Neural wavefunctions for SU(2) lattice gauge theory

Thomas Spriggs
TU Delft
Work in collaboration with: Eliska Greplova, Juan Carrasquilla, and Jannes Nys

Hamiltonian lattice gauge theories

In a nutshell

- New method for simulation non-Abelian lattice gauge theories in any* dimension
- I will show: ground state properties of SU(2) pure gauge
- Can be extended to: SU(N), time evolution, fermions

*GPU RAM dependent

In a nutshell

- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$

In a nutshell

- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$

Samples Action

In a nutshell

- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$

Samples Action

- In variational Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim \psi}$

In a nutshell

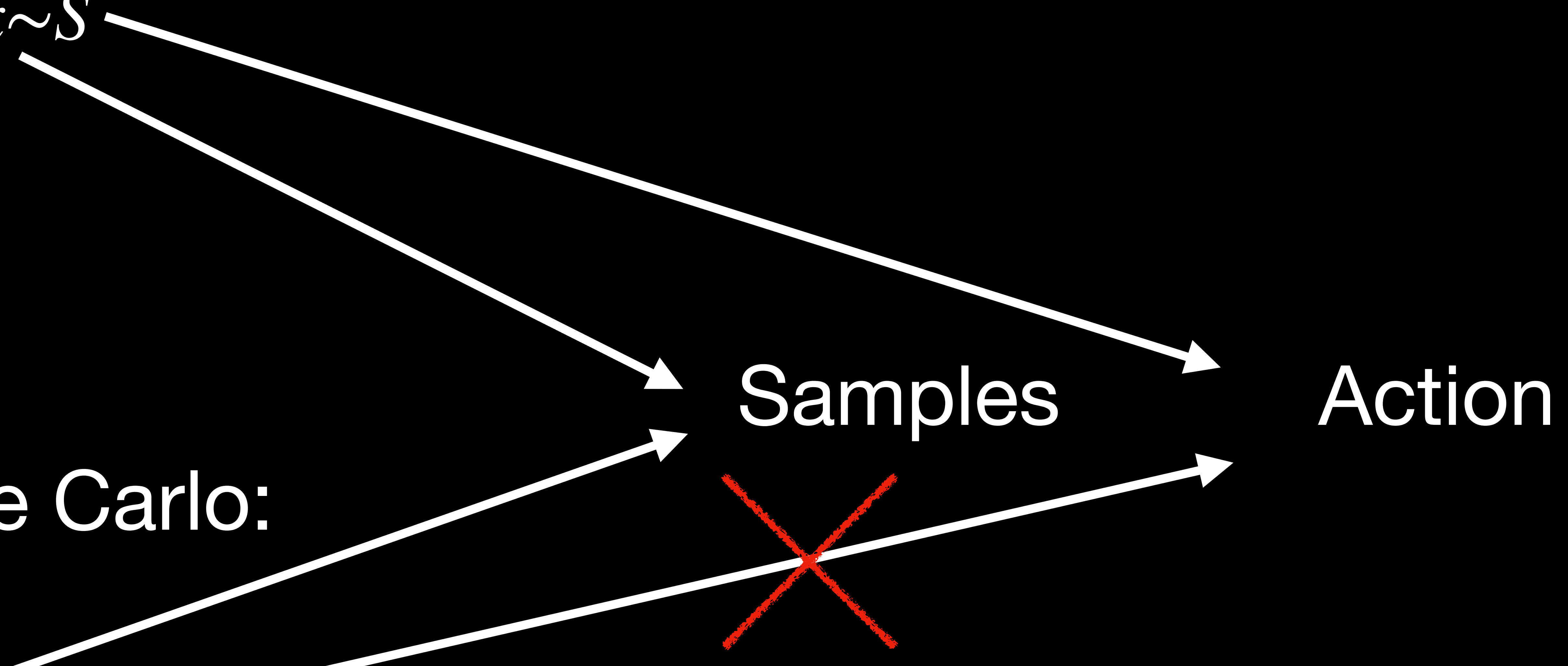
- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$

Samples Action

- In variational Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim \psi}$

In a nutshell

- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$



- In variational Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim \psi}$

In a nutshell

- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$
- In variational Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim \psi}$

Samples

Action

Learned distribution

In a nutshell

- In Euclidean (quantum) Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim S}$

Samples   Action

- In variational Monte Carlo:
  - $\langle O \rangle = \mathbb{E}[O(x)]_{x \sim \psi}$

Neural network

In a nutshell: why?

- Use the Hamiltonian framework (continuous time)
- Free of the sign problem (probably*)
- Maintains continuous representation
- Real time evolution of lattice gauge theories (future direction)

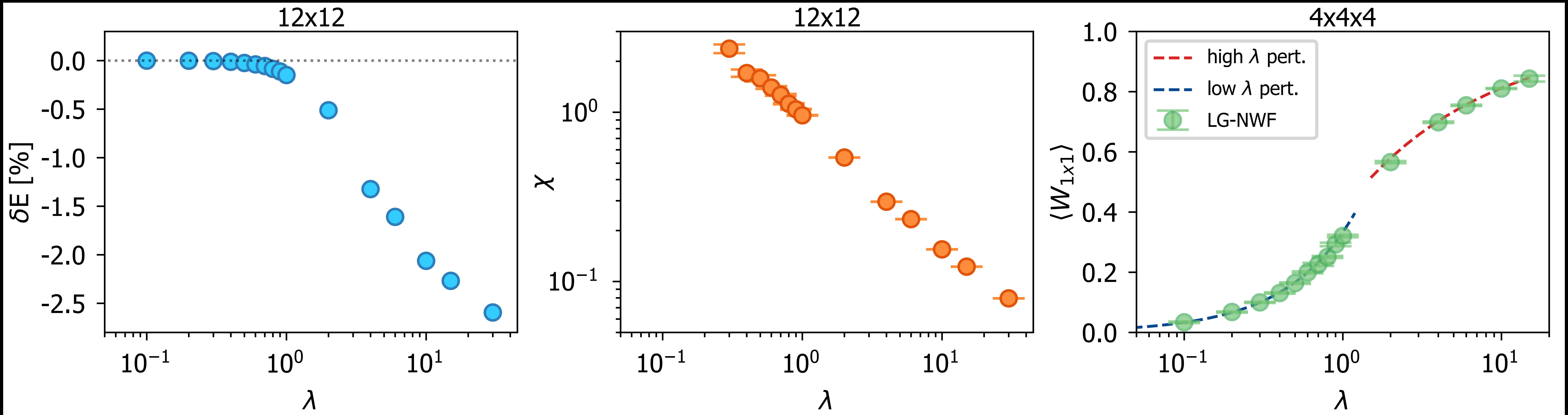
* hard to know exactly if the distribution can be learned in all cases without something like the sign problem appearing

In a nutshell: why not?

- Hard to design the distribution to give you the right observables
- Sampling from this distribution still remains difficult

g range ~ 2.5 - 0.6

Large λ is weak coupling (small g)



Energy improvement over a reference ansatz (low is good)

Creutz ratio

Plaquette

Related neural wavefunction work

- U(1) (Abelian)

  Luo, D. *et al.* arXiv:2211.03198

- $\mathbf{Z}_2$ (discrete and Abelian)

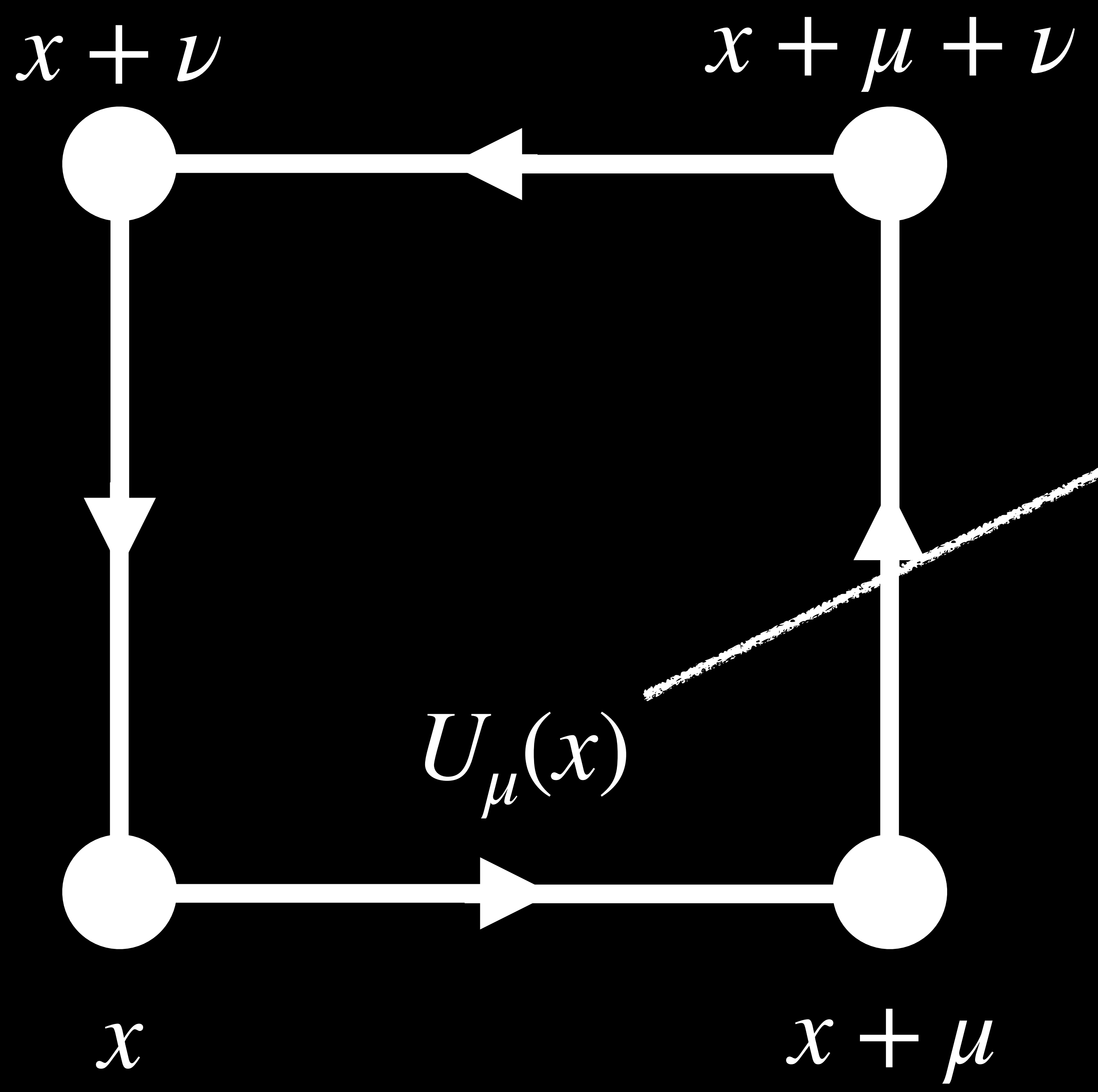
  Apte, A. *et al.* Phys. Rev. B **110**, 165133

