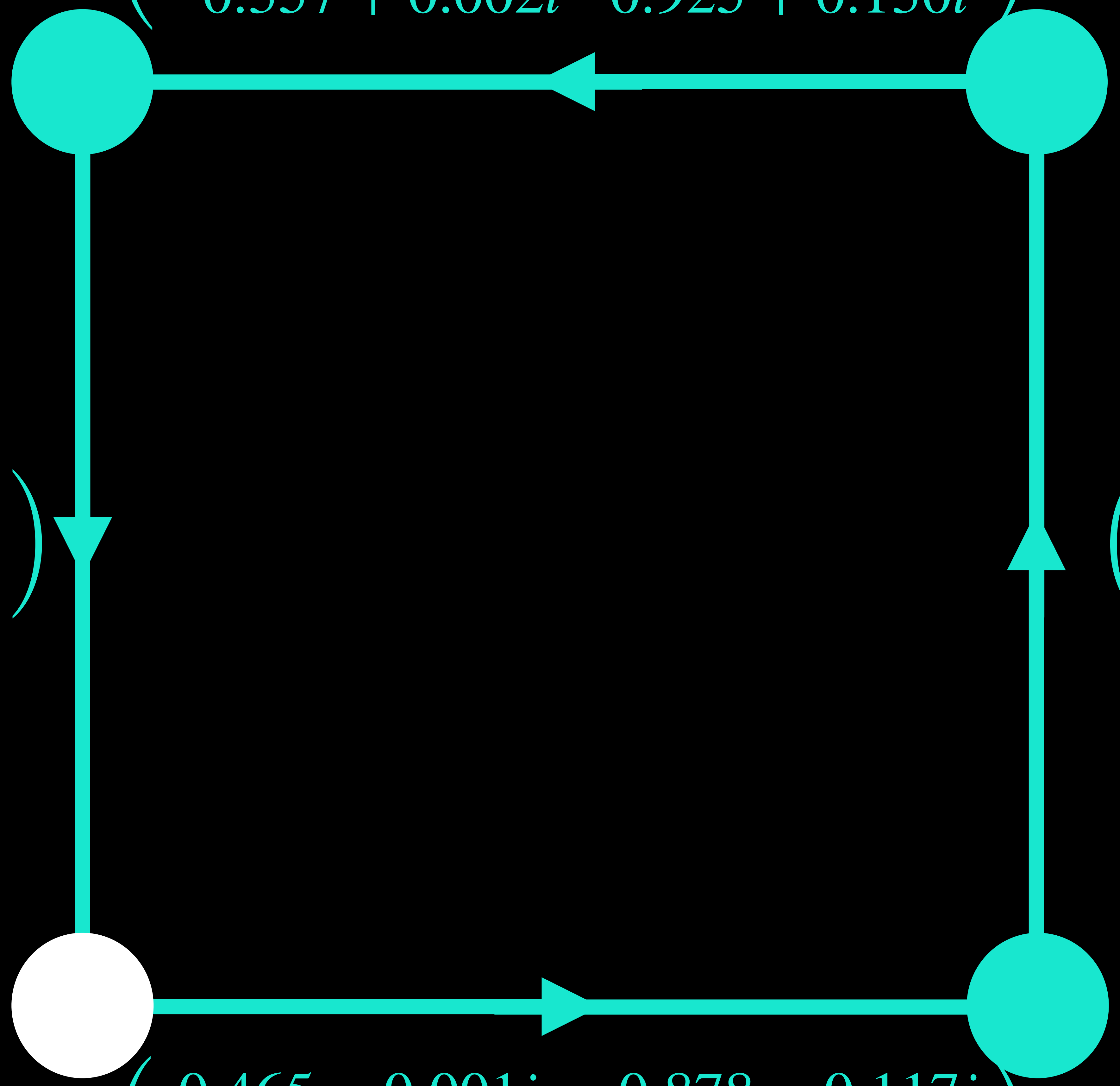
$$\begin{pmatrix} 0.978 + 0.185i & -0.040 - 0.093i \\ 0.040 - 0.093i & 0.978 - 0.185i \end{pmatrix}$$

$$\begin{pmatrix} 0.079 - 0.569i & 0.819 + 0.014i \\ -0.819 + 0.014i & 0.079 + 0.569i \end{pmatrix}$$

SU(2) example

Gauge symmetry

$$\begin{pmatrix} 0.925 - 0.130i & 0.357 + 0.002i \\ -0.357 + 0.002i & 0.925 + 0.130i \end{pmatrix}$$



$$\begin{pmatrix} 0.294 + 0.756i & 0.057 + 0.582i \\ -0.057 + 0.582i & 0.294 - 0.756i \end{pmatrix}$$

$$\begin{pmatrix} 0.272 + 0.343i & 0.711 + 0.550i \\ -0.711 + 0.550i & 0.272 - 0.343i \end{pmatrix}$$

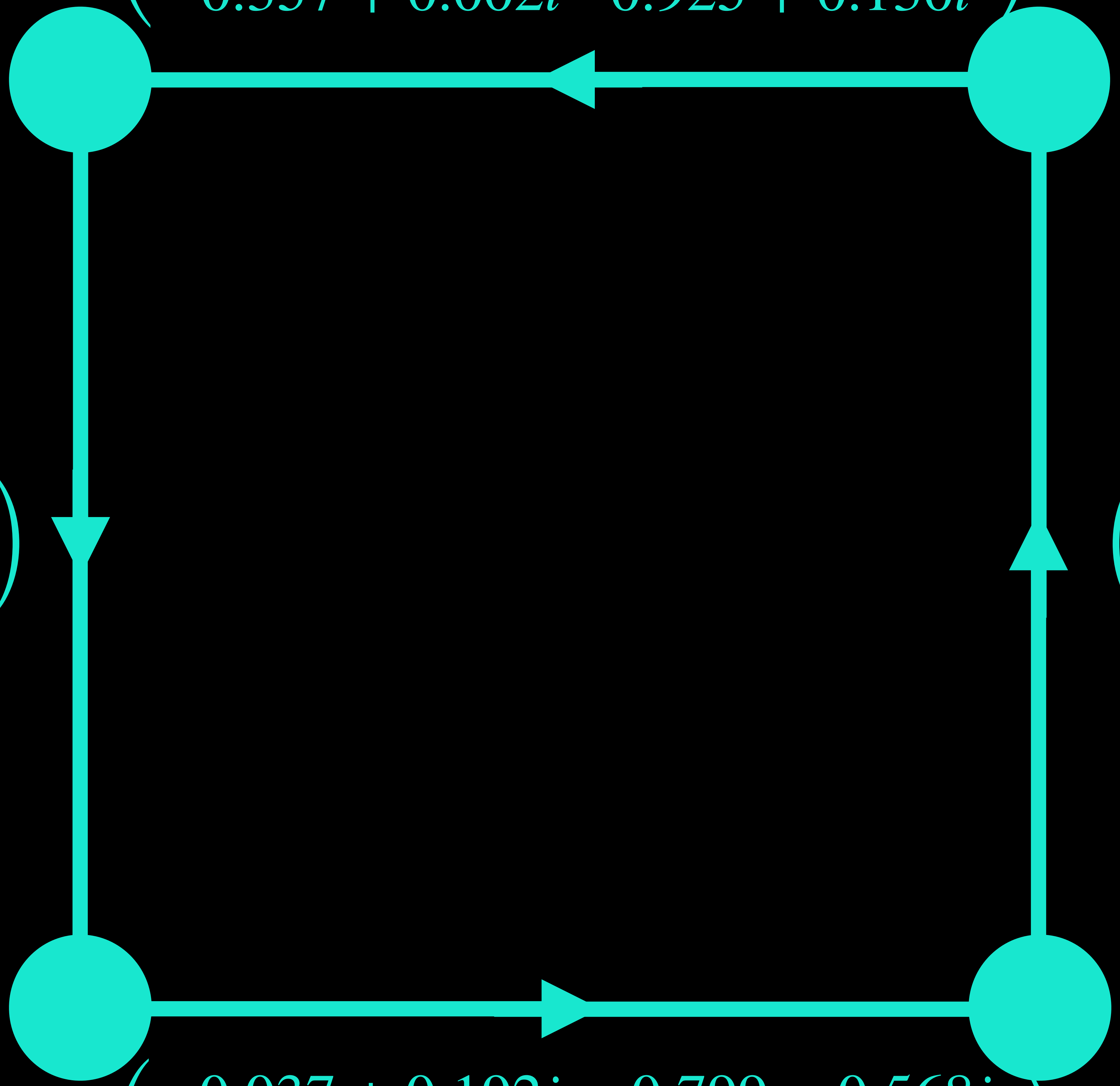
$$\begin{pmatrix} 0.465 - 0.001i & 0.878 - 0.117i \\ -0.878 - 0.117i & 0.465 + 0.001i \end{pmatrix}$$

SU(2) example

Gauge symmetry

$$\begin{pmatrix} 0.925 - 0.130i & 0.357 + 0.002i \\ -0.357 + 0.002i & 0.925 + 0.130i \end{pmatrix}$$

$$\begin{pmatrix} 0.693 + 0.227i & 0.302 + 0.614i \\ -0.302 + 0.614i & 0.693 - 0.227i \end{pmatrix}$$

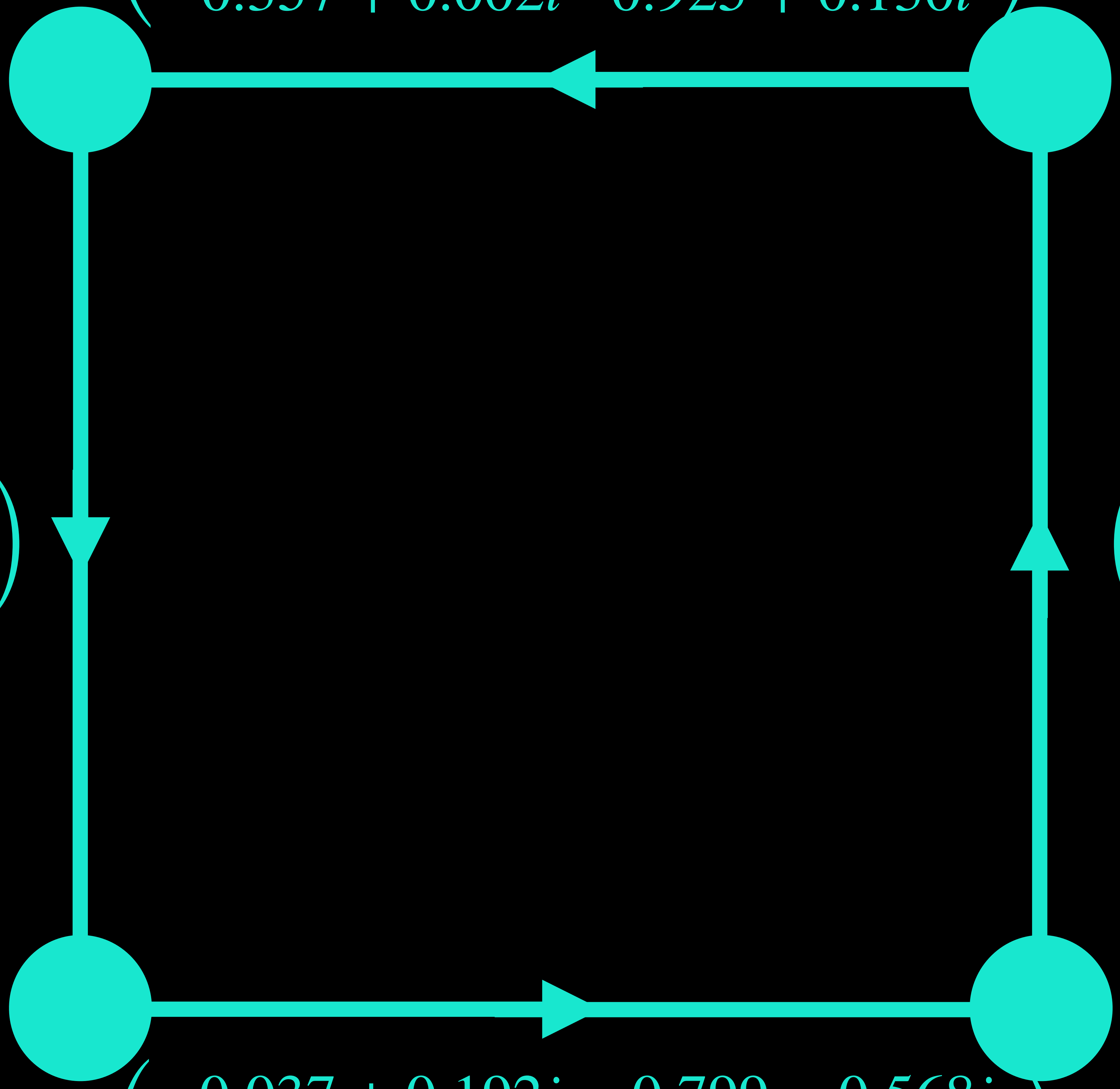


$$\begin{pmatrix} 0.272 + 0.343i & 0.711 + 0.550i \\ -0.711 + 0.550i & 0.272 - 0.343i \end{pmatrix}$$

$$\begin{pmatrix} -0.037 + 0.192i & 0.799 - 0.568i \\ -0.799 - 0.568i & -0.037 - 0.192i \end{pmatrix}$$