- Non-neural-network:
  - SU(2) Chin, S. A. *et al.* (1985) *Phys Rev D* **31**, 3201
  - SU(3) Chin, S. A. *et al.* (1988) *Phys Rev D* **37**, 3001

SU(2) lattice gauge theory

SU(2) Hamiltonian

Basic degrees of freedom

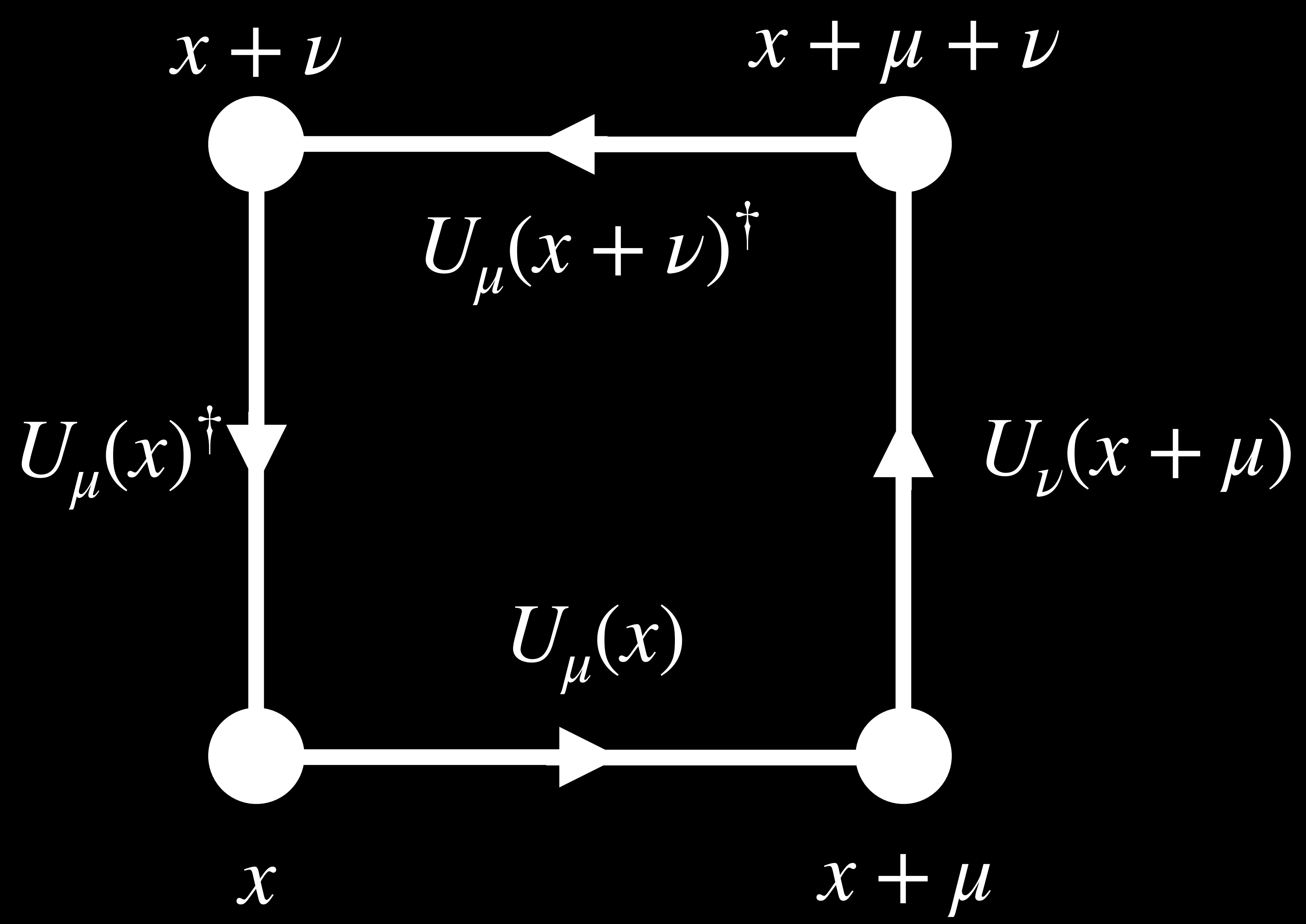


$U_\mu(x) \in SU(2)$

e.g.

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

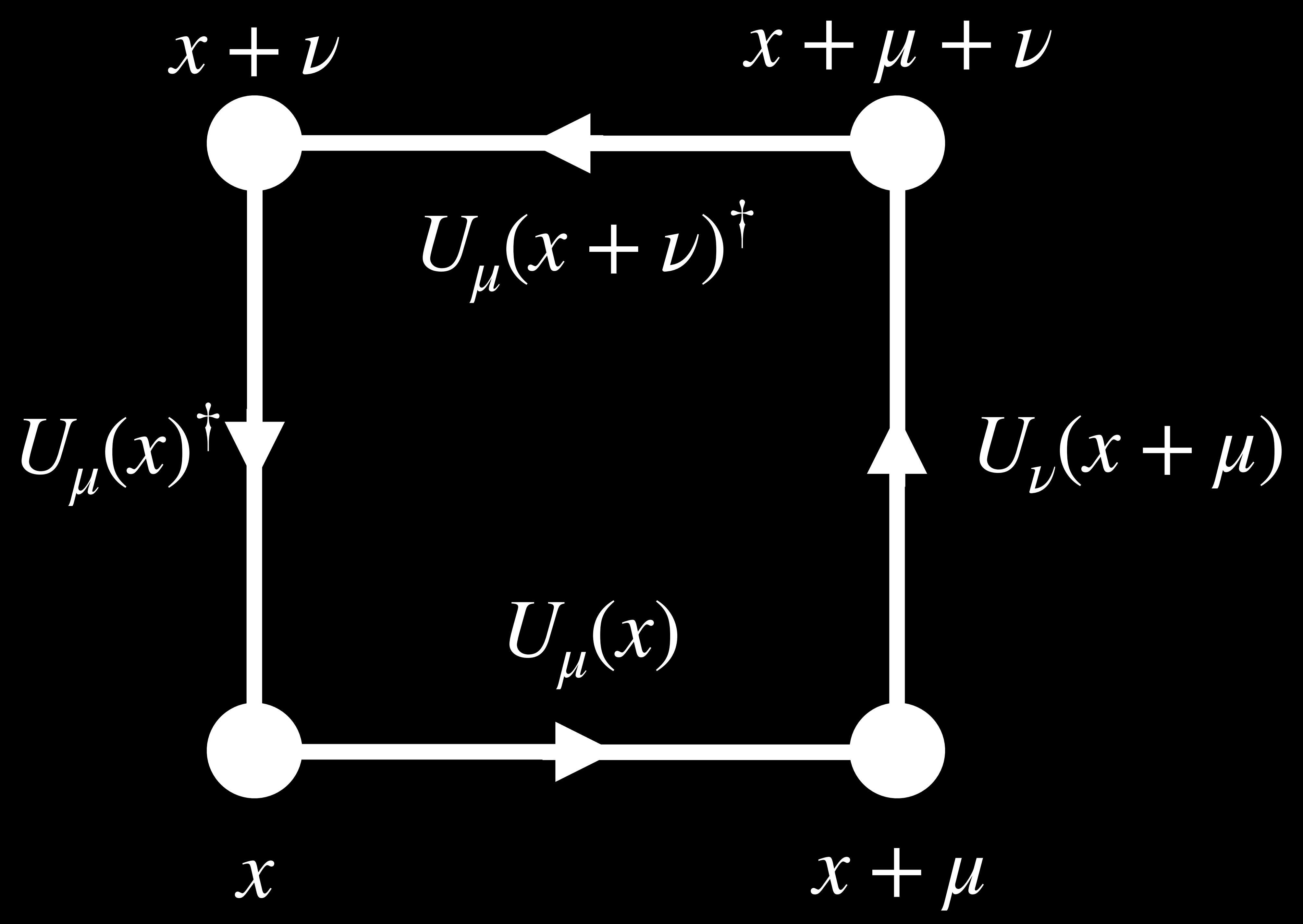
SU(2) Hamiltonian



Plaquette:

non-Abelian

SU(2) Hamiltonian

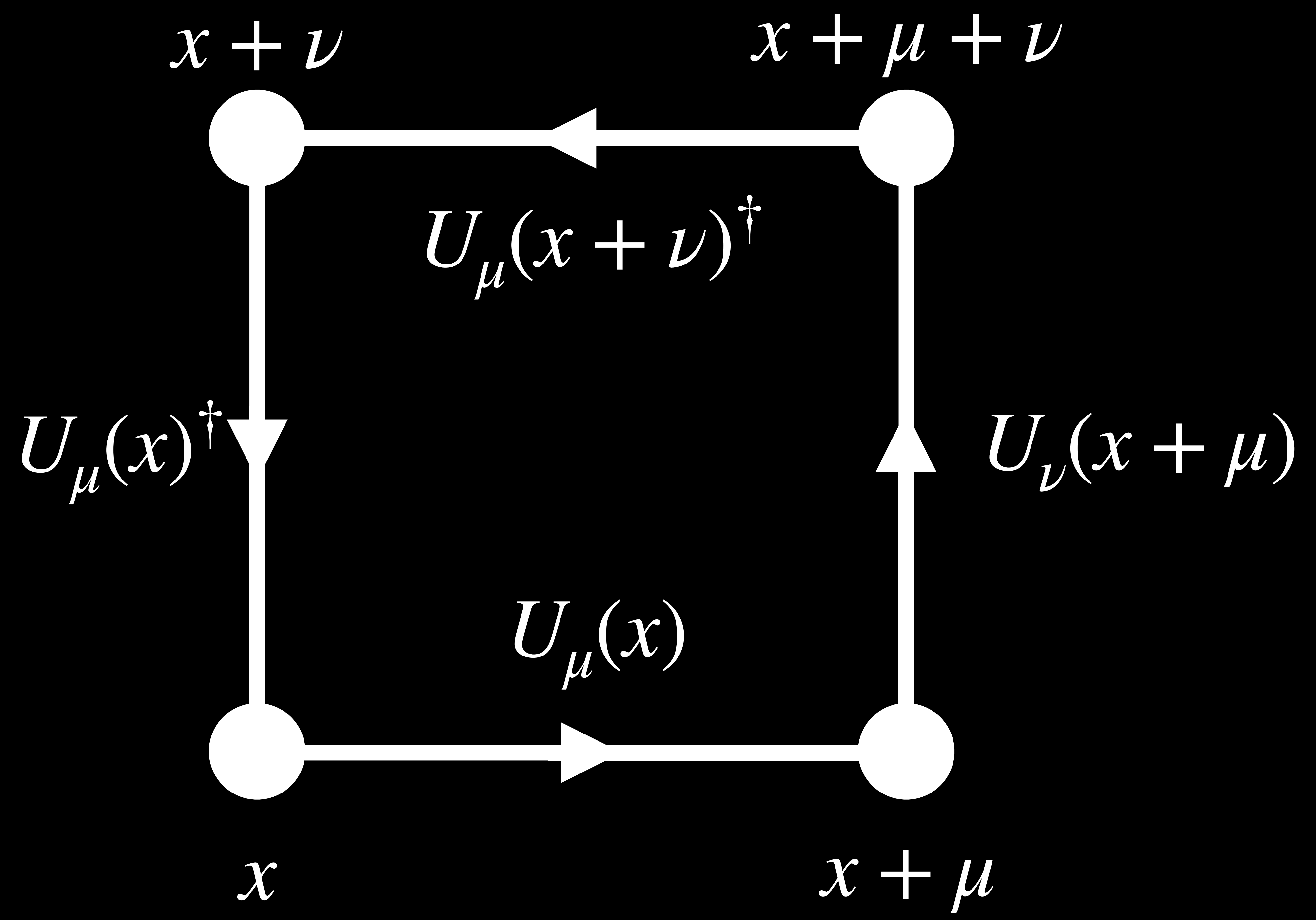


Plaquette:

$$P_{\mu,\nu}(x) = U_\mu(x)U_\nu(x+\mu)U_\mu(x+\nu)^\dagger U_\nu(x)^\dagger$$

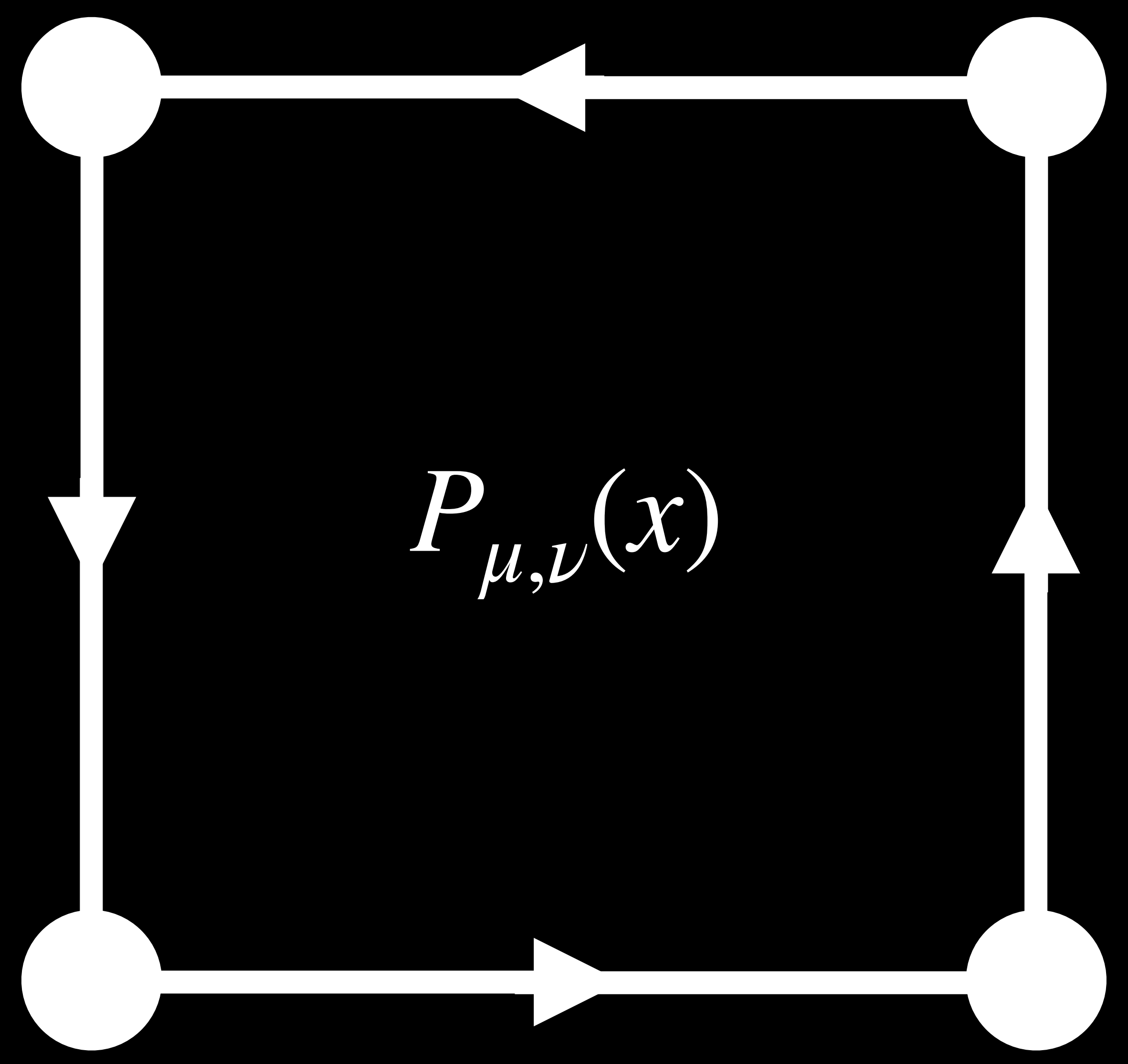
non-Abelian

SU(2) Hamiltonian



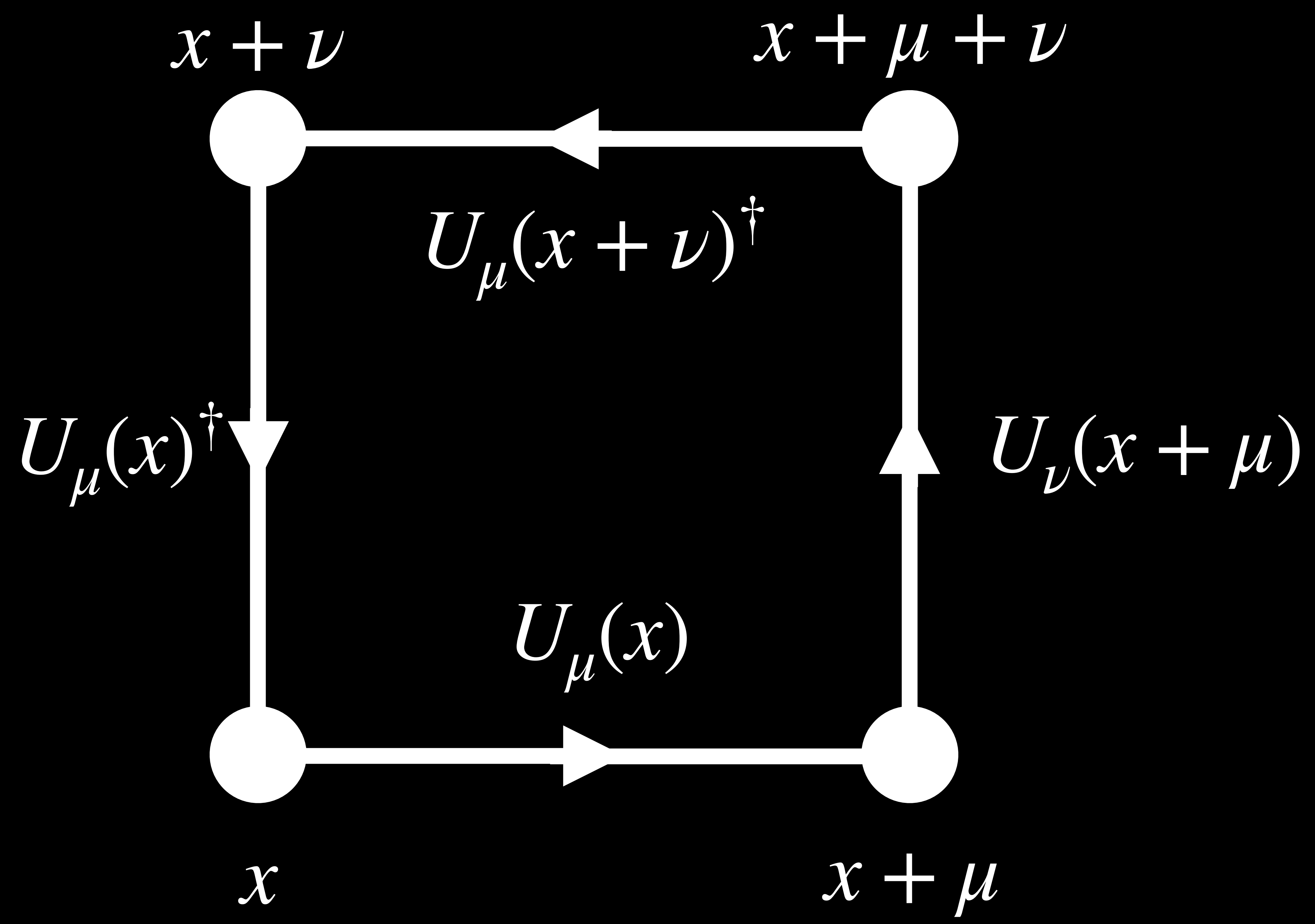
Plaquette:

$$P_{\mu,\nu}(x) = U_\mu(x)U_\nu(x+\mu)U_\mu(x+\nu)^\dagger U_\nu(x)^\dagger$$



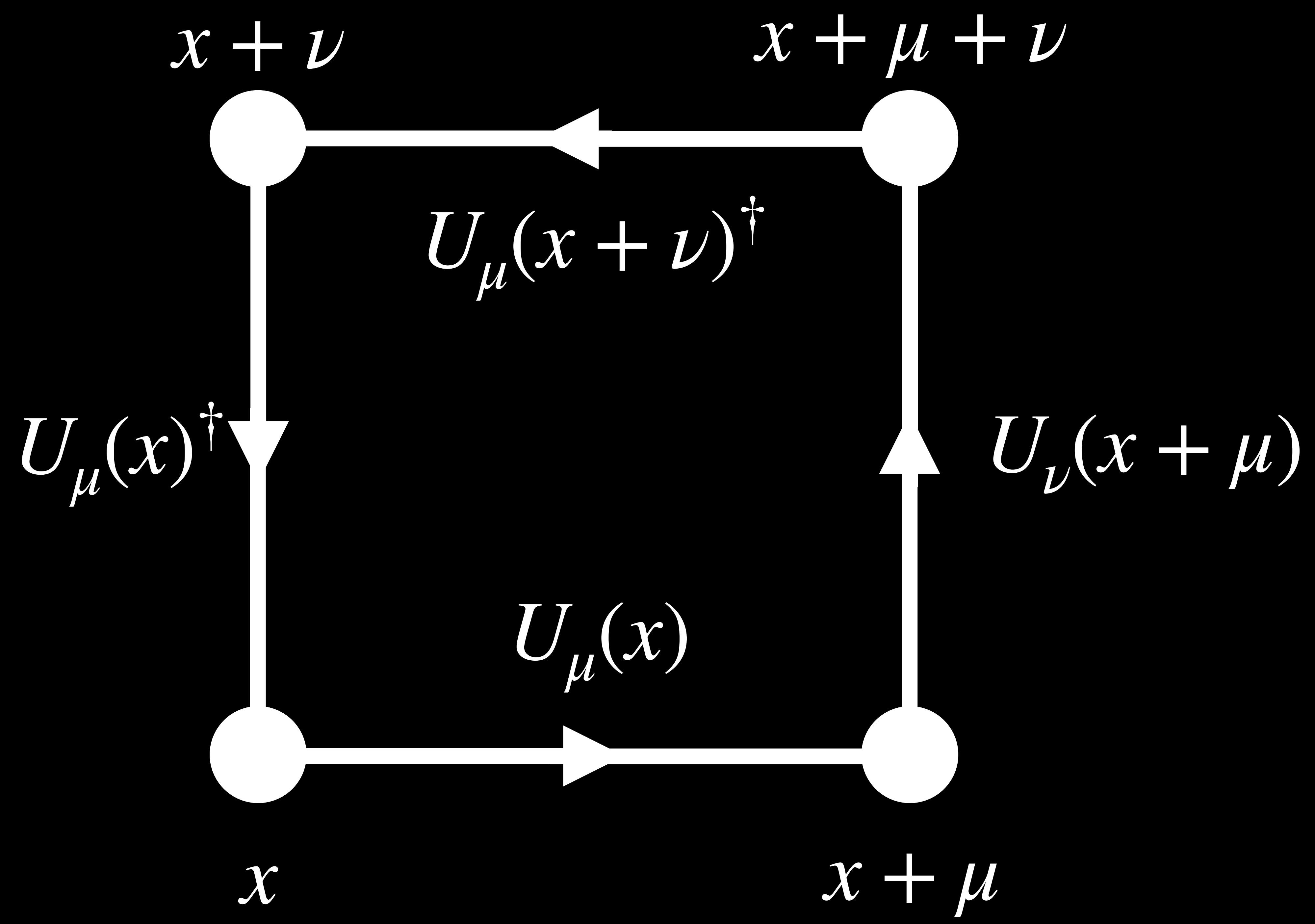
non-Abelian

SU(2) Hamiltonian



Derivative:

SU(2) Hamiltonian

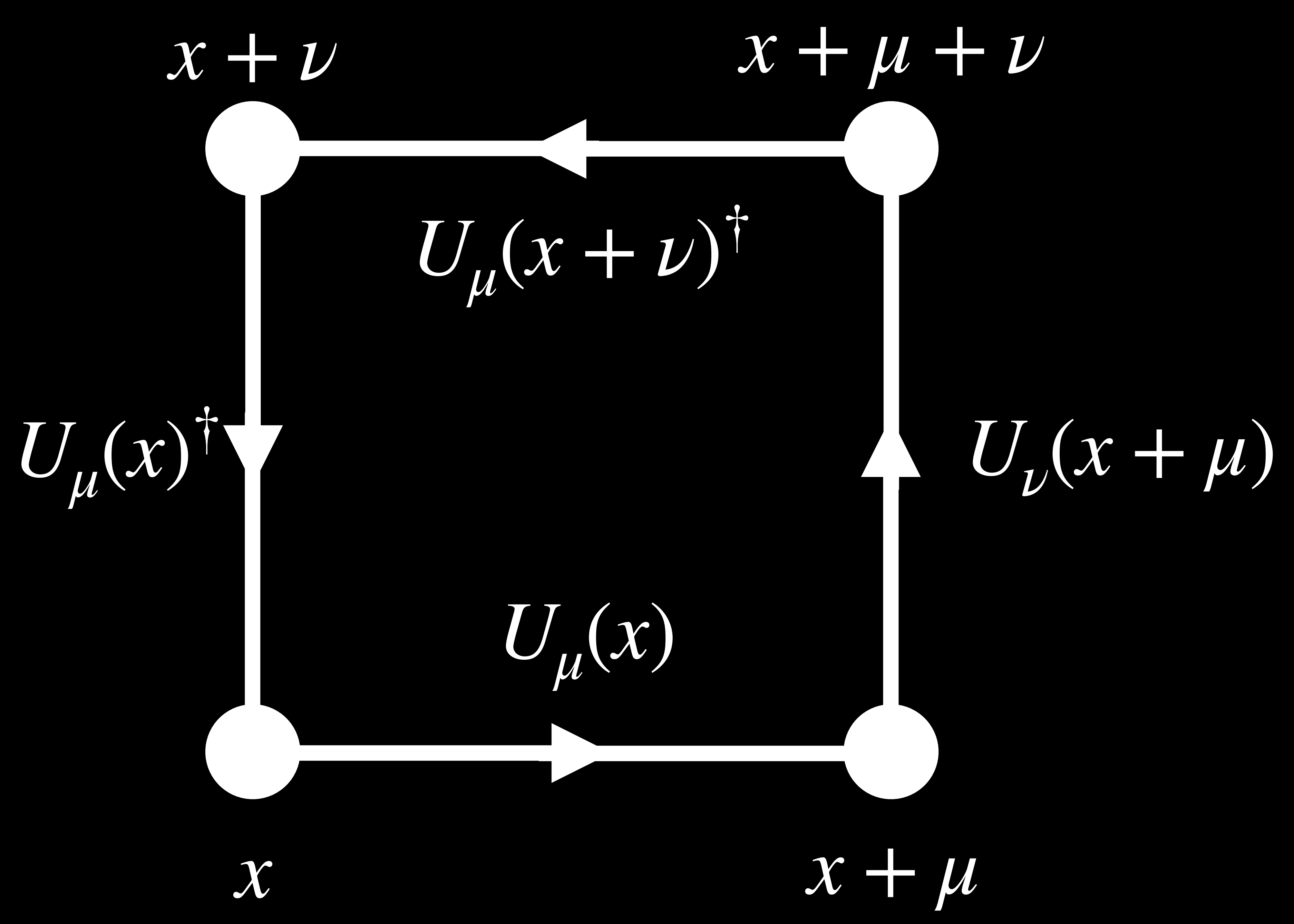


Derivative:

$$U_\mu(x) = \exp(-i\sigma_a A_\mu^a(x))$$

$$\nabla \sim \frac{\partial}{\partial A_\mu}$$

SU(2) Hamiltonian



Hamiltonian:

$$\mathscr{H} = -\frac{1}{2}\sum_l \nabla_l^2 + \lambda \sum_p \left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

kinetic $+\lambda$ potential

electric $+\lambda$ magnetic

Measure ground state properties of the SU(2) LGT

$$\mathcal{H} = -\frac{1}{2}\sum_{l} \nabla_l^2 + \lambda \sum_{p} \left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

Measure ground state properties of the SU(2) LGT

$$\mathcal{H} = -\frac{1}{2}\sum_{l} \nabla_l^2 + \lambda \sum_{p} \left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

Variational Monte Carlo

Some definitions: wavefunction

Some definitions: wavefunction

- Let $|\psi\rangle = \sum_i a_i |\phi_i\rangle$

Some definitions: wavefunction

- Let $|\psi\rangle = \sum_i a_i |\phi_i\rangle$
  - 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 $|\psi\rangle = \sum_i a_i|\phi_i\rangle$
  - 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 $|\psi\rangle = \sum_i a_i |\phi_i\rangle$
  - 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

$$|\psi_\theta\rangle = \sum_i f_\theta(\phi_i)|\phi_i\rangle$$

Some definitions: variational wavefunction

$$|\psi_\theta\rangle = \sum_i f_\theta(\phi_i)\,|\phi_i\rangle$$

Some definitions: variational wavefunction

$$|\psi_\theta\rangle = \sum_i f_\theta(\phi_i)\,|\,\phi_i\rangle$$

$$\langle\phi_0\,|\,\psi_\theta\rangle = \sum_i \langle\phi_0\,|f_\theta(\phi_i)\,|\,\phi_i\rangle = f_\theta(\phi_0)$$

Some definitions: variational wavefunction

$$|\psi_\theta\rangle = \sum_i f_\theta(\phi_i)\,|\,\phi_i\rangle$$

$$\langle\phi_0\,|\,\psi_\theta\rangle = \sum_i \langle\phi_0\,|f_\theta(\phi_i)\,|\,\phi_i\rangle = f_\theta(\phi_0)$$

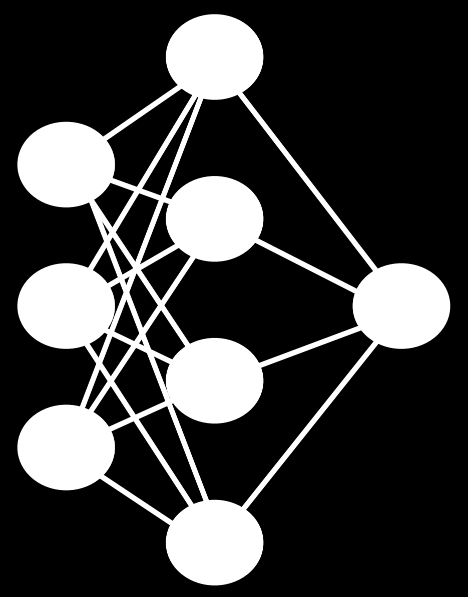
"Configuration in, amplitude out"

Some definitions: variational wavefunction

$$|\psi_\theta\rangle = \sum_i f_\theta(\phi_i)\,|\phi_i\rangle$$

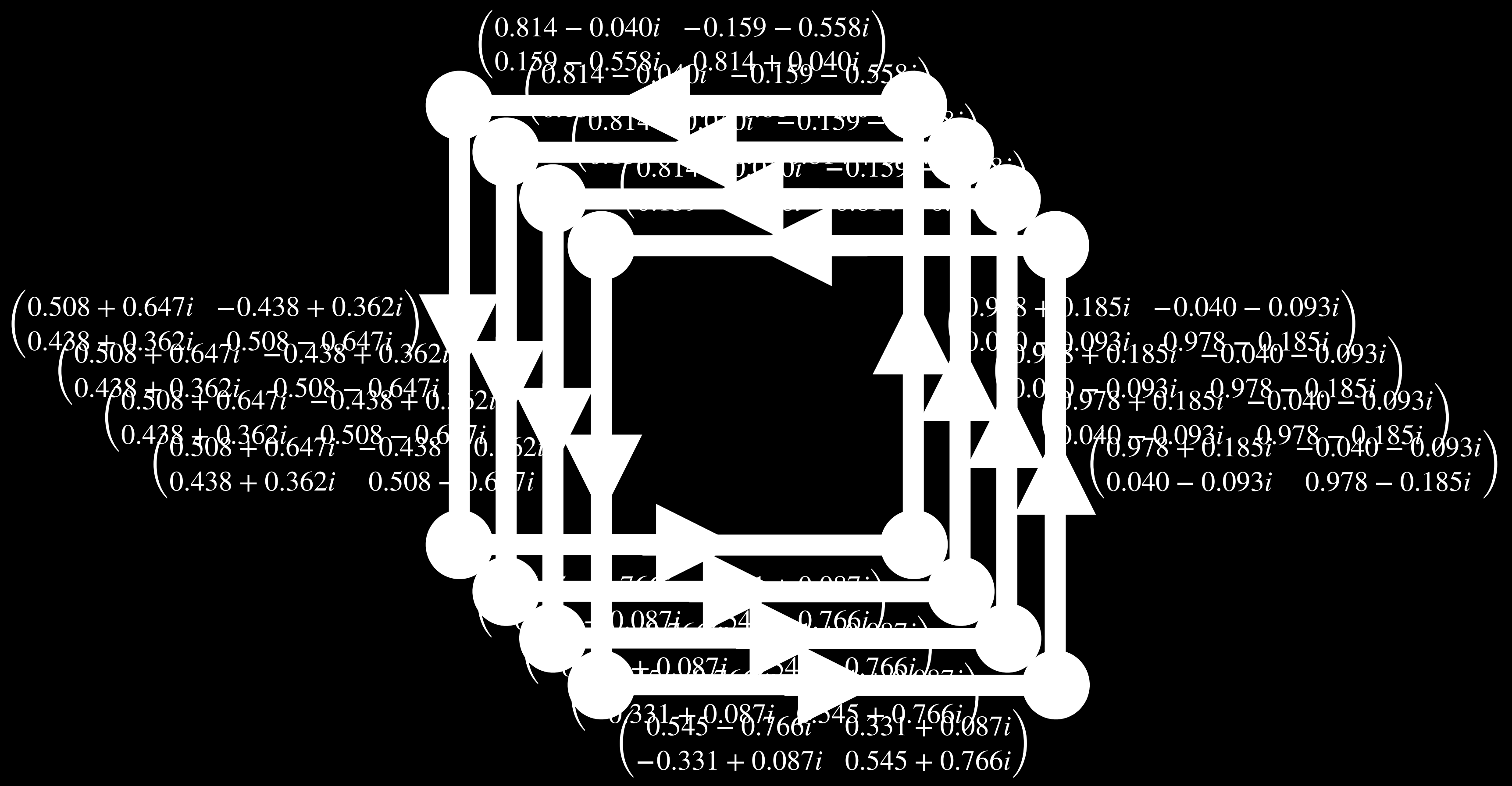
$$\langle\phi_0|\psi_\theta\rangle = \sum_i \langle\phi_0|f_\theta(\phi_i)|\phi_i\rangle = f_\theta(\phi_0)$$

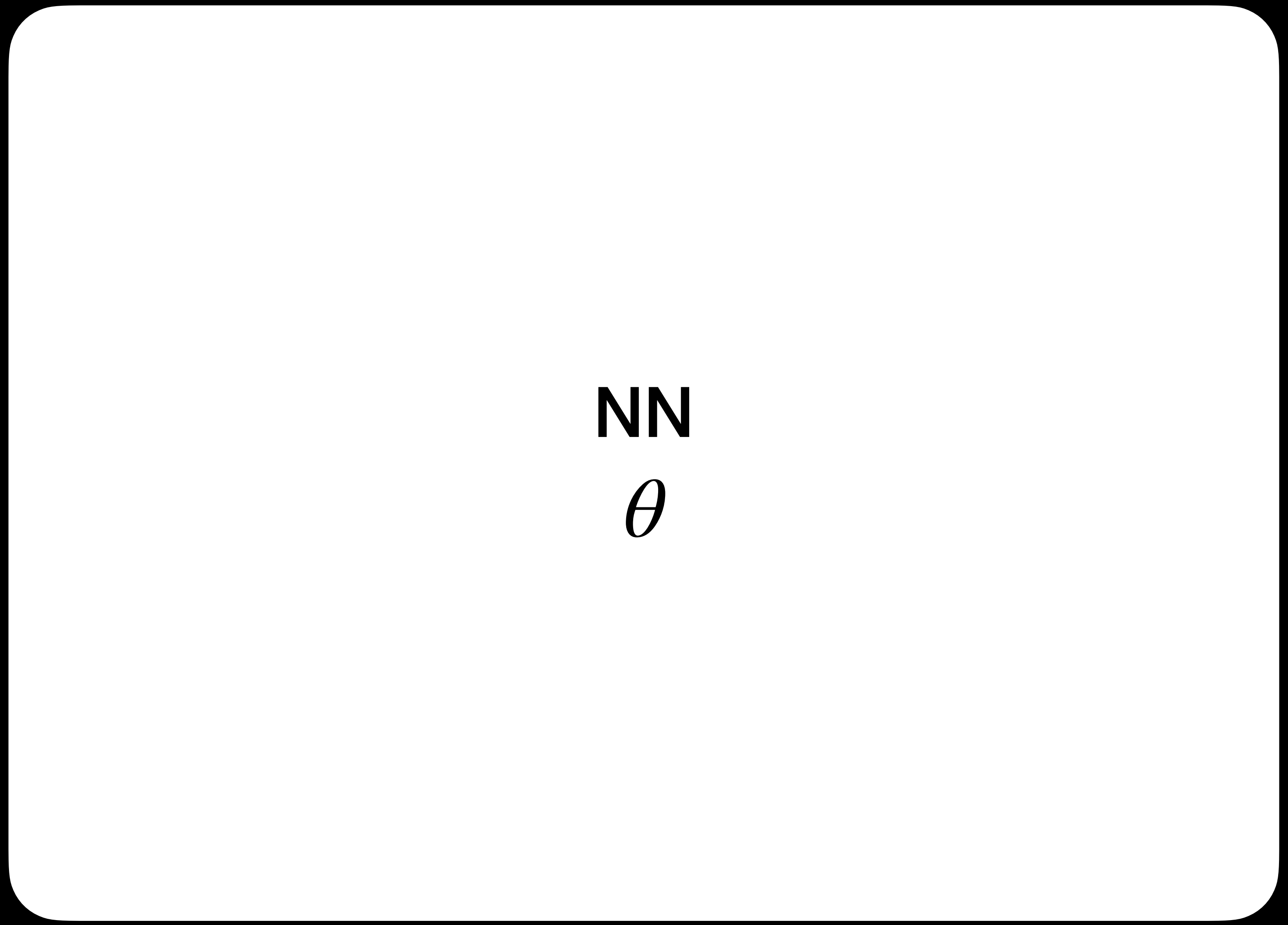
"Configuration in, amplitude out"

$\langle\phi_0|$ →  → $f_\theta(\phi_0)$

NQS workflow for LGTs

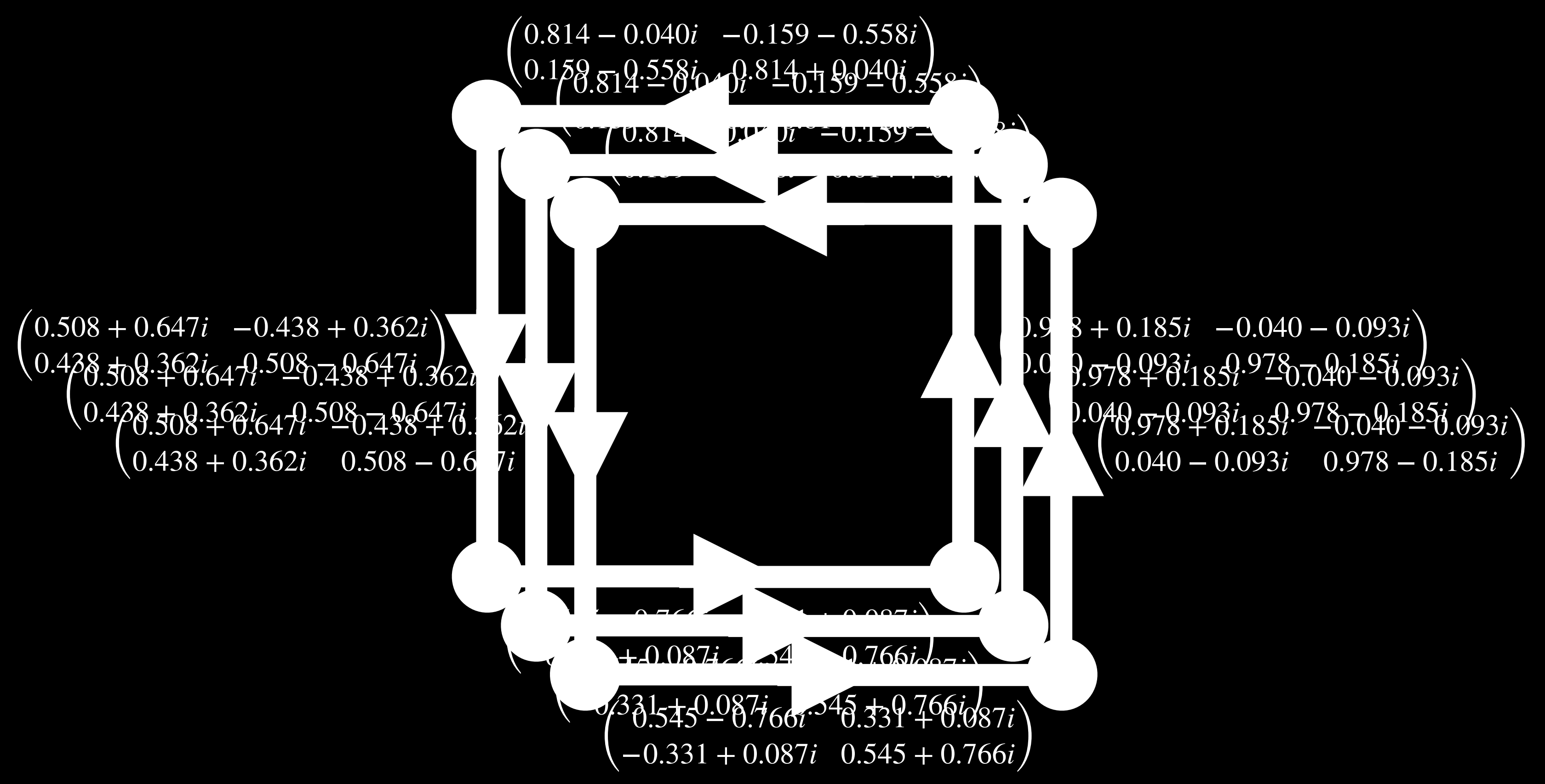
$\{\mathbf{U}_i\} \sim |\psi_\theta\rangle$