SU(2) example

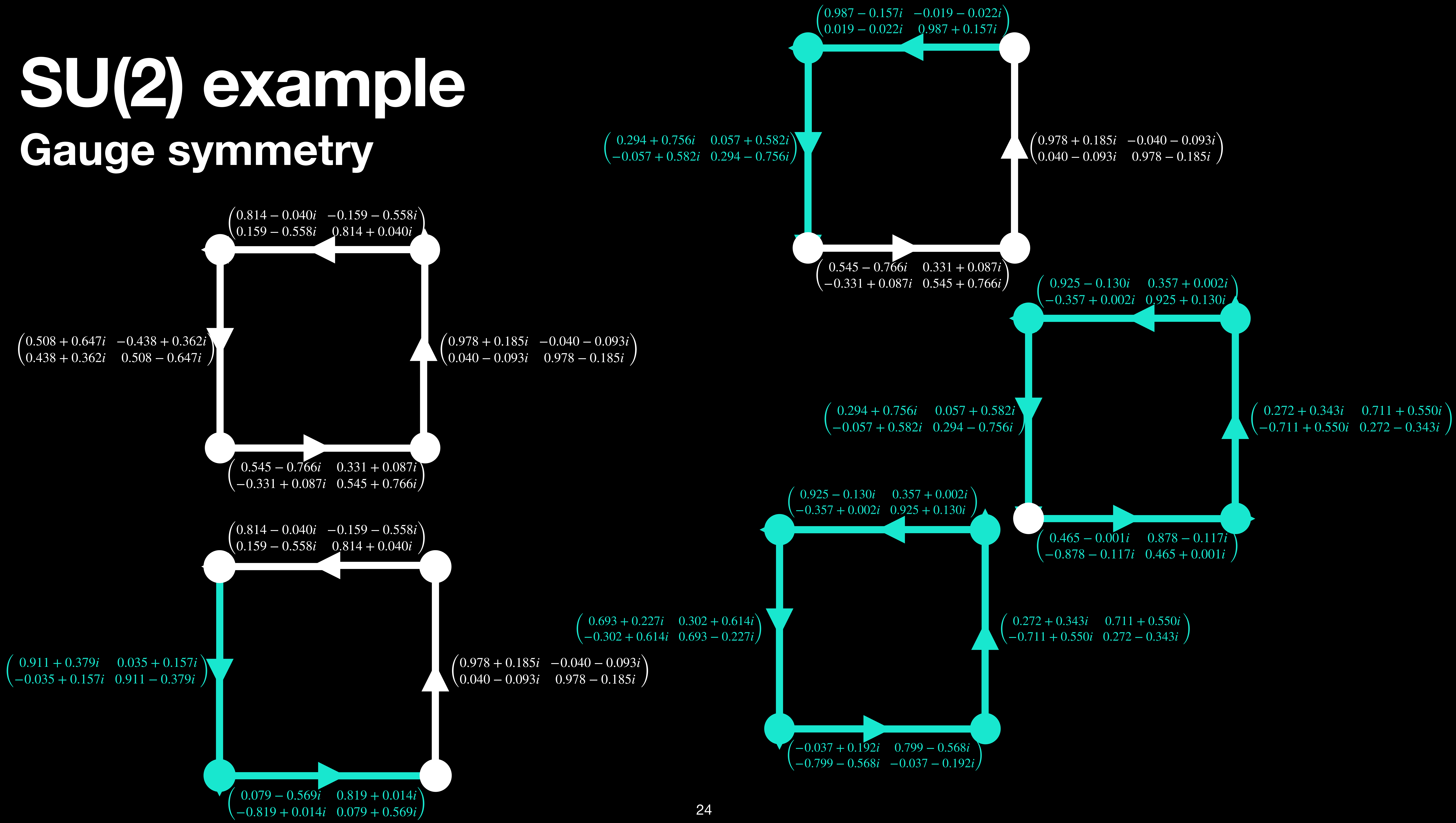
Gauge symmetry



$$U_\mu(x) \rightarrow \Omega(x)U_\mu(x)\Omega^\dagger(x+\mu)$$

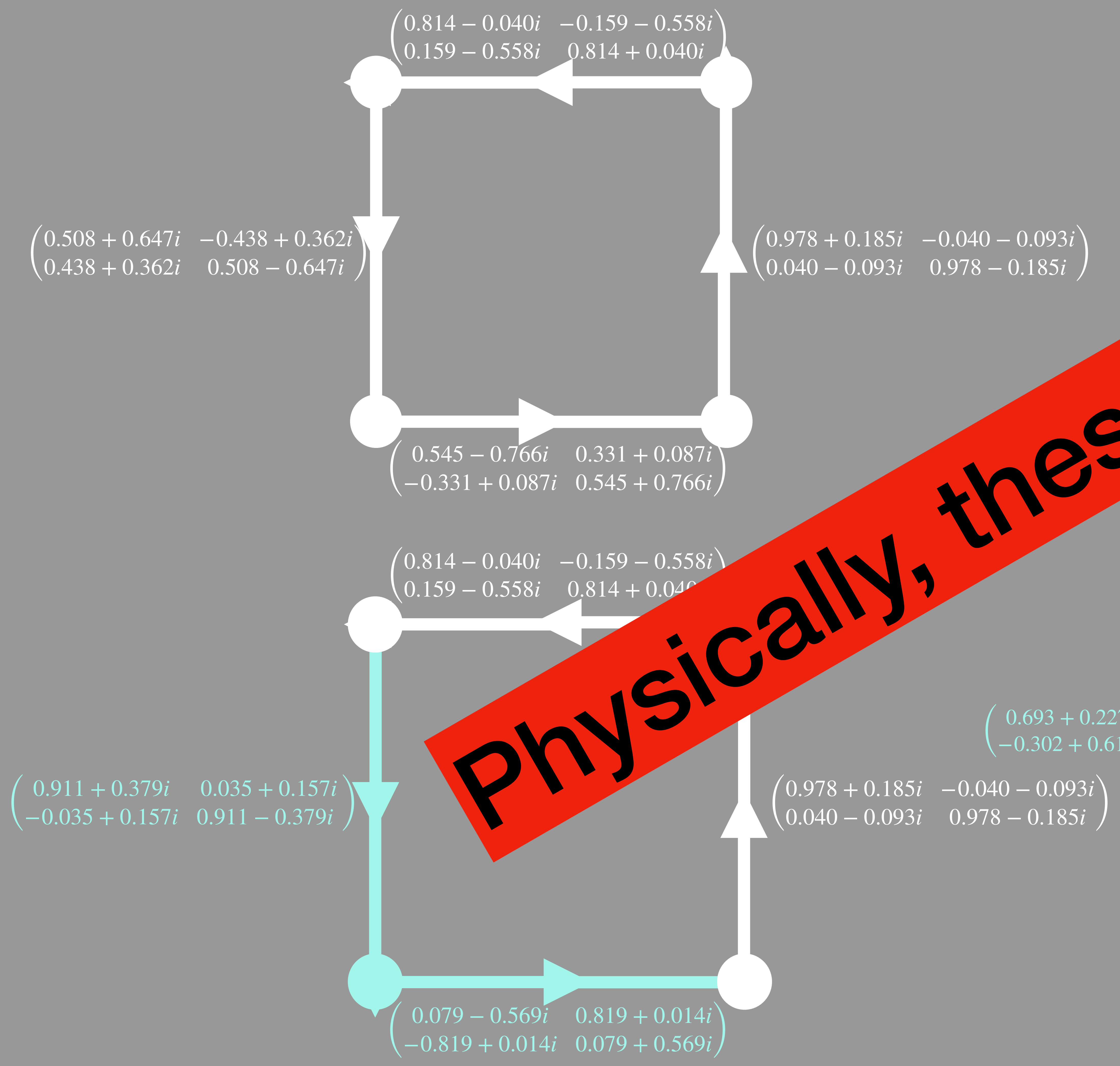
SU(2) example

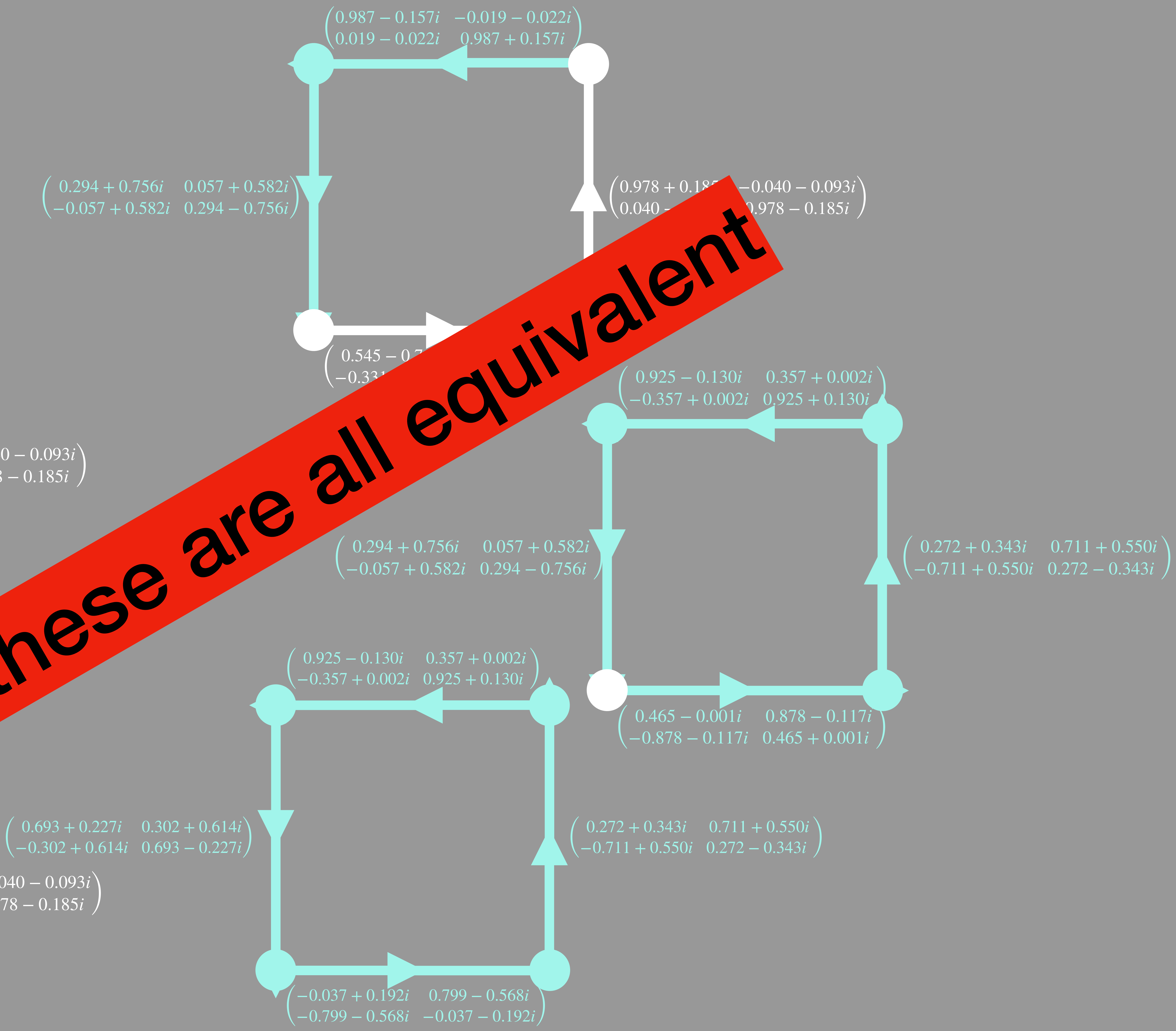
Gauge symmetry



SU(2) example

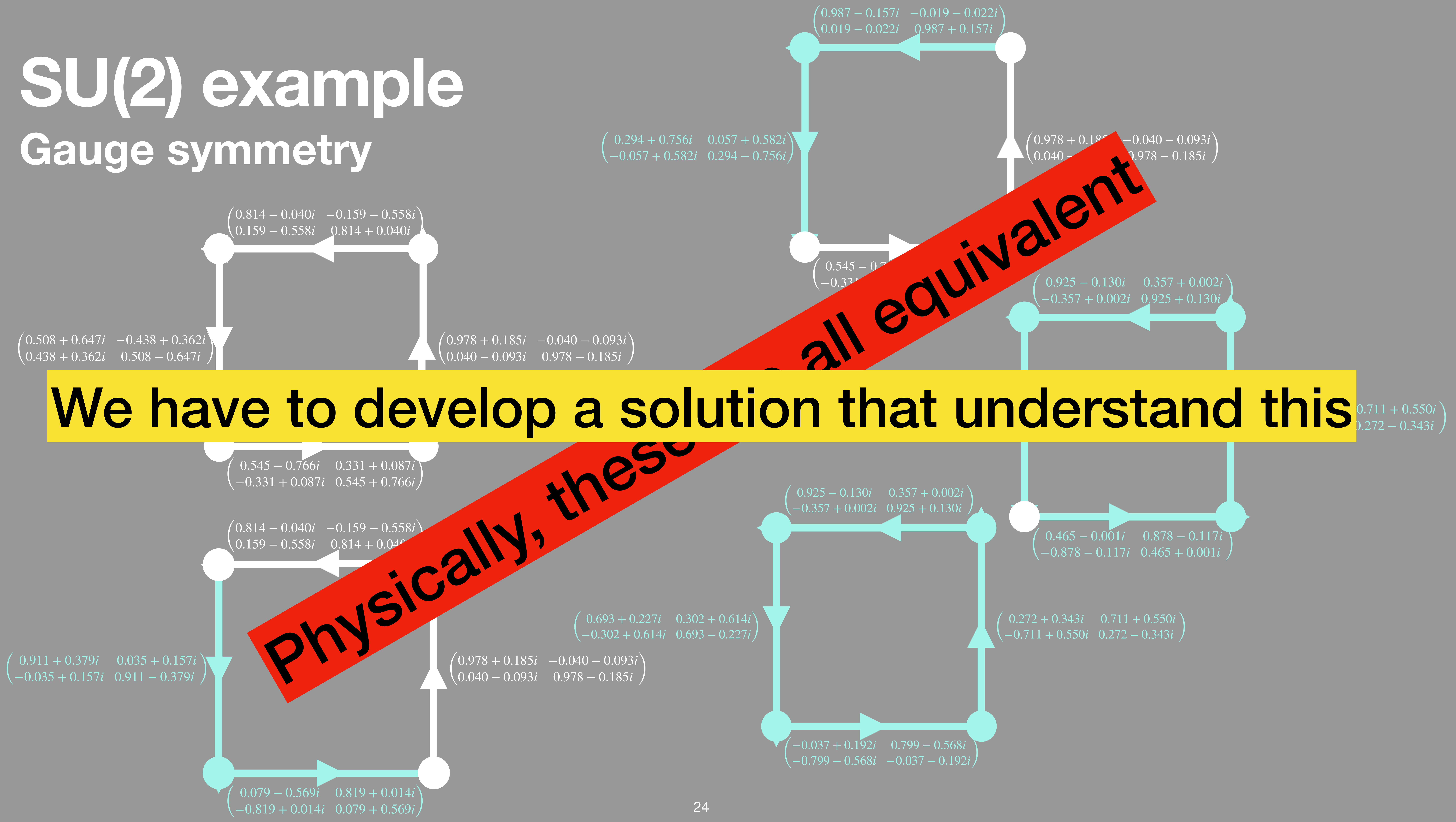
Gauge symmetry





SU(2) example

Gauge symmetry

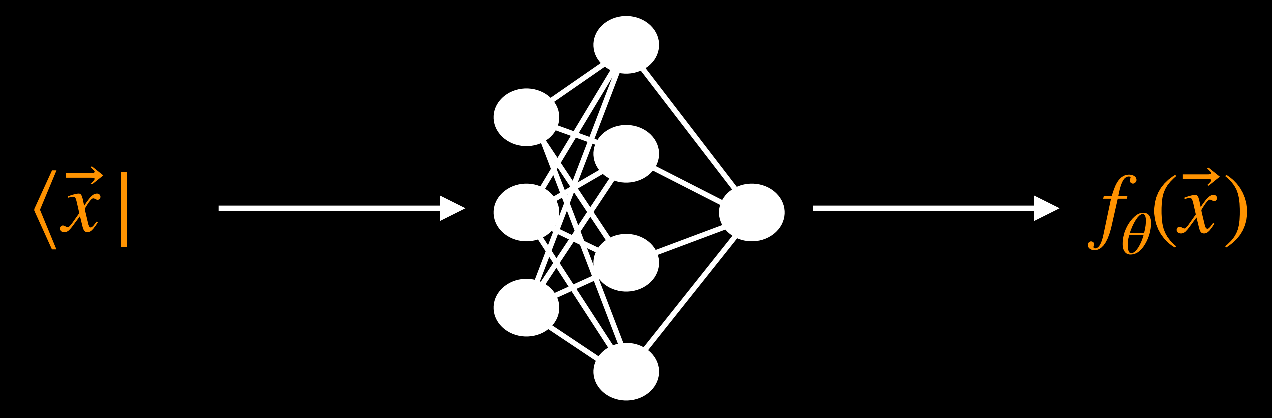


Physically, these all equivalent

We have to develop a solution that understand this

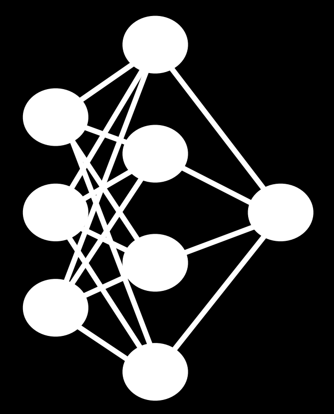
Some ansatze

Physics-informed

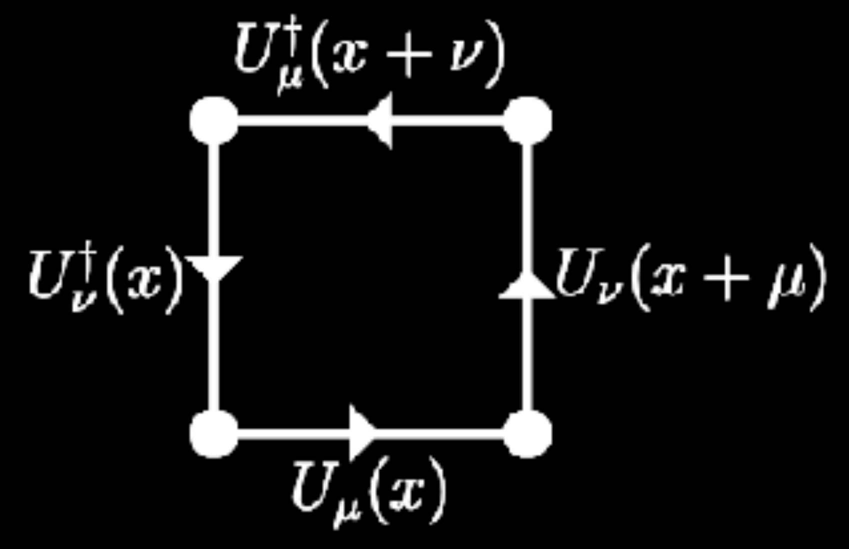


Greensite, J. P. (1979). Calculation of the Yang-Mills vacuum wave functional. *Nuclear Physics B*, *158*(2-3), 469-496.
Chin, S. A., Van Roosmalen, O. S., Umland, E. A., & Koonin, S. E. (1985). Exact ground-state properties of the SU (2) Hamiltonian lattice gauge theory. *Physical Review D*, *31*(12), 3201.

Physics-informed

$\langle \vec{x} |$ $\longrightarrow$  $\longrightarrow$ $f_\theta(\vec{x})$

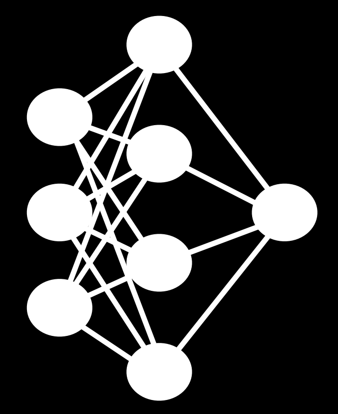
$$f(\mathbf{U}) = \prod_p e^{\alpha \frac{1}{2} \mathrm{Tr}\left(P_p\right)}$$

$\sum \mathrm{Tr}$ 

Greensite, J. P. (1979). Calculation of the Yang-Mills vacuum wave functional. *Nuclear Physics B*, *158*(2-3), 469-496.
Chin, S. A., Van Roosmalen, O. S., Umland, E. A., & Koonin, S. E. (1985). Exact ground-state properties of the SU (2) Hamiltonian lattice gauge theory. *Physical Review D*, *31*(12), 3201.

Physics-informed

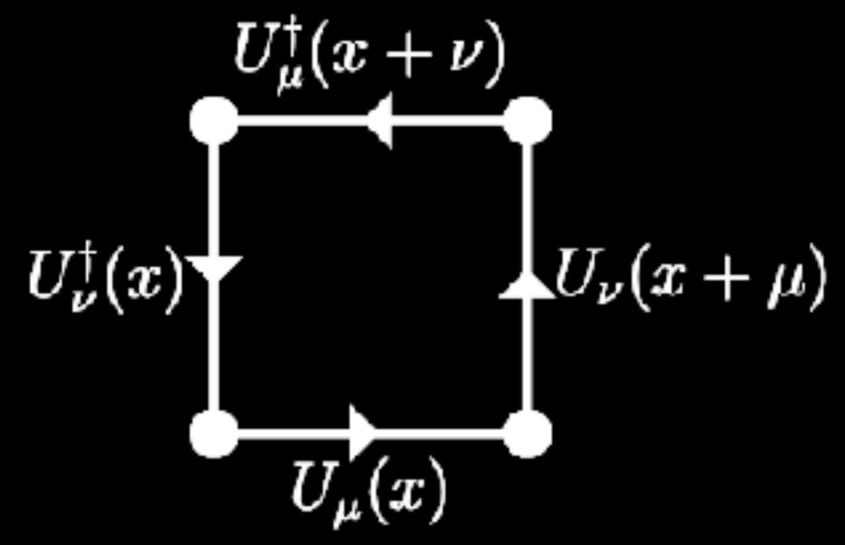
$\langle \vec{x} |$ $\longrightarrow$



$\longrightarrow$ $f_\theta(\vec{x})$

$$f(\mathbf{U}) = \prod_p e^{\alpha \frac{1}{2} \mathrm{Tr}\left(P_p\right)}$$

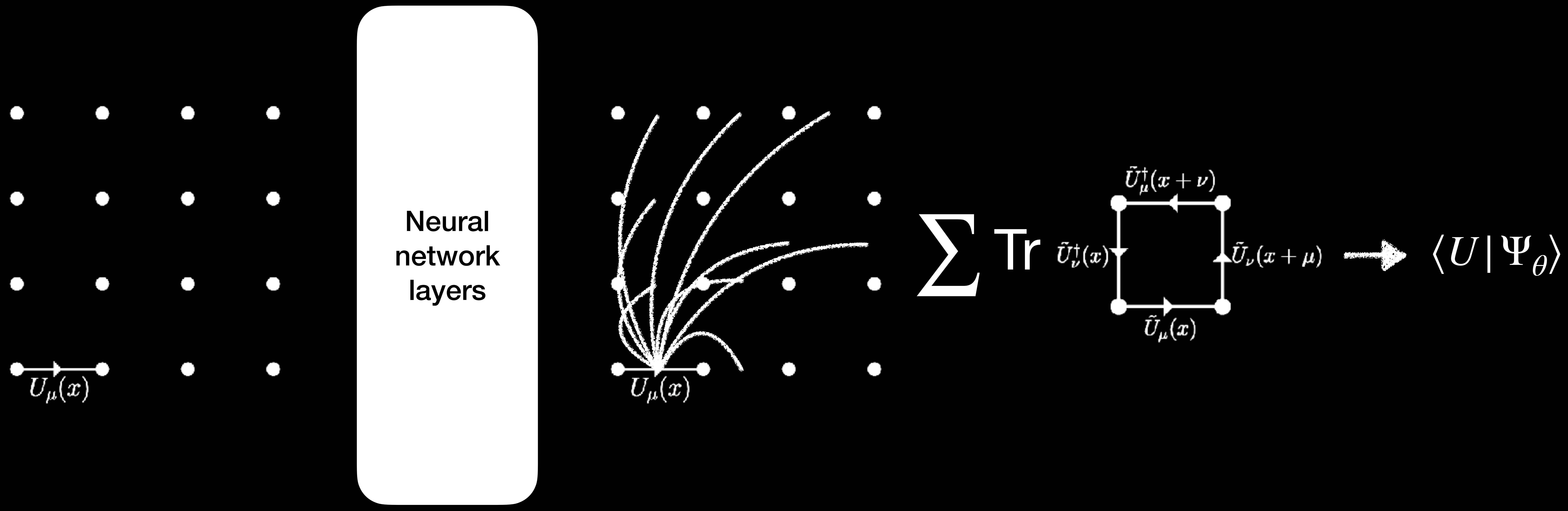
$\sum \mathrm{Tr}$



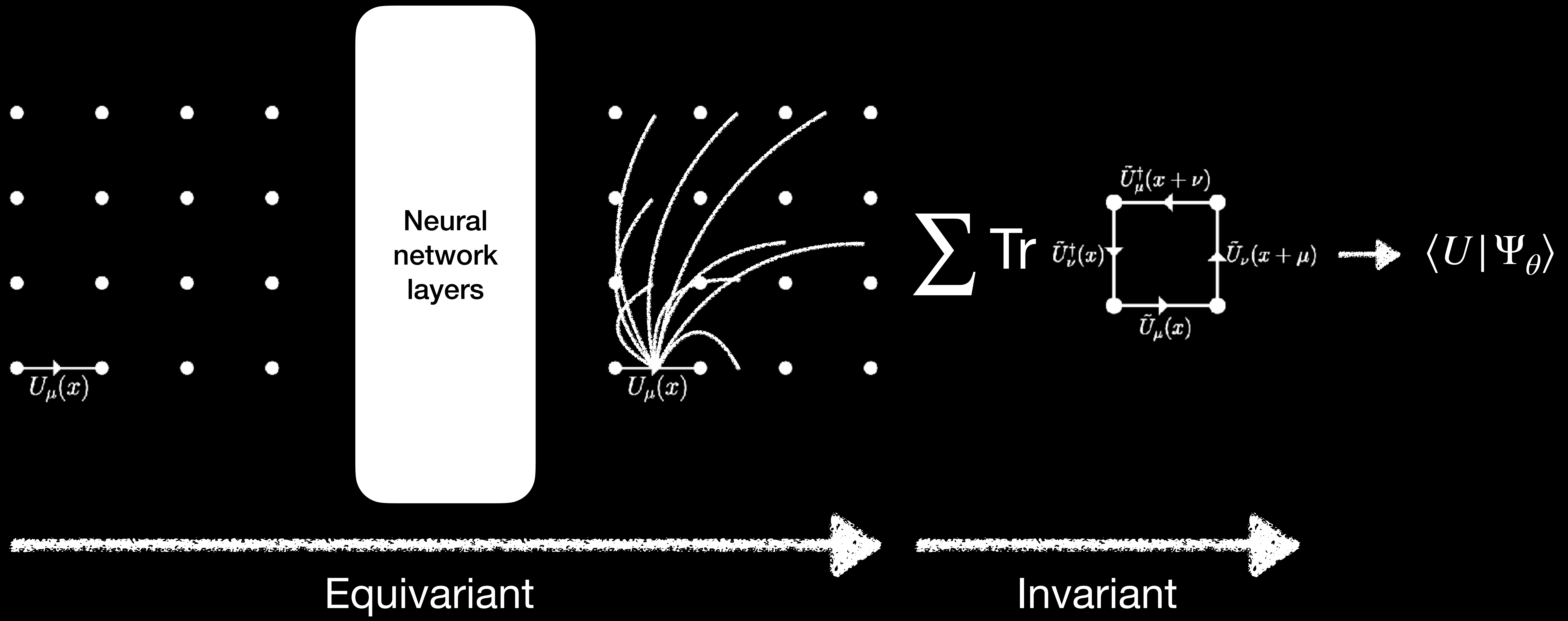
Reliance on hand-crafted features

Greensite, J. P. (1979). Calculation of the Yang-Mills vacuum wave functional. *Nuclear Physics B*, *158*(2-3), 469-496.
Chin, S. A., Van Roosmalen, O. S., Umland, E. A., & Koonin, S. E. (1985). Exact ground-state properties of the SU (2) Hamiltonian lattice gauge theory. *Physical Review D*, *31*(12), 3201.