NQS workflow for LGTs

$\{\mathbf{U}_i\} \sim |\psi_\theta\rangle$



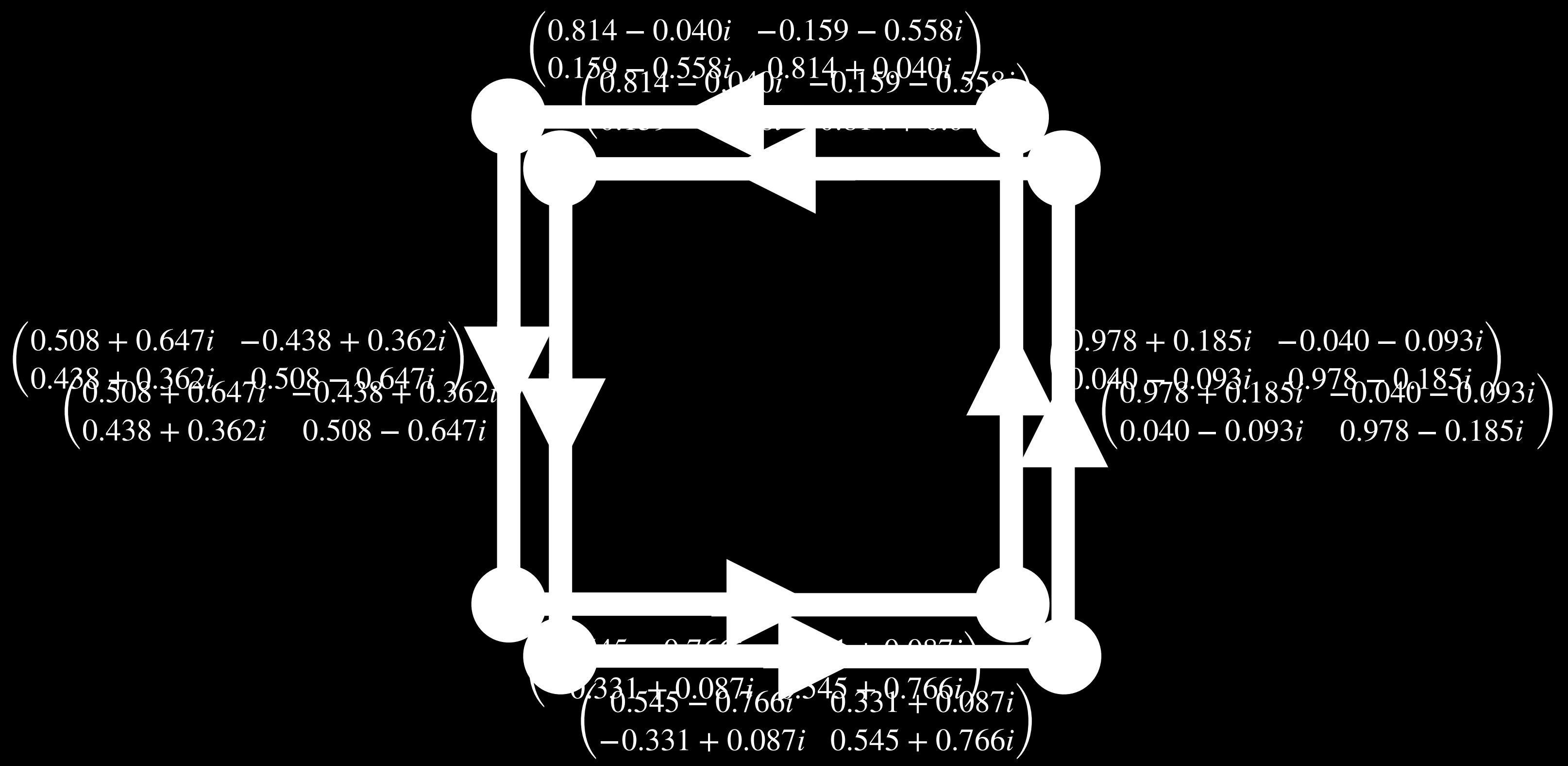
NN

θ

$f_\theta(\mathbf{U_i})$

NQS workflow for LGTs

$$\{\mathbf{U}_i\} \sim |\psi_\theta\rangle$$



NN
θ

$$f_\theta(\mathbf{U_i})$$

$$f_\theta(\mathbf{U_{i'}})$$

NQS workflow for LGTs

NN
θ

$f_\theta(\mathbf{U}_\mathbf{i})$

$f_\theta(\mathbf{U}_{\mathbf{i}'})$

$f_\theta(\mathbf{U}_{\mathbf{i}''})$

$f_\theta(\mathbf{U}_{\mathbf{i}'''})$

NQS workflow for LGTs

NN
θ

$f_\theta(\mathbf{U_i})$

$f_\theta(\mathbf{U_{i'}})$

$f_\theta(\mathbf{U_{i''}})$

$f_\theta(\mathbf{U_{i'''}})$

$$|\psi_\theta\rangle = \sum_i f_\theta(\mathbf{U}_i)\,|\,\mathbf{U}_i\rangle$$

NQS workflow for LGTs

NN
θ

$f_\theta(\mathbf{U_i})$

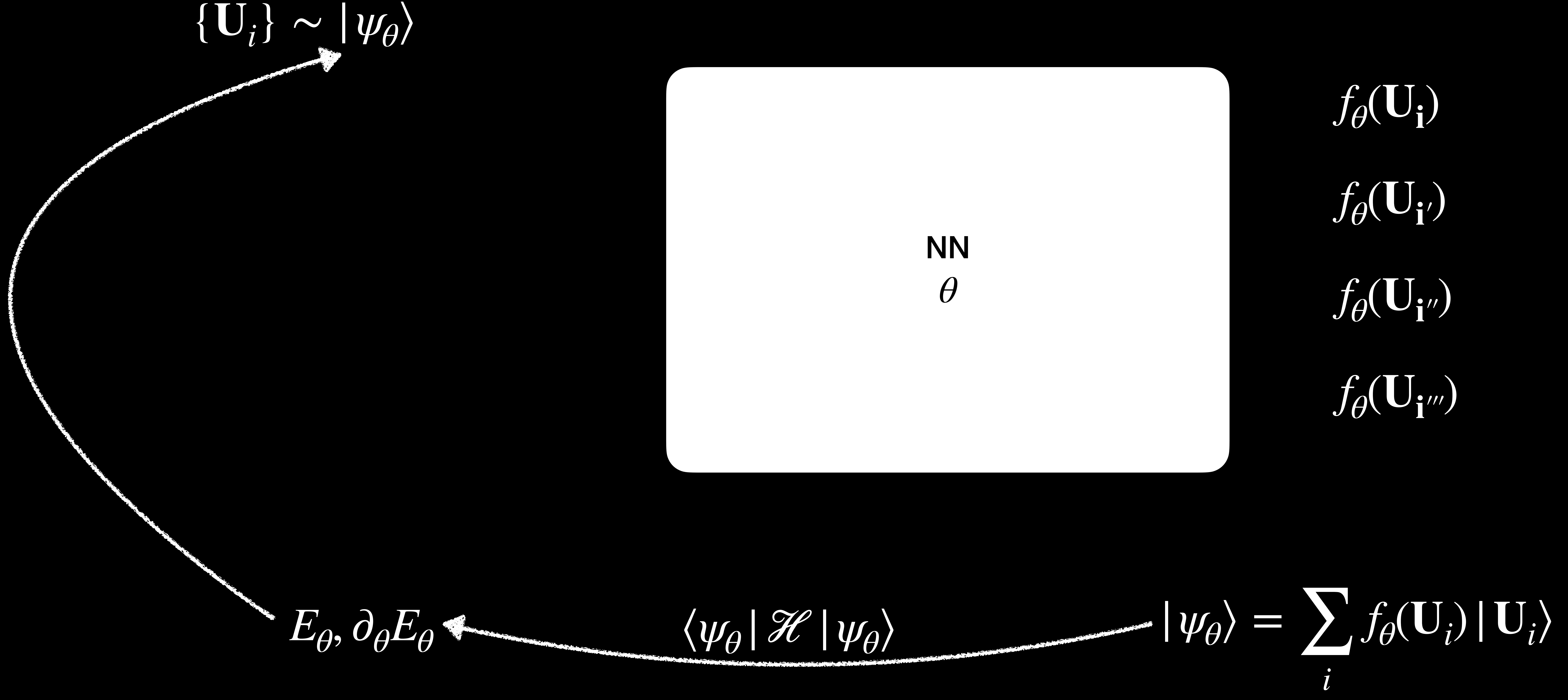
$f_\theta(\mathbf{U_{i'}})$

$f_\theta(\mathbf{U_{i''}})$

$f_\theta(\mathbf{U_{i'''}})$

$E_\theta, \partial_\theta E_\theta$ ← $\langle \psi_\theta | \mathscr{H} | \psi_\theta \rangle$ ← $|\psi_\theta\rangle = \sum_i f_\theta(\mathbf{U}_i) | \mathbf{U}_i \rangle$

NQS workflow for LGTs



NQS workflow for LGTs

$\{\mathbf{U}_i\} \sim |\psi_\theta\rangle$

NN
θ

$f_\theta(\mathbf{U_i})$

$f_\theta(\mathbf{U_{i'}})$

$f_\theta(\mathbf{U_{i''}})$

$f_\theta(\mathbf{U_{i'''}})$

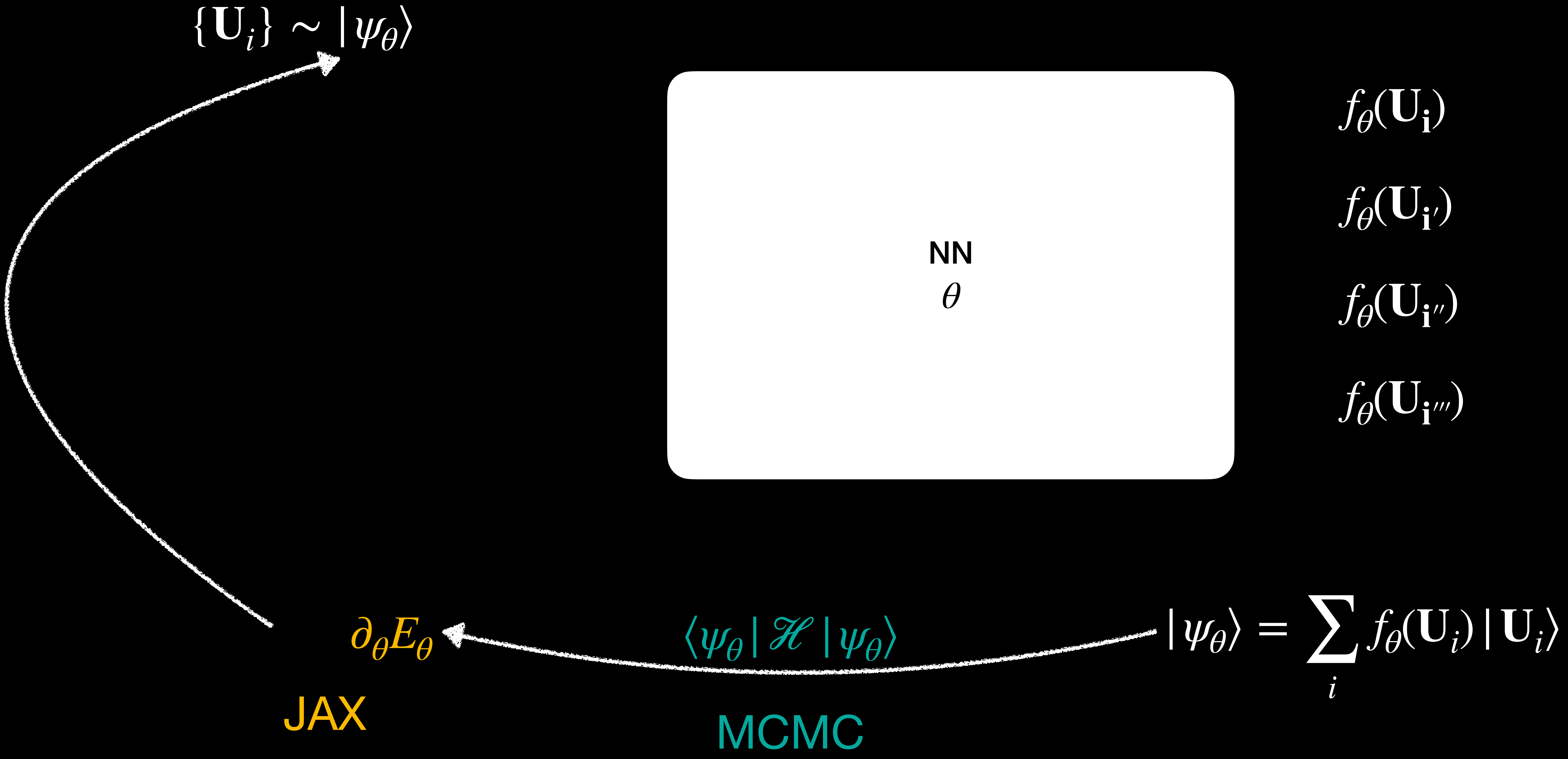
$E_\theta, \partial_\theta E_\theta$

$\langle\psi_\theta|\mathscr{H}|\psi_\theta\rangle$

$$|\psi_\theta\rangle = \sum_i f_\theta(\mathbf{U}_i)\,|\mathbf{U}_i\rangle$$

MCMC

NQS workflow for LGTs



NQS workflow for LGTs

Magnetic basis?