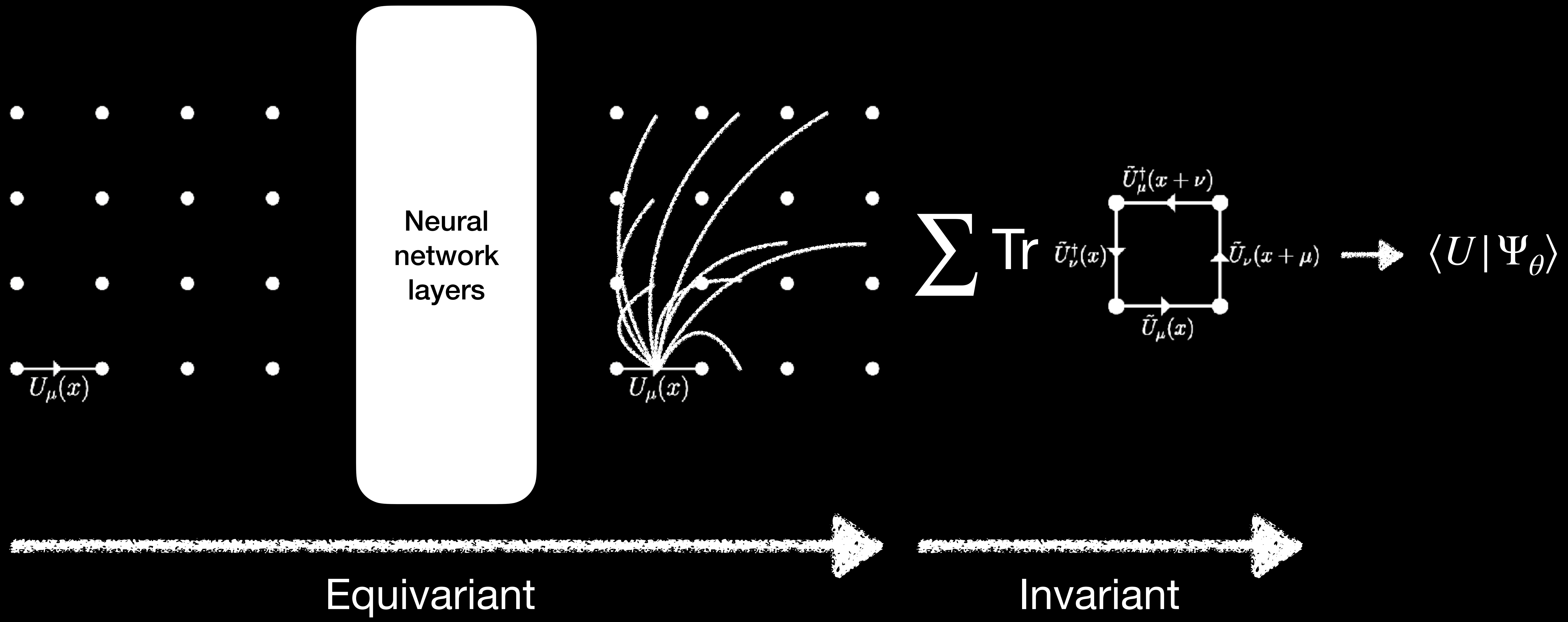
Neural network + physics-informed



Neural network + physics-informed



Neural network + physics-informed

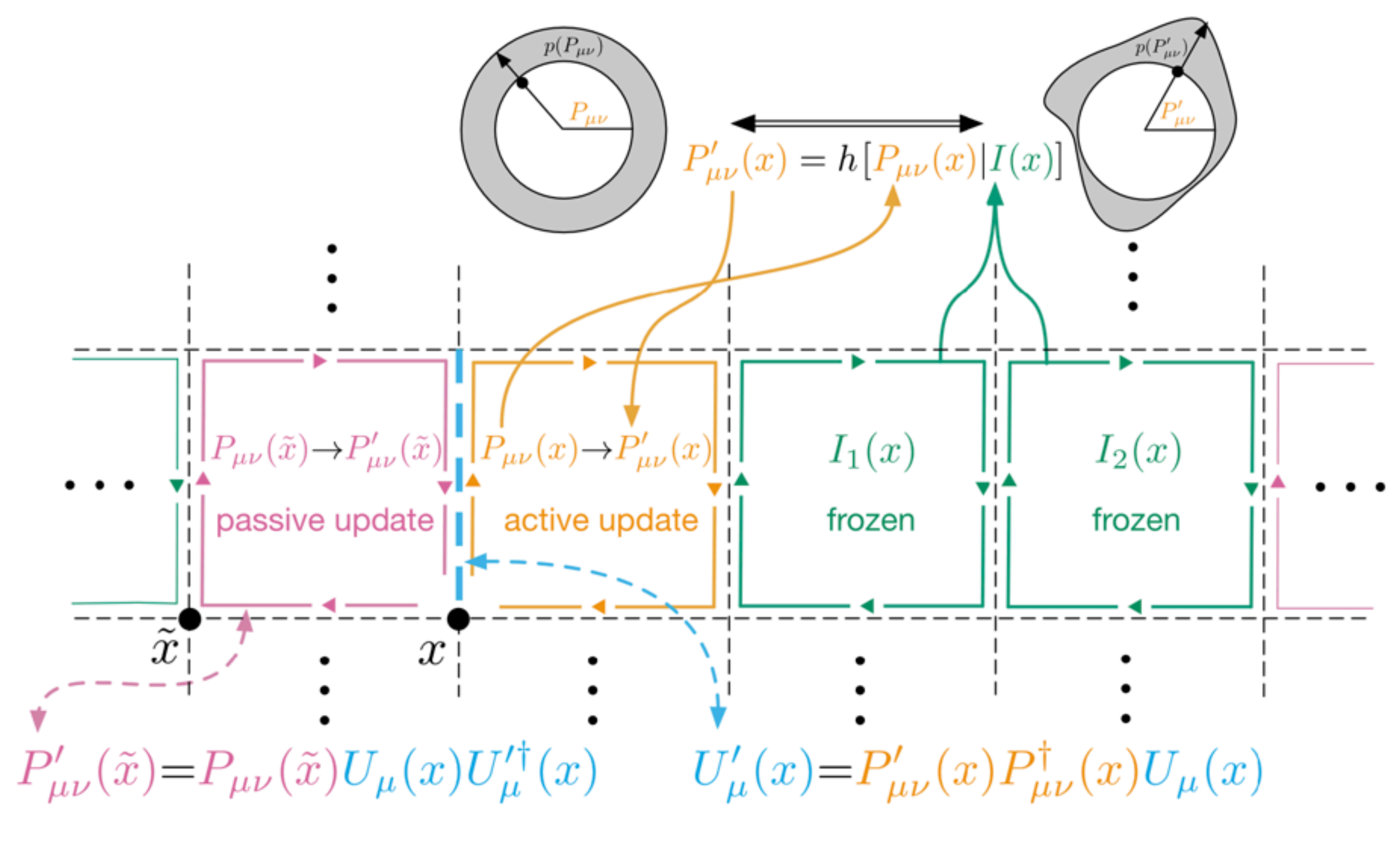


$U_\mu(x) \to \Omega(x) U_\mu(x) \Omega^\dagger(x+\mu)$ **before and after NN**

Choices for network

All gauge equivariant

- “Normalising flow networks”



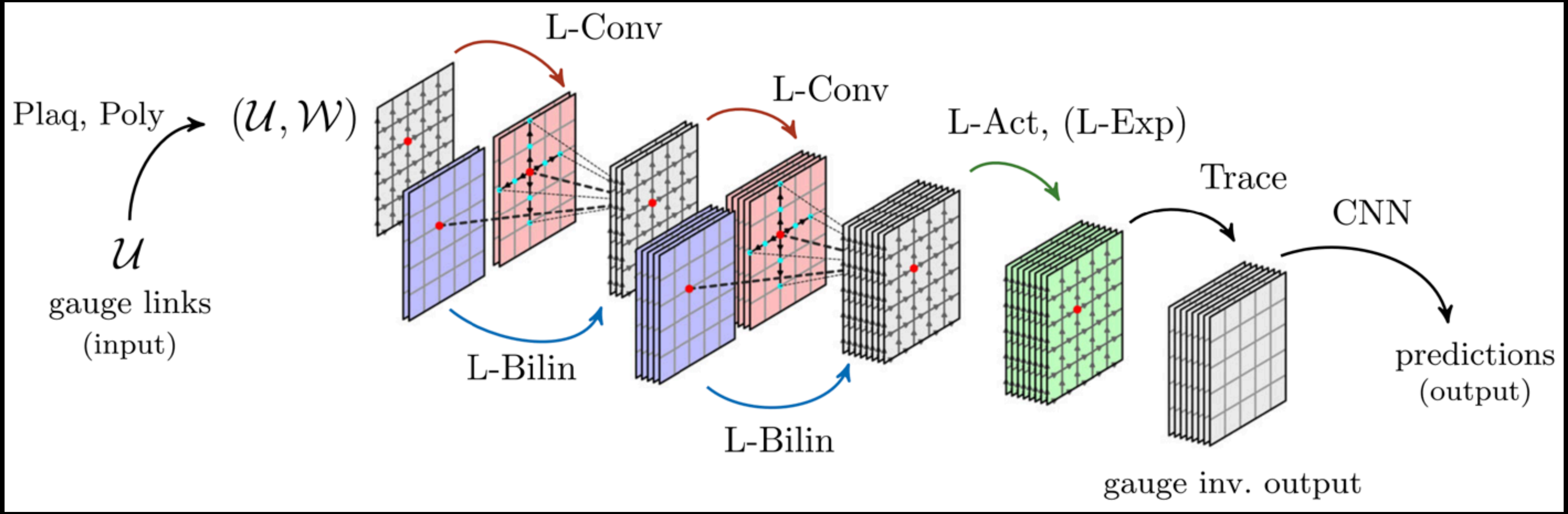
Boyda, D., Kanwar, G., Racanière, S., Rezende, D. J., Albergo, M. S., Cranmer, K., ... & Shanahan, P. E. (2021). Sampling using SU (N) gauge equivariant flows. *Physical Review D*, *103*(7), 074504.

Kanwar, G., Albergo, M. S., Boyda, D., Cranmer, K., Hackett, D. C., Racaniere, S., ... & Shanahan, P. E. (2020). Equivariant flow-based sampling for lattice gauge theory. *Physical Review Letters*, *125*(12), 121601.

Choices for network

All gauge equivariant

- Lattice Gauge Equivariant Convolutional Neural Networks (LCNN)



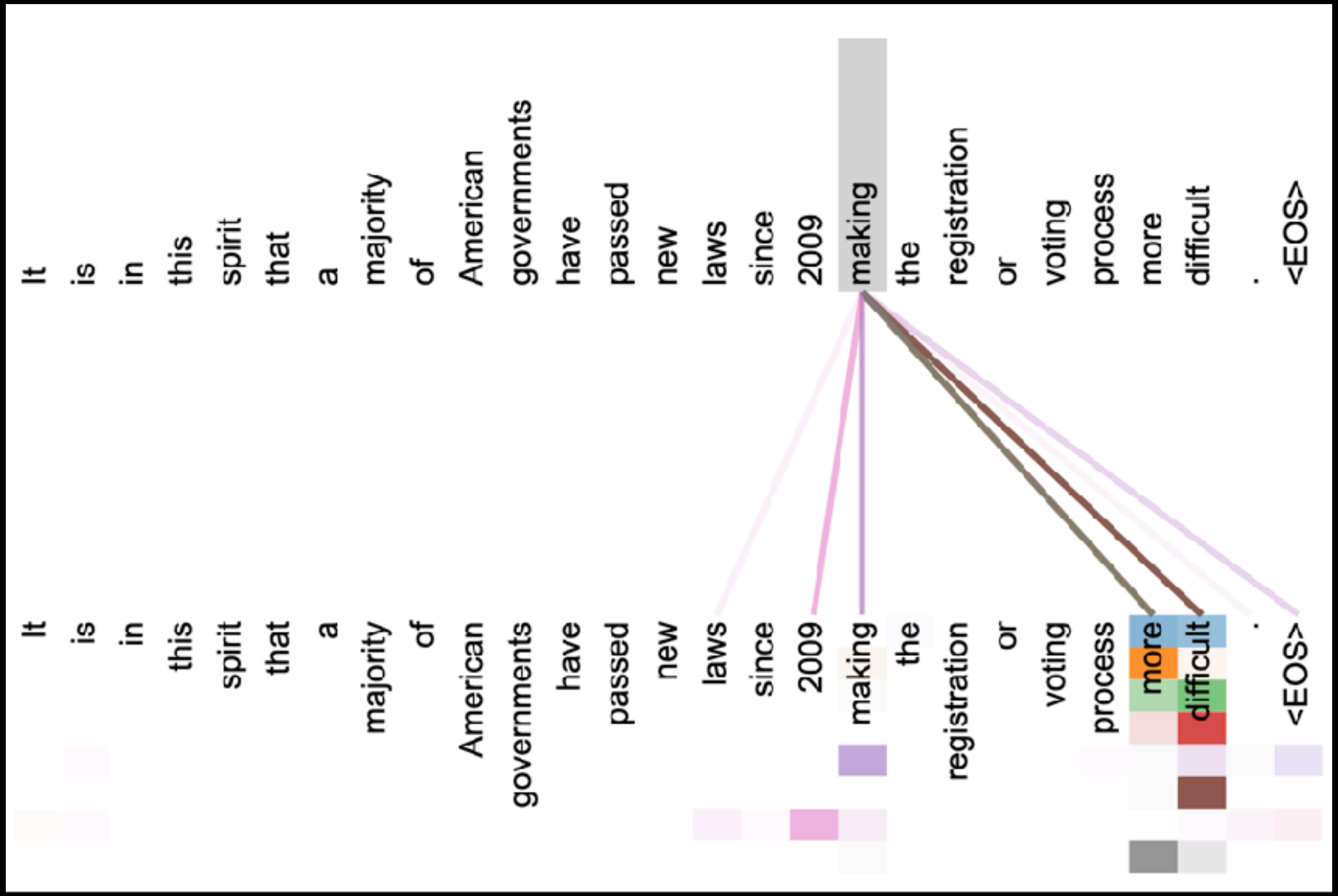
Favoni, M., Ipp, A., Müller, D. I., & Schuh, D. (2022). Lattice gauge equivariant convolutional neural networks. *Physical Review Letters*, *128*(3), 032003.

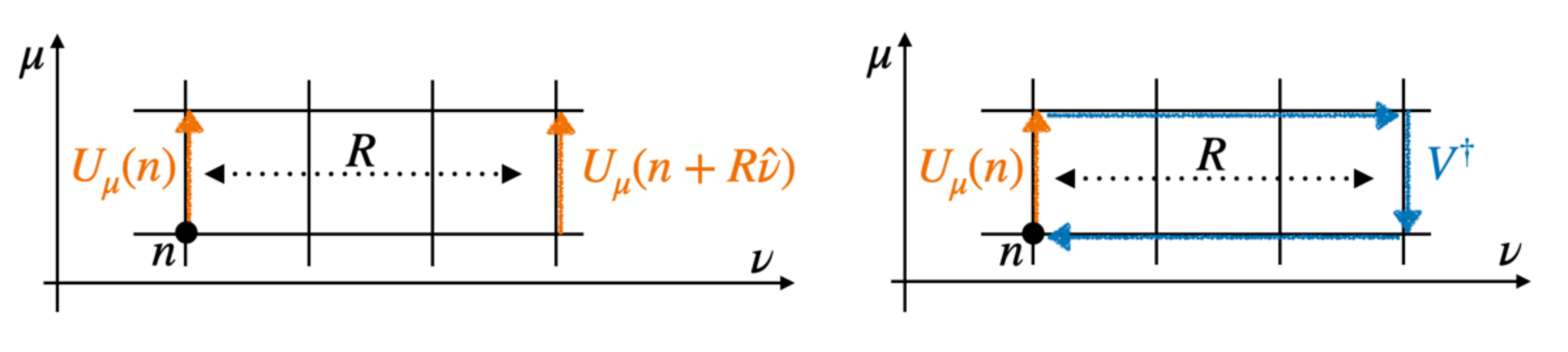
Choices for network

All gauge equivariant

- Gauge equivariant transformer

Vaswani, A., ... & Polosukhin, I. (2017). Attention is all you need



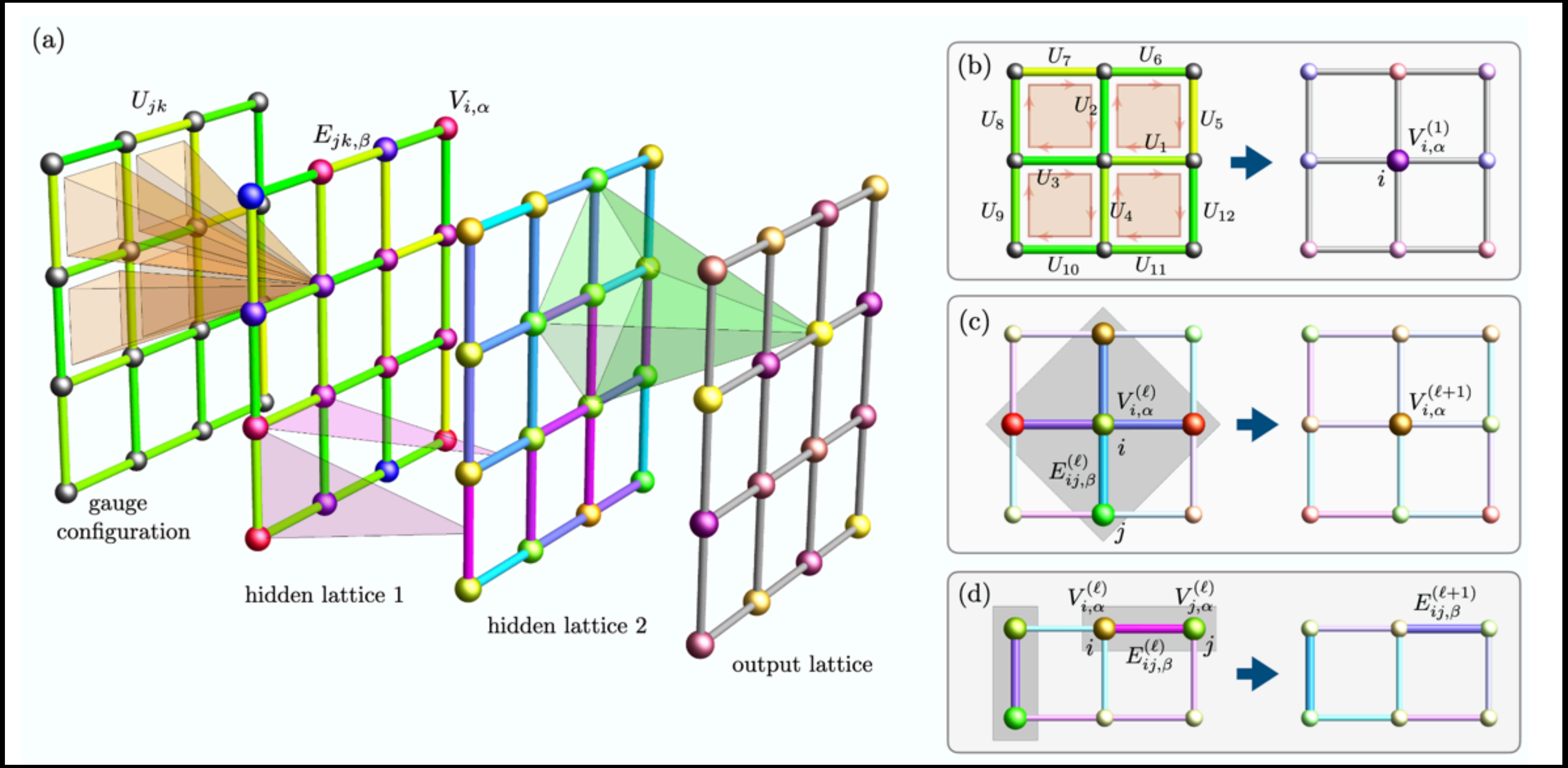


Nagai, Y., Ohno, H., & Tomiya, A. (2025). CASK: A Gauge Covariant Transformer for Lattice Gauge Theory. *arXiv preprint arXiv:2501.16955*.

Choices for network

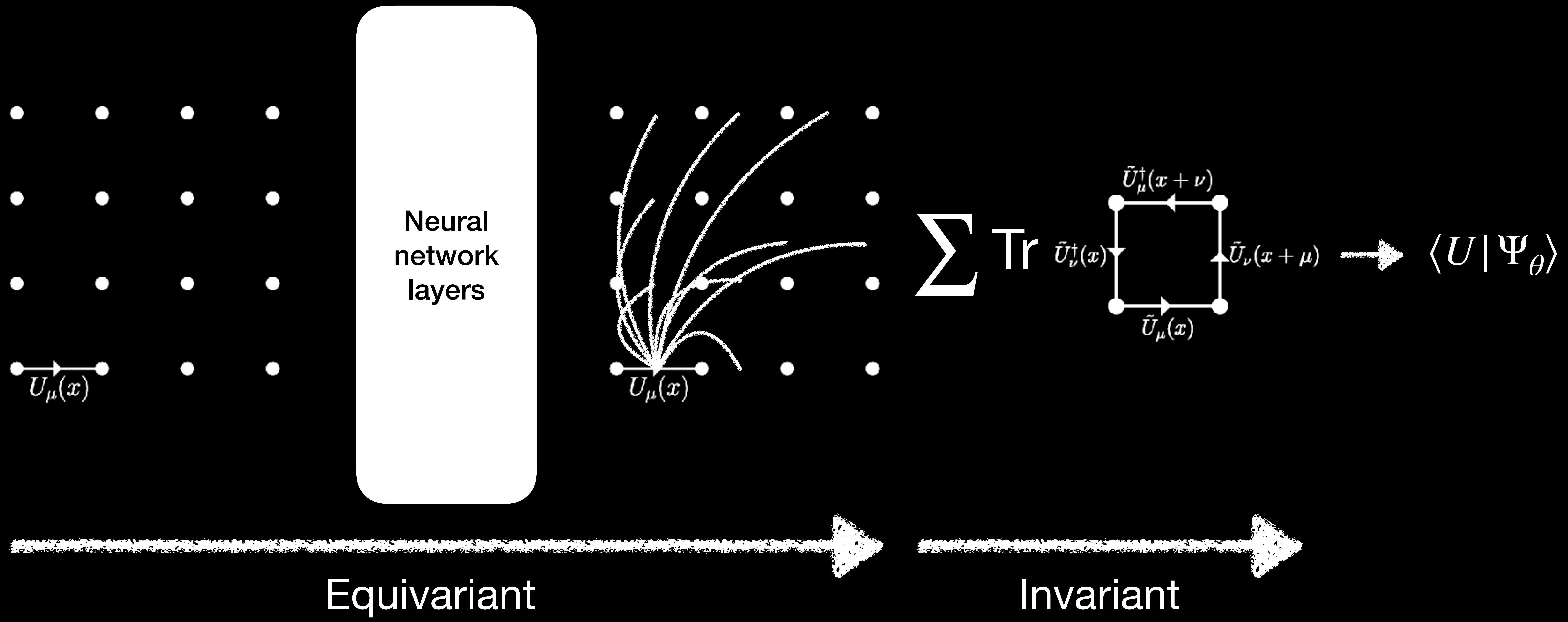
All gauge equivariant

- Gauge equivariant graph neural network

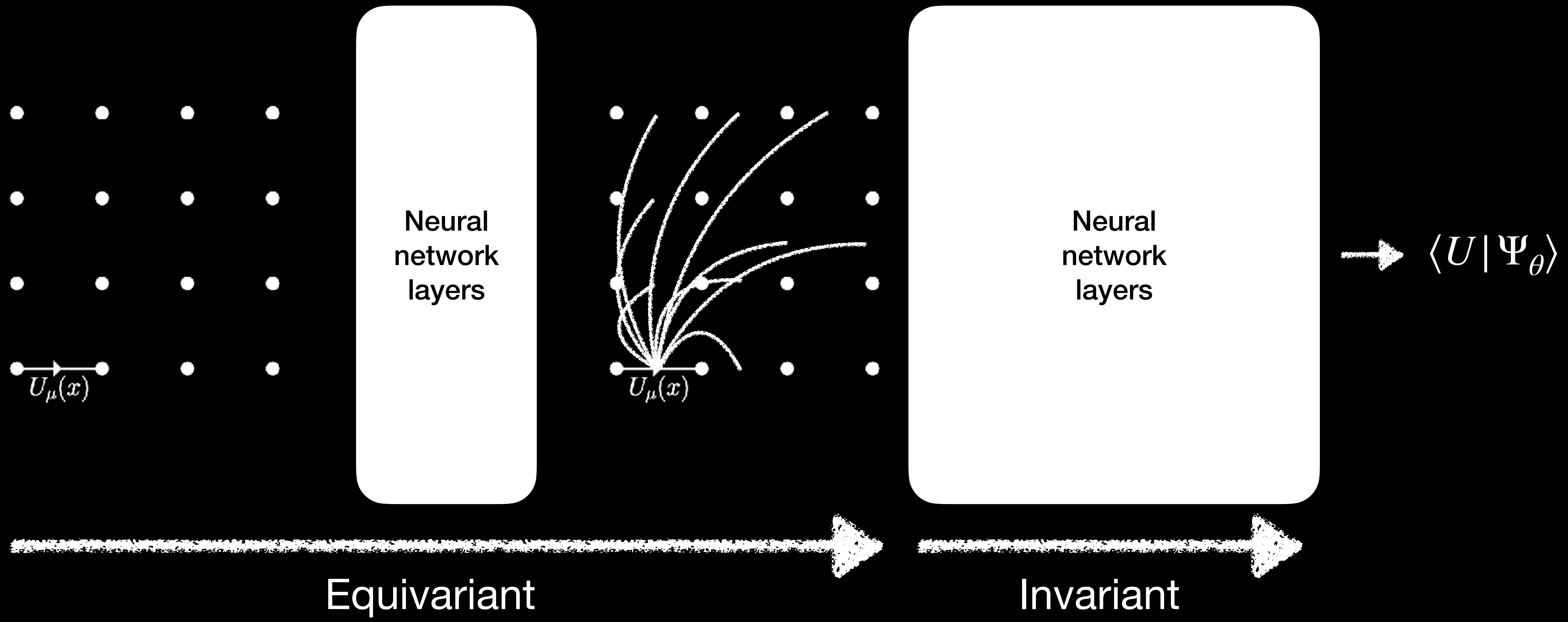


Rayat, A., Li, Y., & Chern, G. W. (2026). Gauge-Equivariant Graph Neural Networks for Lattice Gauge Theories. *arXiv preprint arXiv:2604.20797*.