$\{\mathbf{U}_i\} \sim |\psi_\theta\rangle$

NN
θ

$f_\theta(\mathbf{U_i})$

$f_\theta(\mathbf{U_{i'}})$

$f_\theta(\mathbf{U_{i''}})$

$f_\theta(\mathbf{U_{i'''}})$

$\partial_\theta E_\theta$

JAX

$\langle\psi_\theta|\mathscr{H}|\psi_\theta\rangle$

MCMC

$$|\psi_\theta\rangle = \sum_i f_\theta(\mathbf{U}_i)\,|\mathbf{U}_i\rangle$$

VMC in other contexts

- Frustrated quantum spin systems

Viteritti, L. *et al*. Phys. Rev. Lett. **130**, 236401

- Ultra-cold Fermi gasses

Kim, J. *et al.* Commun Phys **7**, 148

- Helium-4

Linteau, D. *et al.* Phys. Rev. Lett. **134**, 246001

- Neutron stars

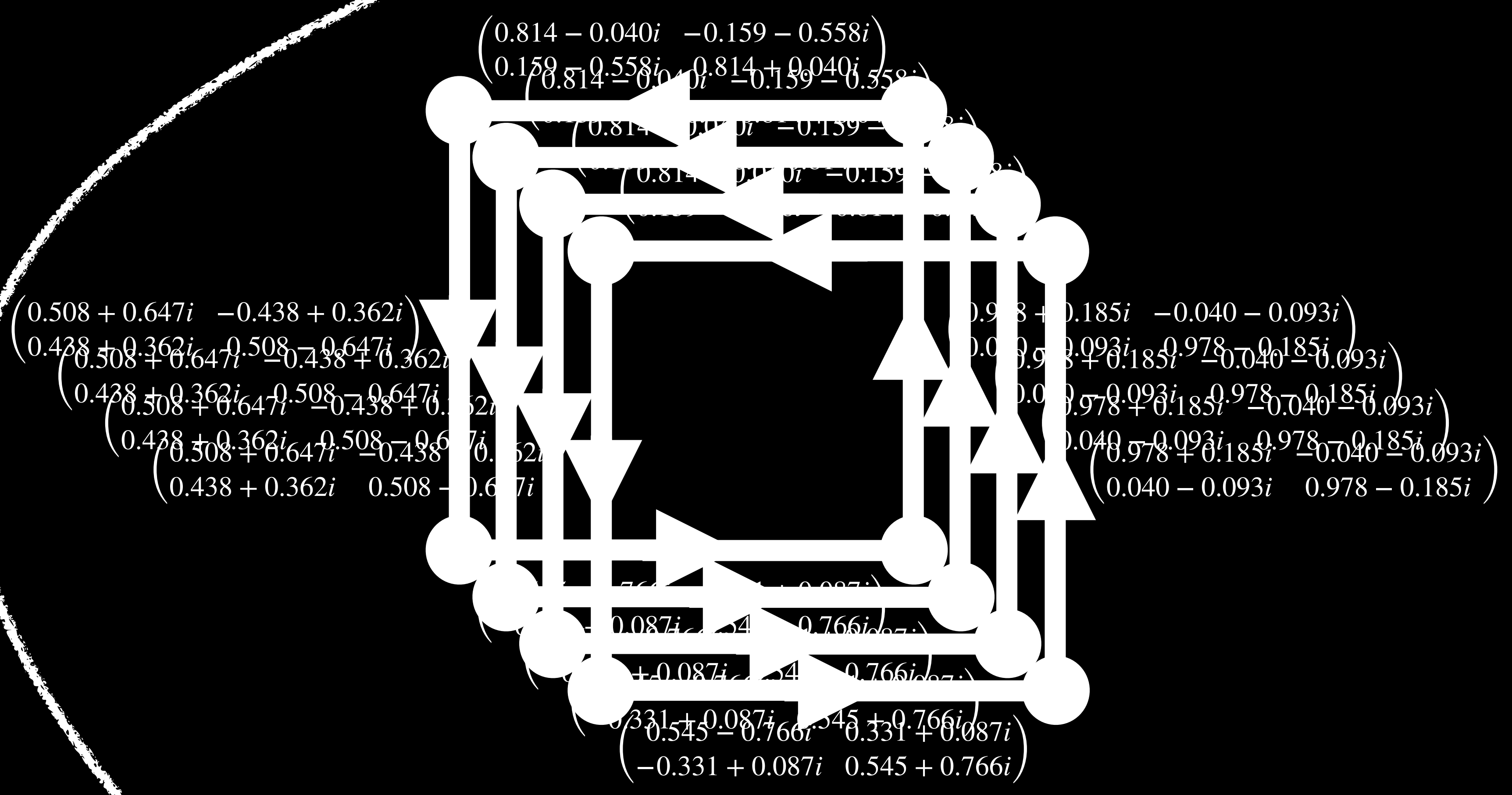
Fore, B. *et al.* Phys. Rev. Research **5**, 033062

- “Foundation” models

Scherbela, M. *et al.* arXiv 2303.09949 & Rende, R. Nat Commun **16**, 7213

NQS workflow for LGTs

$\{\mathbf{U}_i\} \sim |\psi_\theta\rangle$



NN

θ

$f_\theta(\mathbf{U_i})$

$f_\theta(\mathbf{U_{i'}})$

$f_\theta(\mathbf{U_{i''}})$

$f_\theta(\mathbf{U_{i'''}})$

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

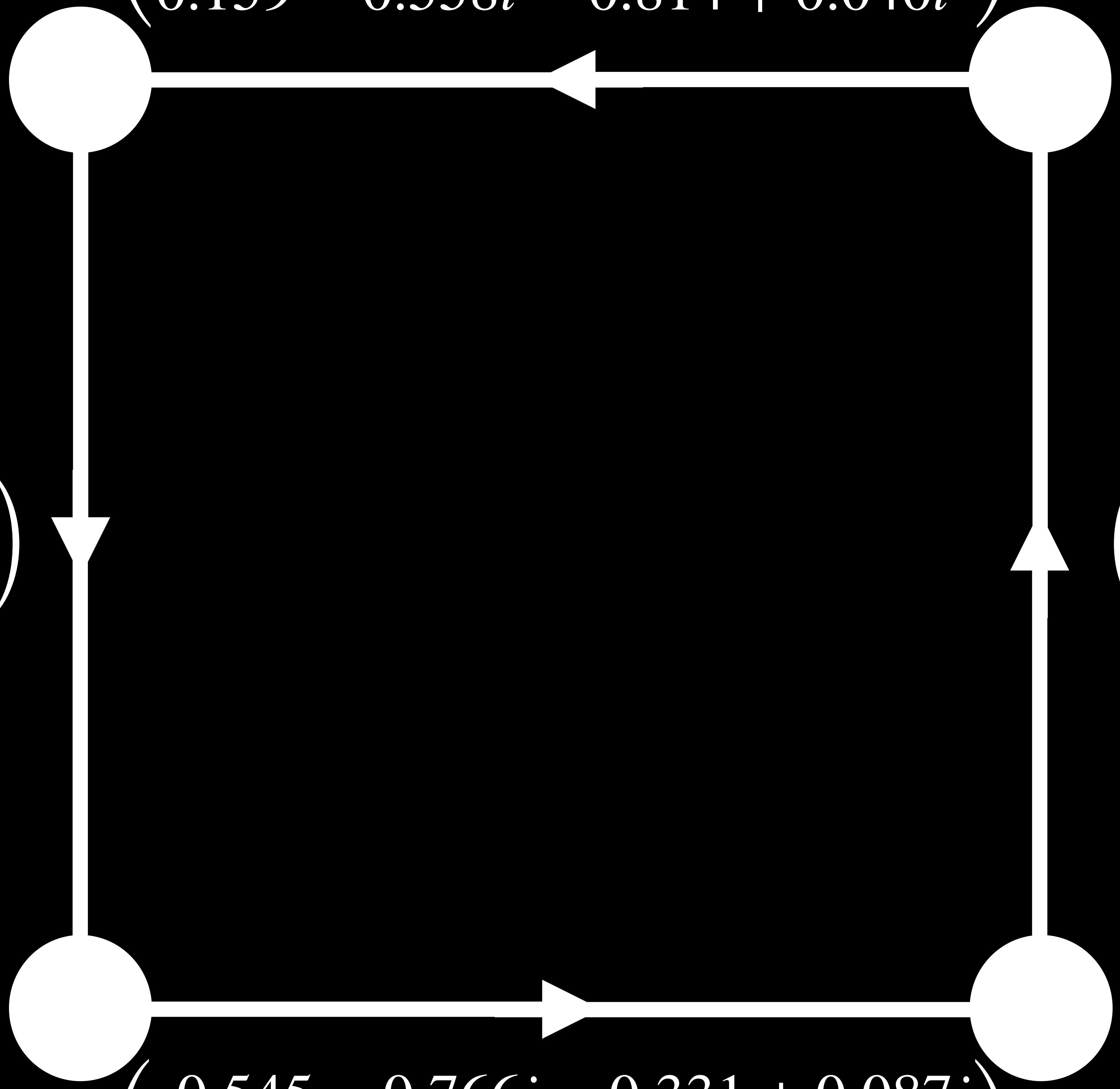
$\langle \psi_\theta | \mathscr{H} | \psi_\theta \rangle$