Neural network + physics-informed



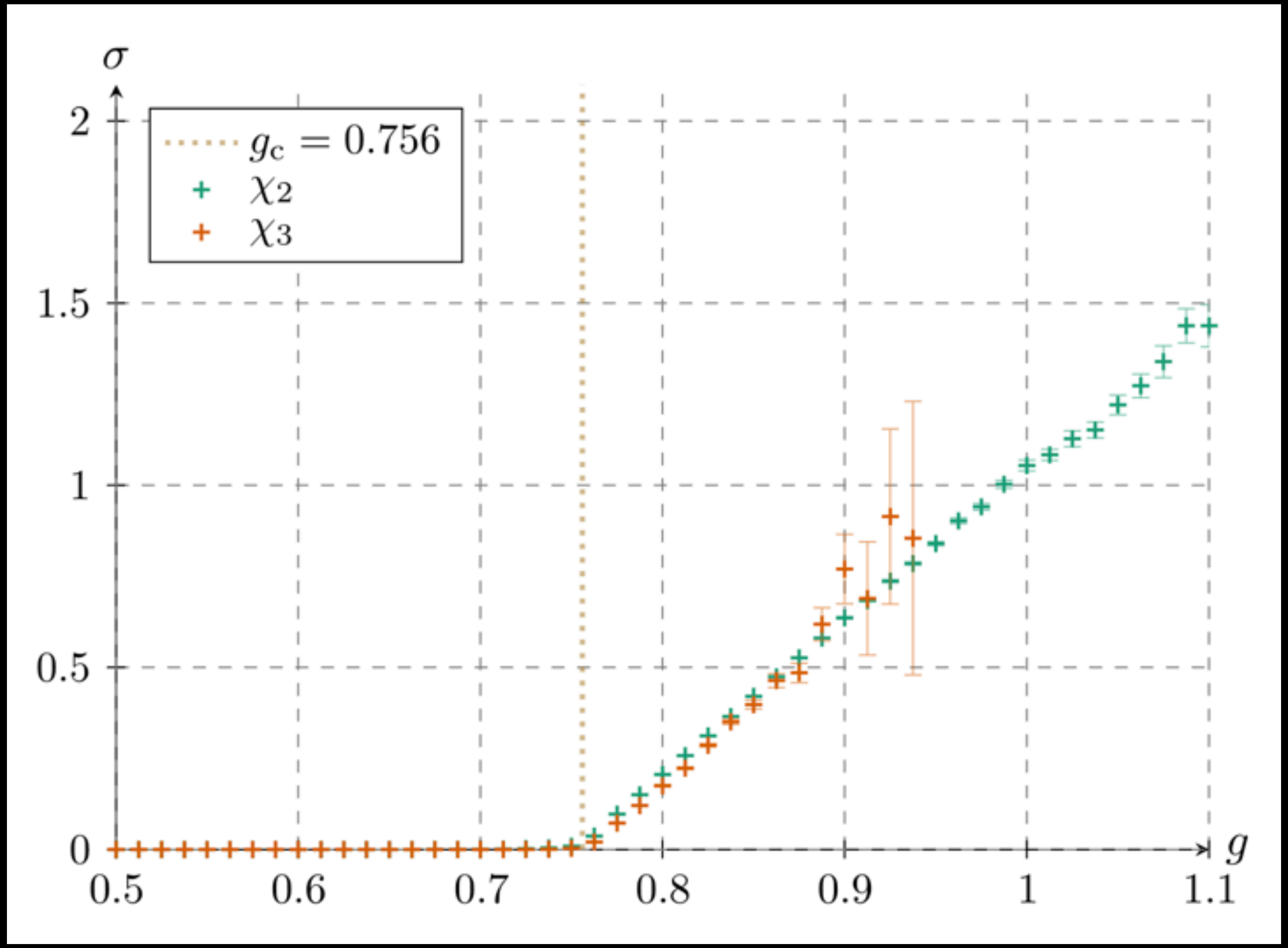
Neural network only



Existing works

Z_2 pure gauge theory

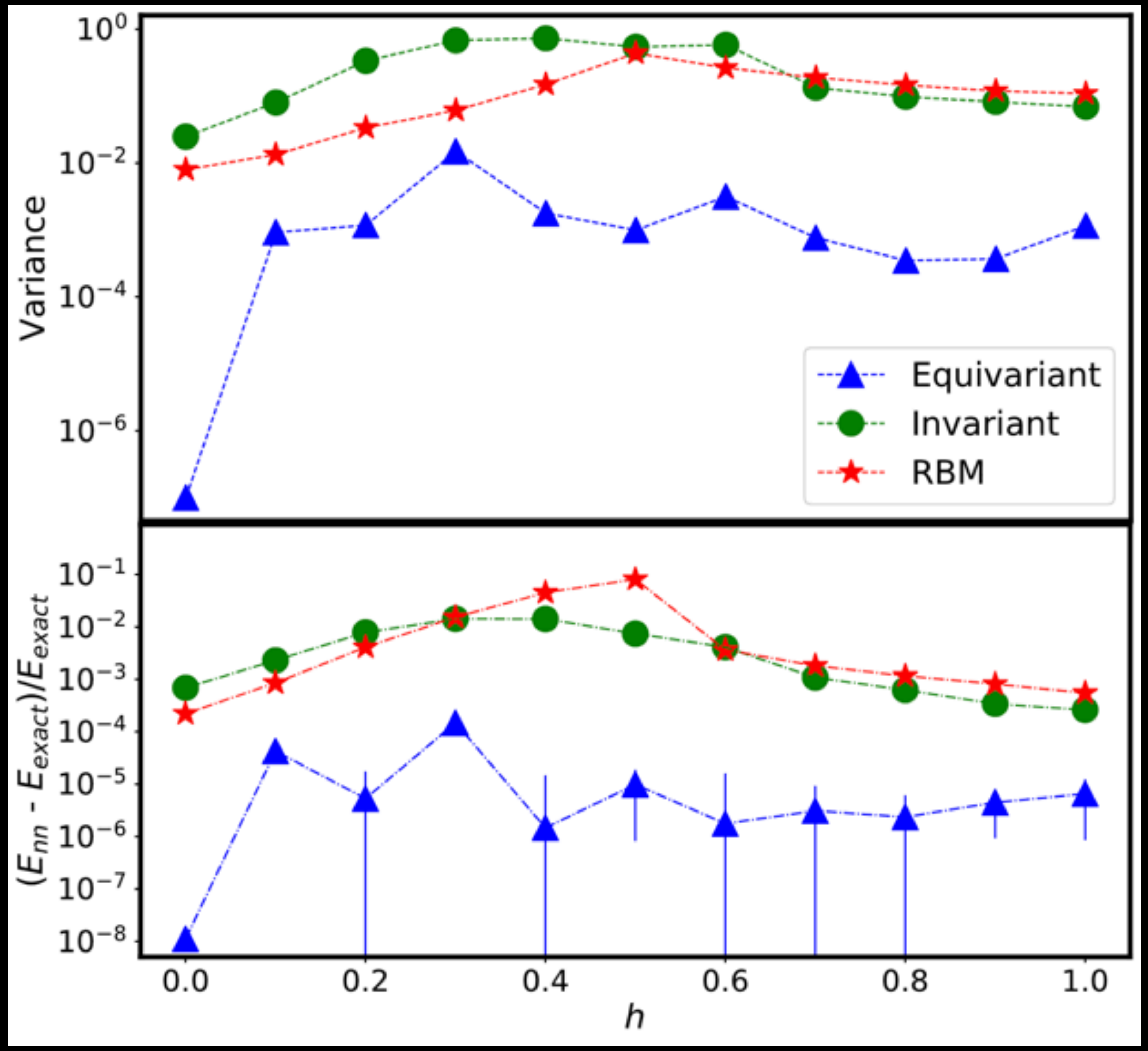
- Using the LCNN architecture
- Detect phase transition
    - Perimeter -> Area law scaling of Wilson loop

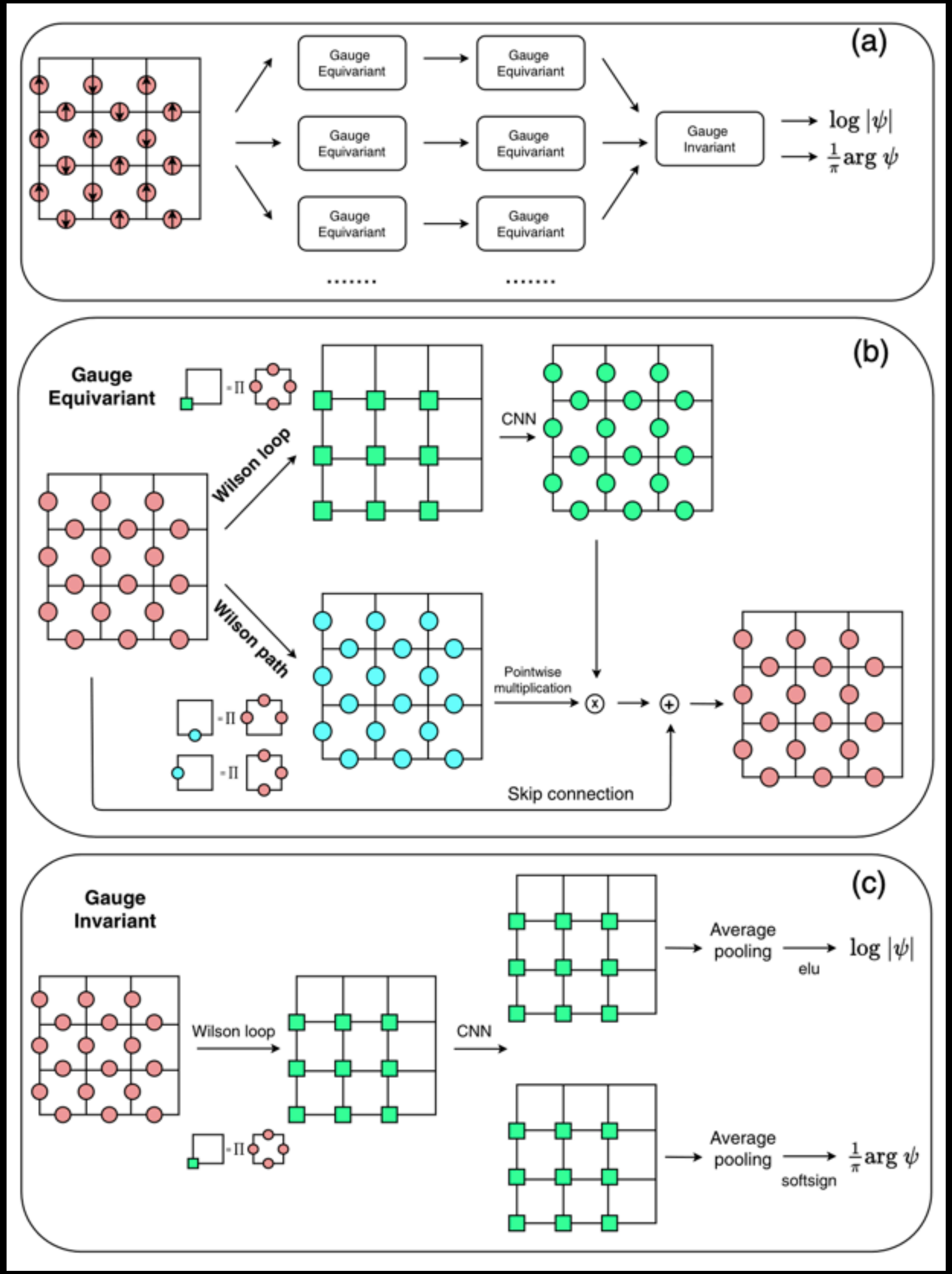


Apte, A., Córdova, C., Huang, T. C., & Ashmore, A. (2024). Deep learning lattice gauge theories. *Physical Review B*, *110*(16), 165133.

Z_2 pure gauge theory

- Using a custom ansatz
- Energy comparable to exact solution

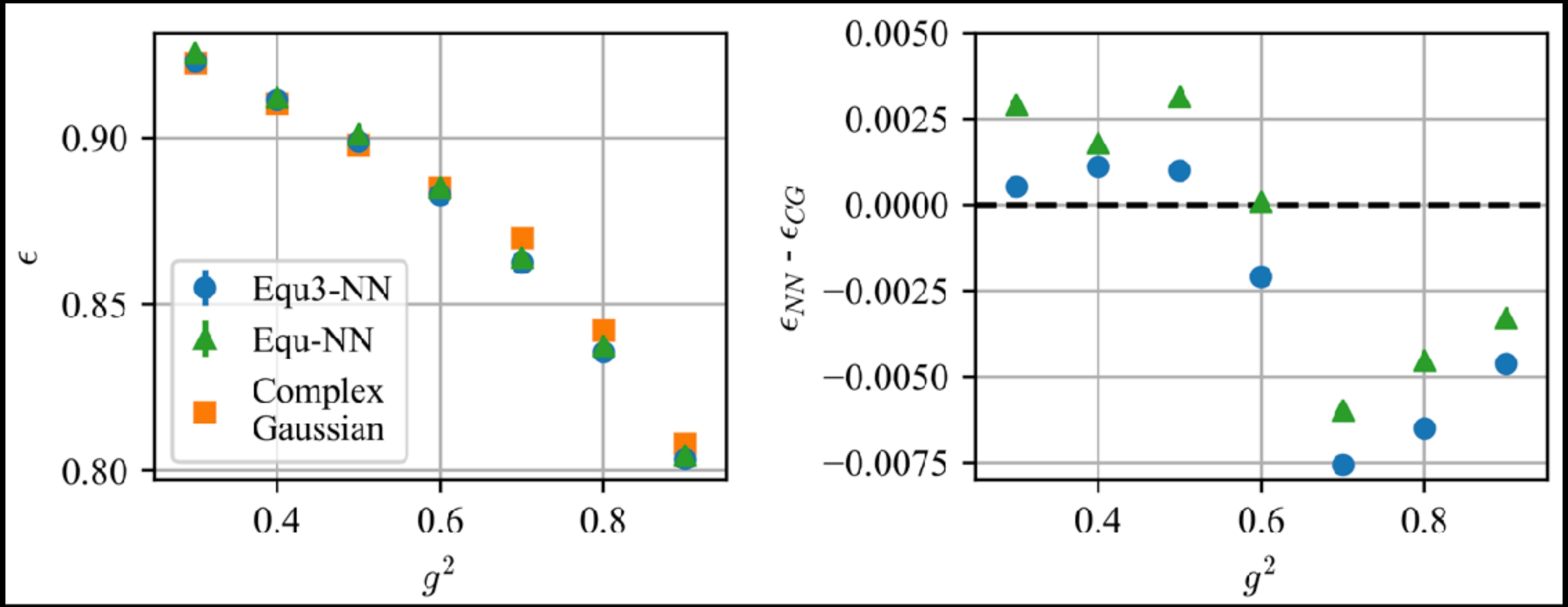


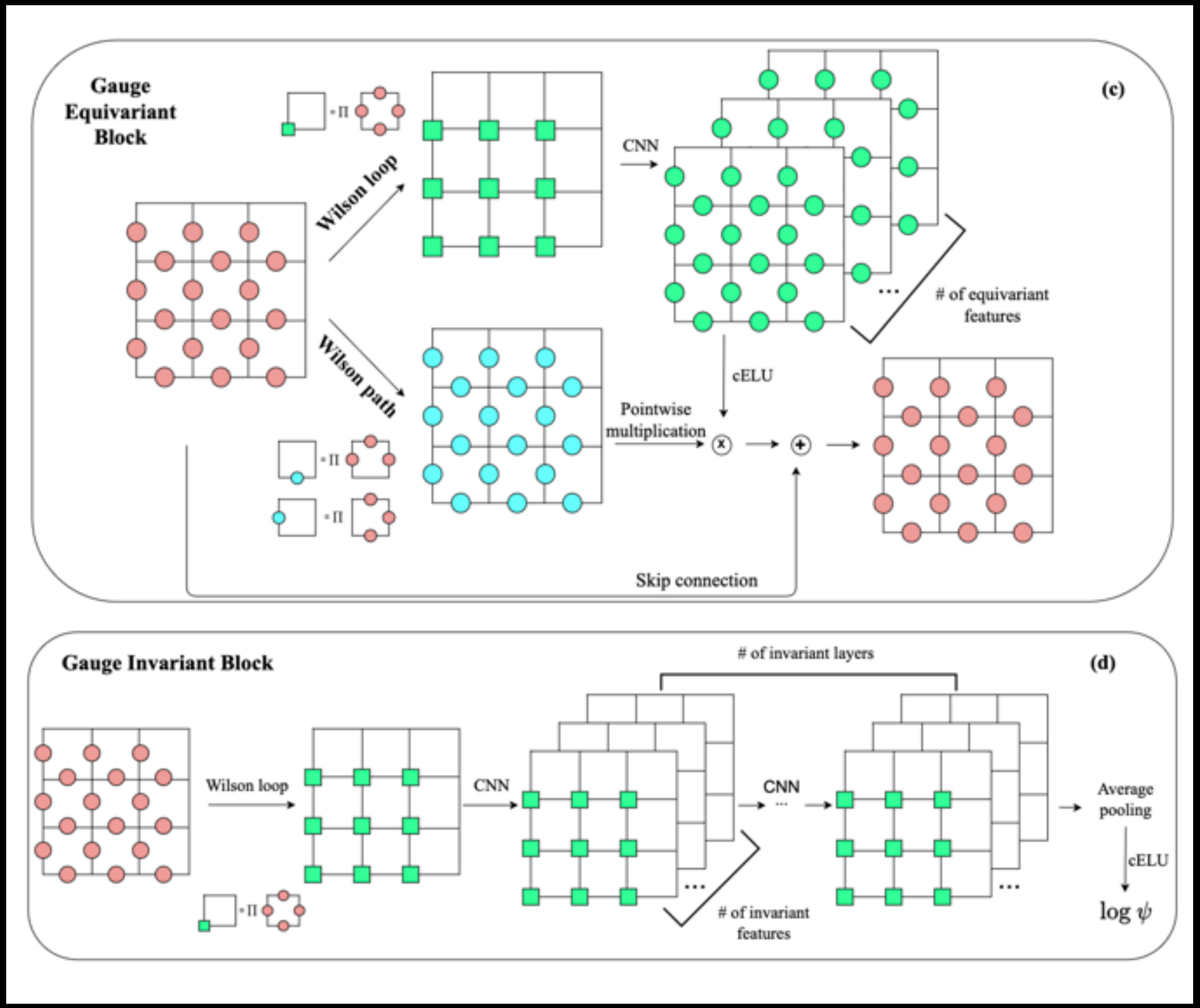


Luo, D., Carleo, G., Clark, B. K., & Stokes, J. (2021). Gauge equivariant neural networks for quantum lattice gauge theories. *Physical review letters*, *127*(27), 276402.

$U(1)$ pure gauge theory

Energy similar to physics informed ansatz

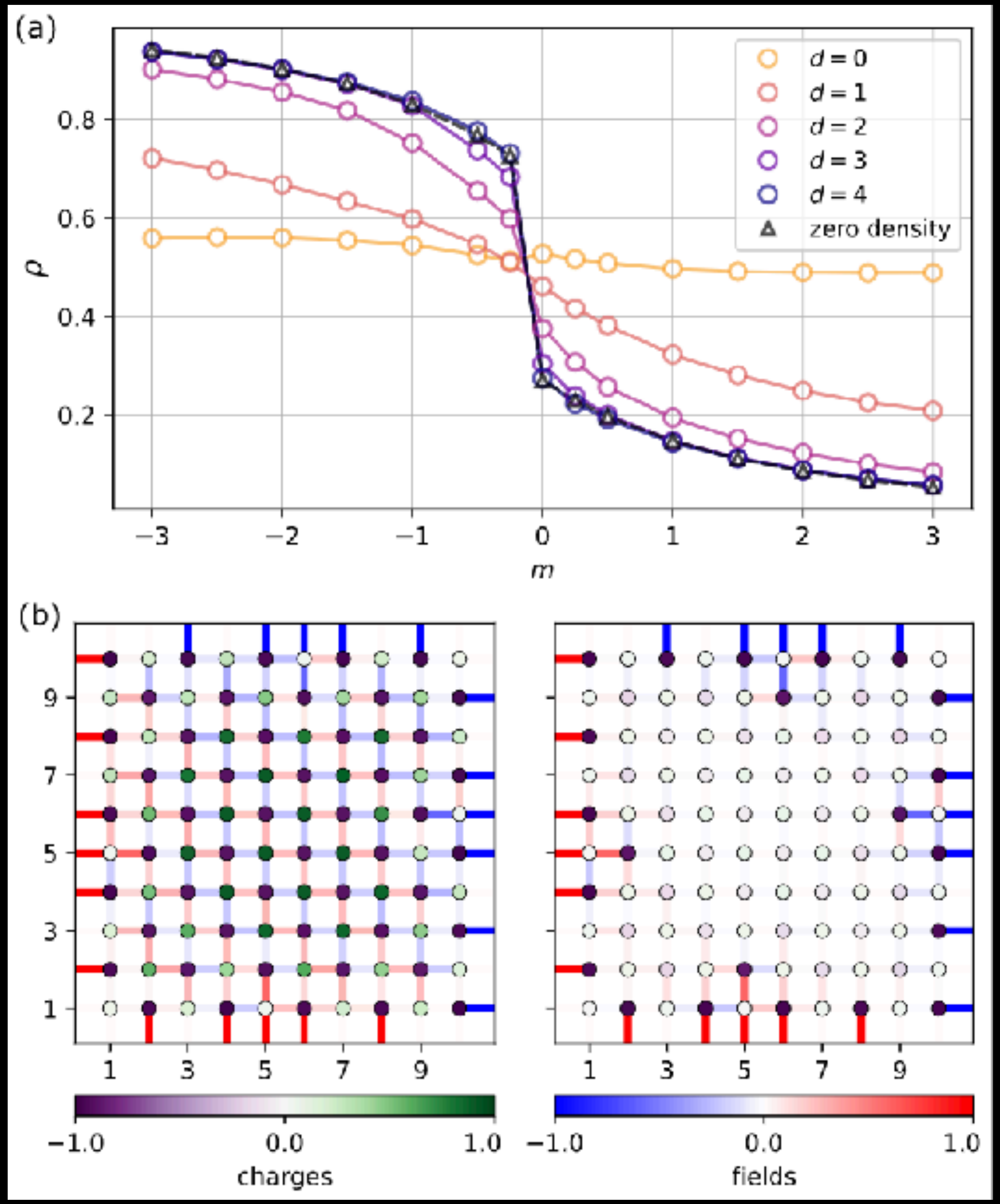




Luo, D., Yuan, S., Stokes, J., & Clark, B. K. (2022). Gauge equivariant neural networks for 2+ 1d u (1) gauge theory simulations in hamiltonian formulation. *arXiv preprint arXiv:2211.03198*.

$U(1)$ gauge theory with matter

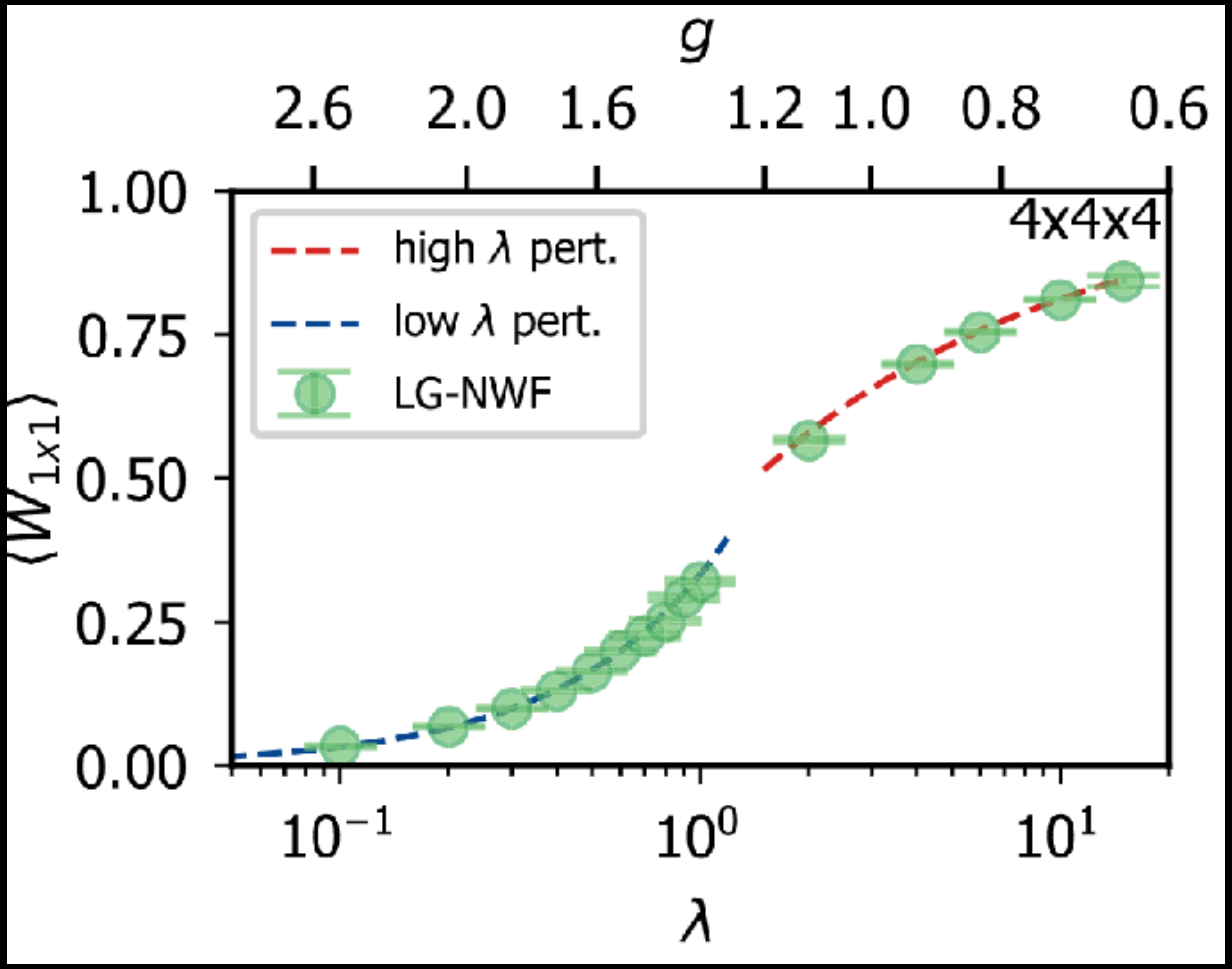
- **Autoregressive** Ansatz (no Markov chains)
- Detect/predict transitions in matter density

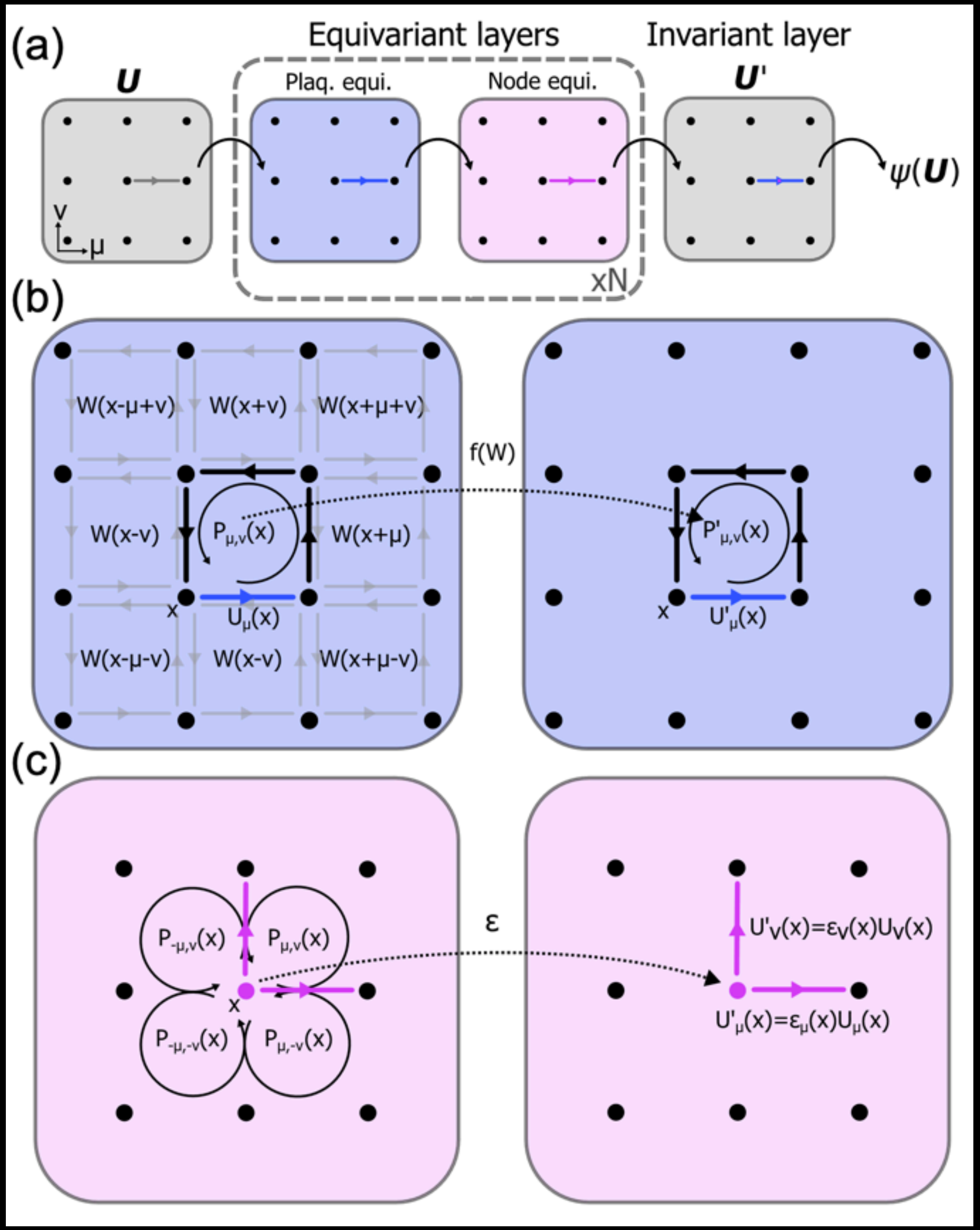


Chen, Z., Luo, D., Hu, K., & Clark, B. K. (2022). Simulating 2+ 1d lattice quantum electrodynamics at finite density with neural flow wavefunctions. *arXiv preprint arXiv:2212.06835*.

$SU(2)$ pure gauge theory

- Both LCNN and flow-like architectures
- 3D





Spriggs, T., Greplova, E., Carrasquilla, J., & Nys, J. (2026). Accurate Ground States of SU (2) Lattice Gauge Theory in 2+ 1 D and 3+ 1 D. *Physical review letters*, *136*(10), 101902.

Getting involved

NetKet:
https://www.netket.org/

- Great tutorials
- Very modular code
  - Mixed learning curve
  - Lots of code/feature reuse
- Regular updates & new features
- https://gitlab.com/QMAI/papers/su2_neural_wavefunctions

At Lattice 2026

Wednesday morning

Vacuum structure and confinement

Ground States of adjoint SU(2) from Neural Newtork assisted Variational Montecarlo

Vacuum structure a... | Vacuum structure and ...

Jul 29, 2026, 11:10 AM
20m
Thurgood Marshall (Adele H. Stamp Student Union)

Speaker

David Ward (National Taiwan University)

Thursday afternoon

Algorithms and machine learning

Diagonalizing the Kogut-Susskind Hamiltonian with Physics-Informed Neural Networks

Algorithms and artifi... | Algorithms and artifici...

Jul 30, 2026, 4:10 PM
20m
Benjamin Banneker B (Adele H. Stamp Student Union)

Speaker

Simone Romiti

Neural Wavefunctions in Quantum Field Theory

Contributed talk | Algorithms and artifi... | Algorithms and artifici...

Jul 30, 2026, 4:30 PM
20m
Benjamin Banneker B (Adele H. Stamp Student Union)

Speaker

Suryansh Rajawat (University of Maryland)

3+1d SU(3) Yang-Mills + theta-term with Neural Network Quantum States

Contributed talk | Algorithms and artifi... | Algorithms and artifici...

Jul 30, 2026, 4:50 PM
20m
Benjamin Banneker B (Adele H. Stamp Student Union)

Speaker

Joshua Lin (Argonne National Laboratory)

Eliska Greplova

TU Delft

ETH zürich

Juan Carrasquilla

Jannes Nys

Thank you for listening

Backup slides

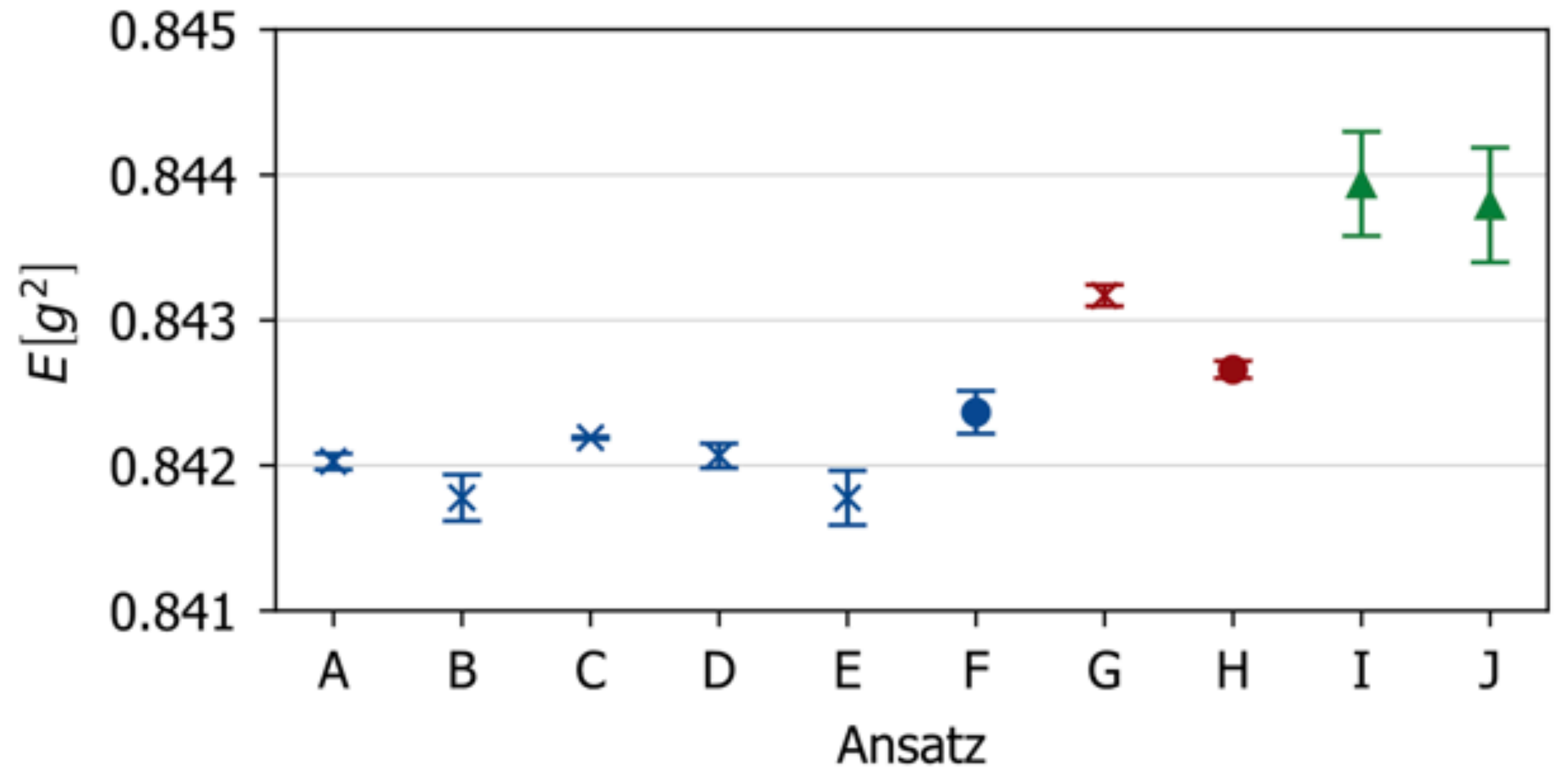


Figure S2. Variational ground state energy for a range of ansatze outlined in Table S1 for $\lambda = 1$. Errors are the Monte Carlo error of the energy over the final set of samples.

Figure S2 label	**Plaquette layer**			**Node layer**	**Num. equivariant blocks**	**Final layer**
	Architecture	**Num. layers**	**Other**			
A	CNN	1	K=3	Yes	1	Jastrow
B	CNN	1	K=1	Yes	1	Jastrow
C	CNN	2	K=3	Yes	1	Jastrow
D	CNN	1	K=3	Yes	2	Jastrow
E	CNN	1	K=3	Yes	1	FCN
F	ViT	1	NH=2, D=4	Yes	1	Jastrow
G	CNN	1	K=3	No	1	Jastrow
H	ViT	1	NH=2, D=4	No	1	Jastrow
I	None	N/A	N/A	No	0	Jastrow
J	None	N/A	N/A	No	0	FCN

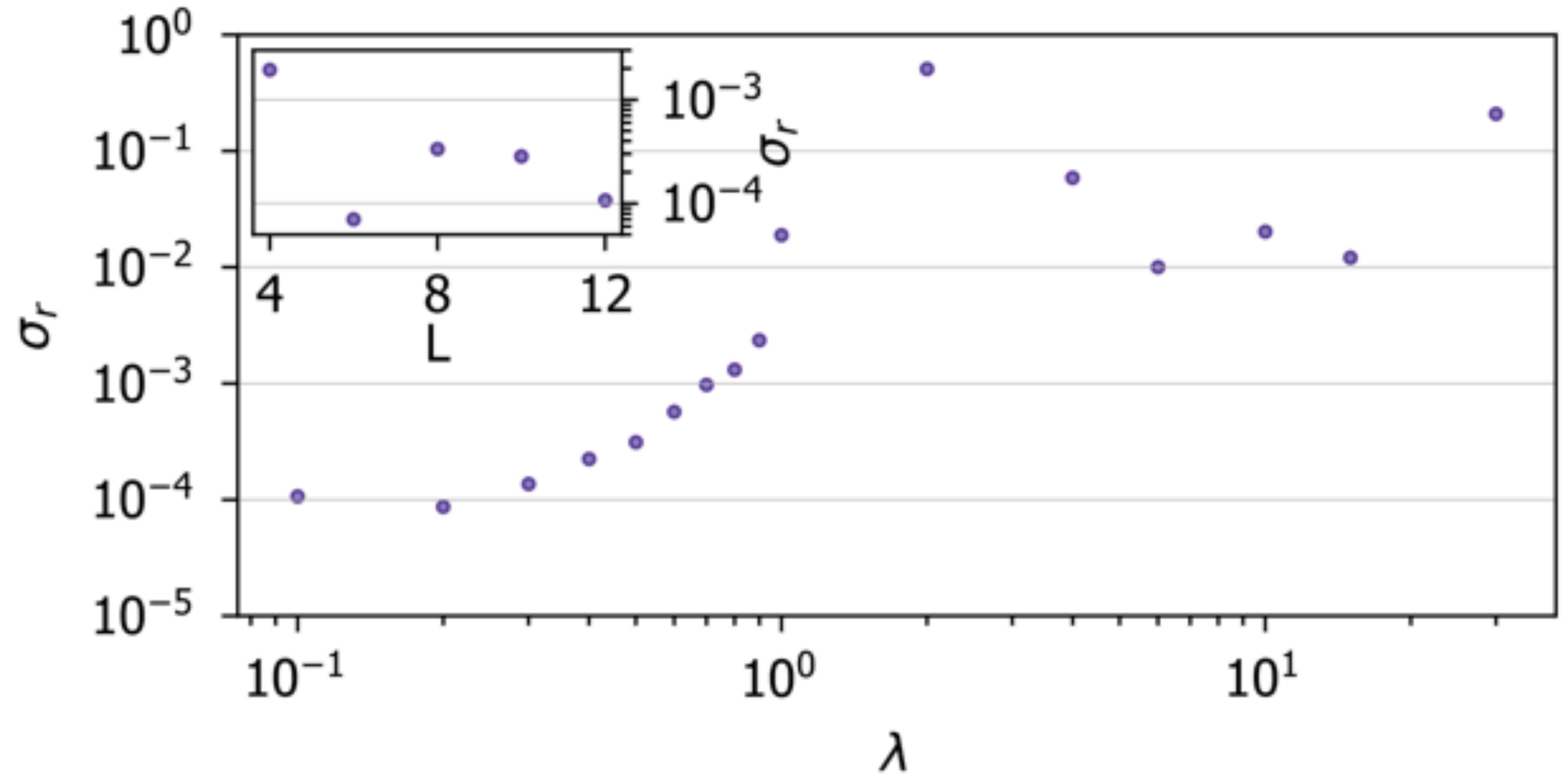


Figure S4. The rescaled variance of the energy, as in Eq. (D22), averaged over the final 20 iterations of VMC. The main panel shows results for 12×12 lattices at various values of λ; the inset shows the rescaled variance but against system size $(L \times L)$ with fixed $\lambda = 0.1$.

the version trained under the alternative Hamiltonian and then fine-tuned and they agree within errors.

During the VMC procedure we also compute the energy variance σ, defined as

$$\sigma = \frac{\langle\Psi|H^2|\Psi\rangle}{\langle\Psi|\Psi\rangle} - \left(\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right)^2 . \tag{D21}$$

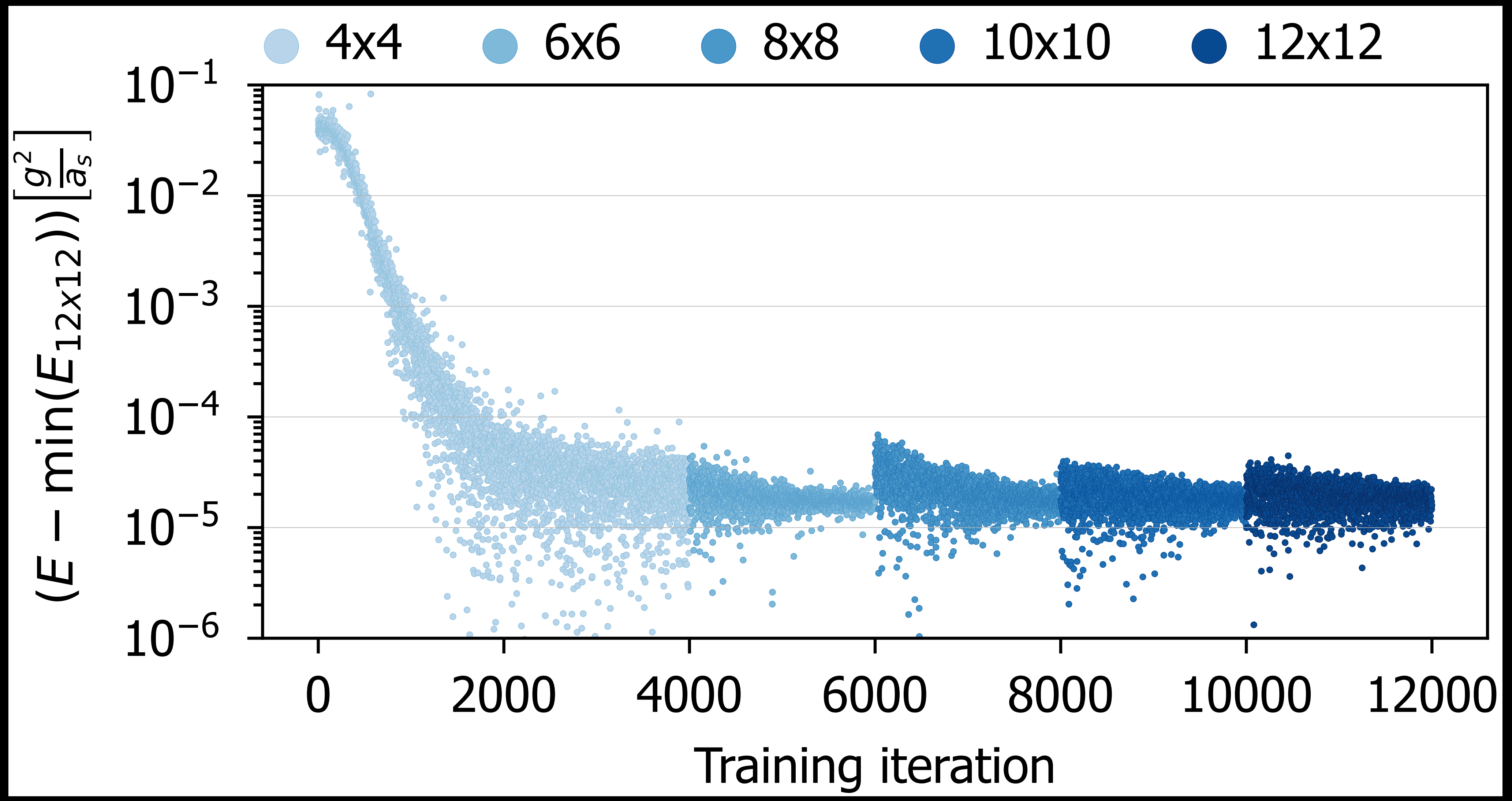
Different gauge groups

Do-able

- $U = U(\theta, \phi, \psi)$ -> 8 angles
- ∇^2 -> More complicated
    - (possibly done but not explained in Chin, S. A., *et al.* Phys. Rev. D **37**, 3001)
- Note: canonical ordering of eigenvalues non-trivial in SU(3)
    - When do go $P_{\mu,\nu}(x) \to \lambda_P \to \lambda'_P \to P'_{\mu,\nu}(x)$ you have to order the eigenvalues before and after so they correspond to the same eigenvectors

Larger lattices

Trivial



- We did transfer learning in this work already
- Equivariant layers are local
  - Convolution layers, fixed kernel size
- Final layer is non-local
  - One new parameter for each new unique distance on the lattice
    - Can be initialised as ‘small’
- Can also be done in arbitrary dimension, 2 and 3 shown in paper

Fermions and time evolution

Hard, open study

- Both done in spin systems with neural wavefunctions
  - Even at the same time:
    - Nys, J., Pescia, G., Sinibaldi, A. and Carleo, G., 2024. Ab-initio variational wave functions for the time-dependent many-electron Schrödinger equation. *Nature communications*, *15*(1), p.9404.
- Gauge invariance and representation for LGTs with fermions remain open questions
- Time evolution is more straightforward (only variational challenges, nothing conceptual)

Other applications

Done, but not for LGTs

- Physics distilled into neural network parameters, θ
    - Fewer parameters than basis states
- Infer phase transitions
    - Hernandes, V., **Spriggs, T.**, Khaleefah, S., Greplova, E. ICLR Workshop on Neural Network Weights as a New Data Modality (2025). *quant-ph arXiv:2503.17140.*
- Excited states and phase transitions
    - Medvidović, M., Orfi, A., Carrasquilla, J., Sels, D. arXiv:2510.15030
- Compute non-stabiliserness
    - **Spriggs, T**., Ahmadi, A., Chen, B. and Greplova, E., 2025. *Machine Learning: Science and Technology*, *6*(1), p.015042

Time evolution

- Time evolution of the quantum state -> time evolution of parameters
- $|\Psi(t)\rangle = |\Phi_{\theta(t)}\rangle$
- Many approaches
  - TVMC, pTVMC, TreTVMC
- TVMC equates a small Totter evolution of the true state to a Taylor expansion of the variational state:
  - $|\Psi(t+\delta t)\rangle = |\Phi_{\theta(t+\delta t)}\rangle \quad \rightarrow \quad \Psi(t)(1 - iH\delta t) = \Phi_{\theta(t)}(1 + \dot{\theta} t)$
  - Minimise Fubini-Study distance between these two expansions, solve for $\dot{\theta}$

Detailed description in SM of Schmitt, M. & Heyl, M. Phys. Rev. Lett. **125**, 100503 https://journals.aps.org/prl/supplemental/10.1103/PhysRevLett.125.100503/supplement.pdf