$$|\psi_\theta\rangle = \sum_i f_\theta(\mathbf{U}_i)\,|\,\mathbf{U}_i\rangle$$

Gauge symmetry

SU(2) Hamiltonian

Gauge symmetry



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

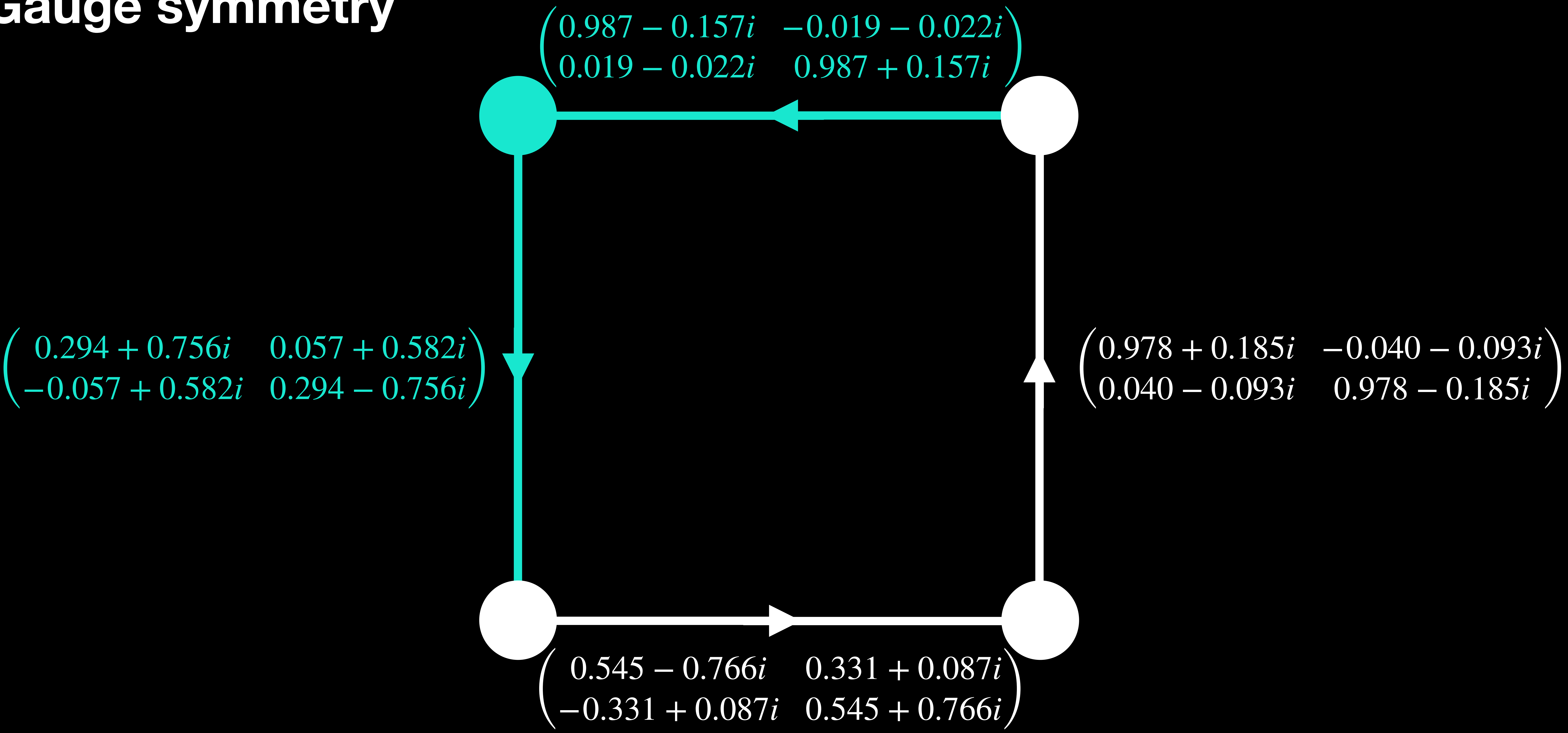
$$\begin{pmatrix} 0.508 + 0.647i & -0.438 + 0.362i \\ 0.438 + 0.362i & 0.508 - 0.647i \end{pmatrix}$$

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

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

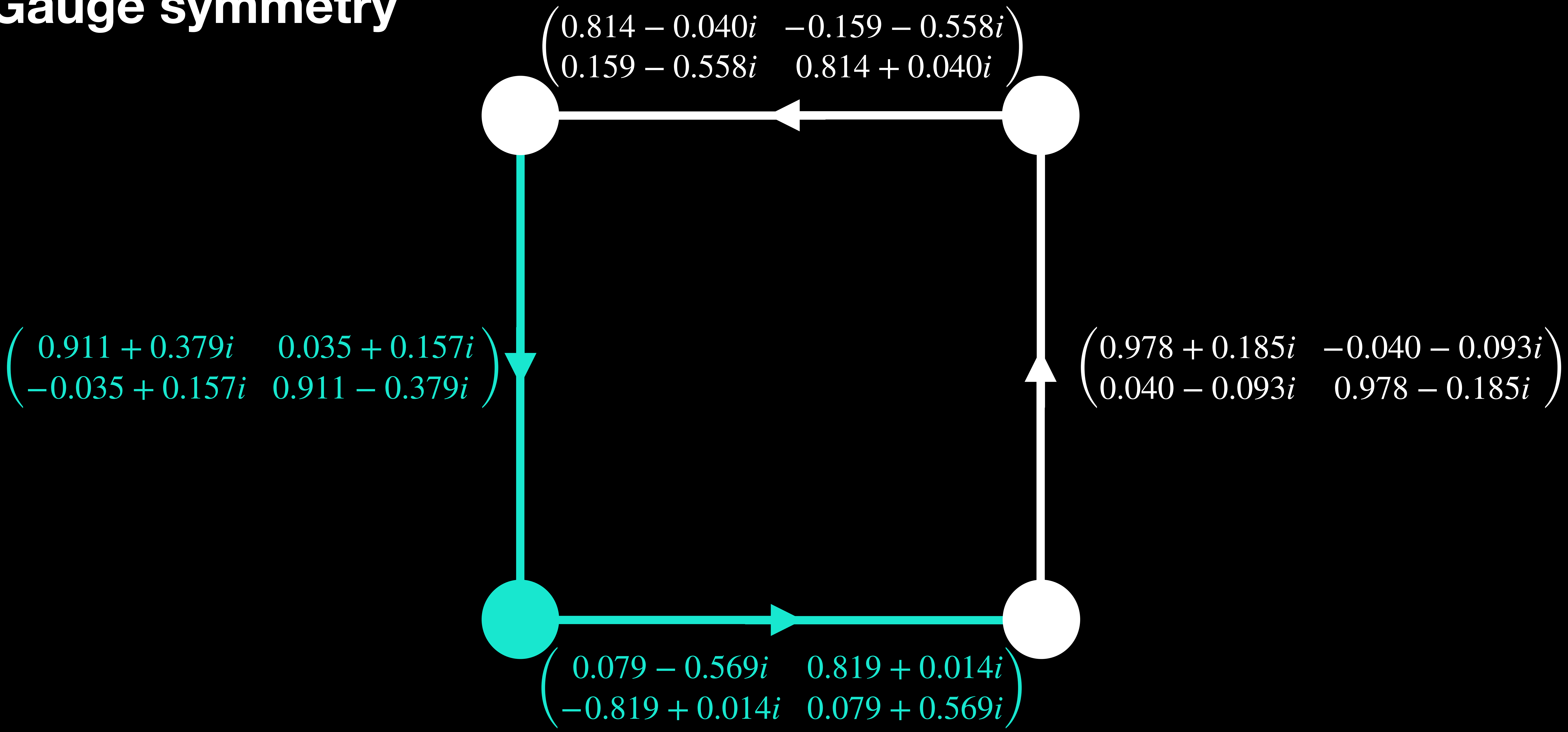
SU(2) Hamiltonian

Gauge symmetry



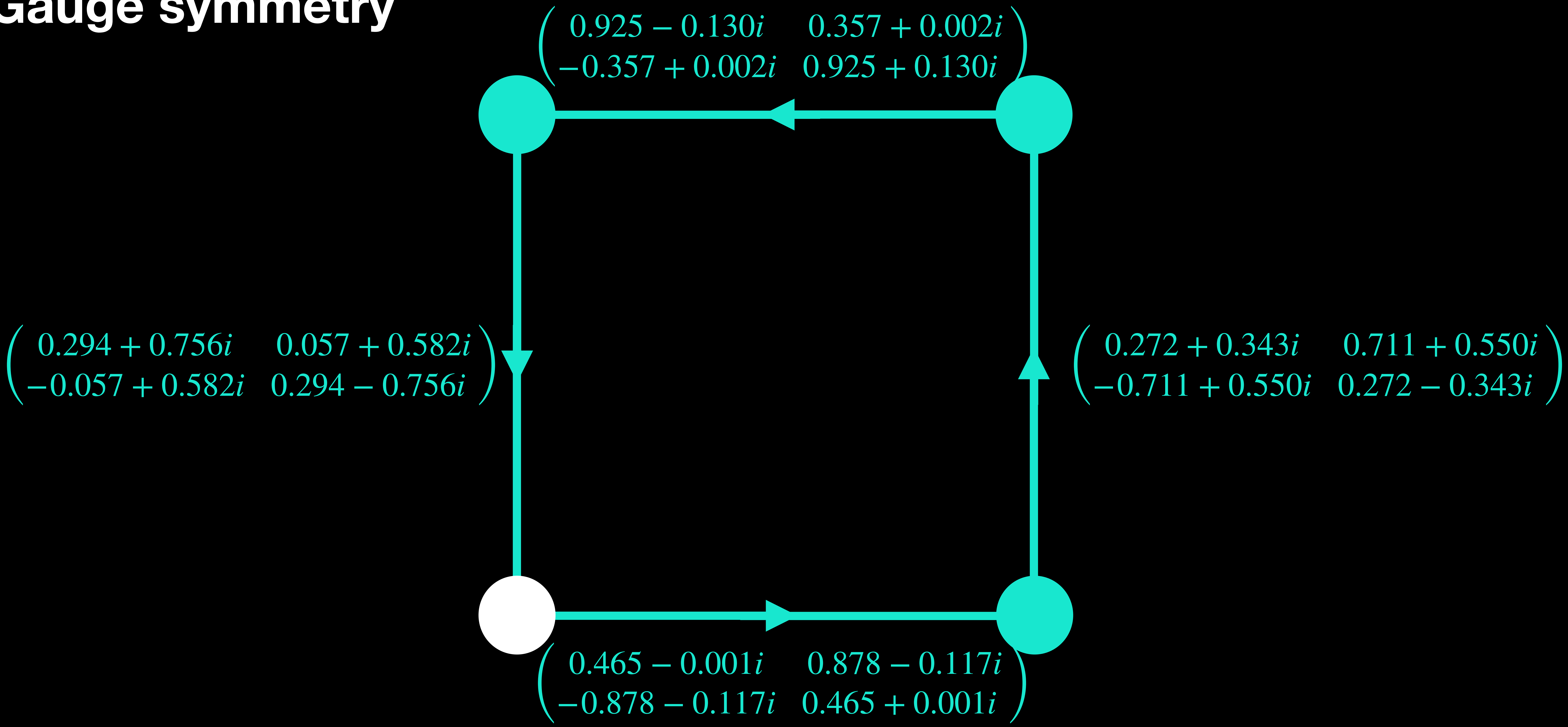
SU(2) Hamiltonian

Gauge symmetry



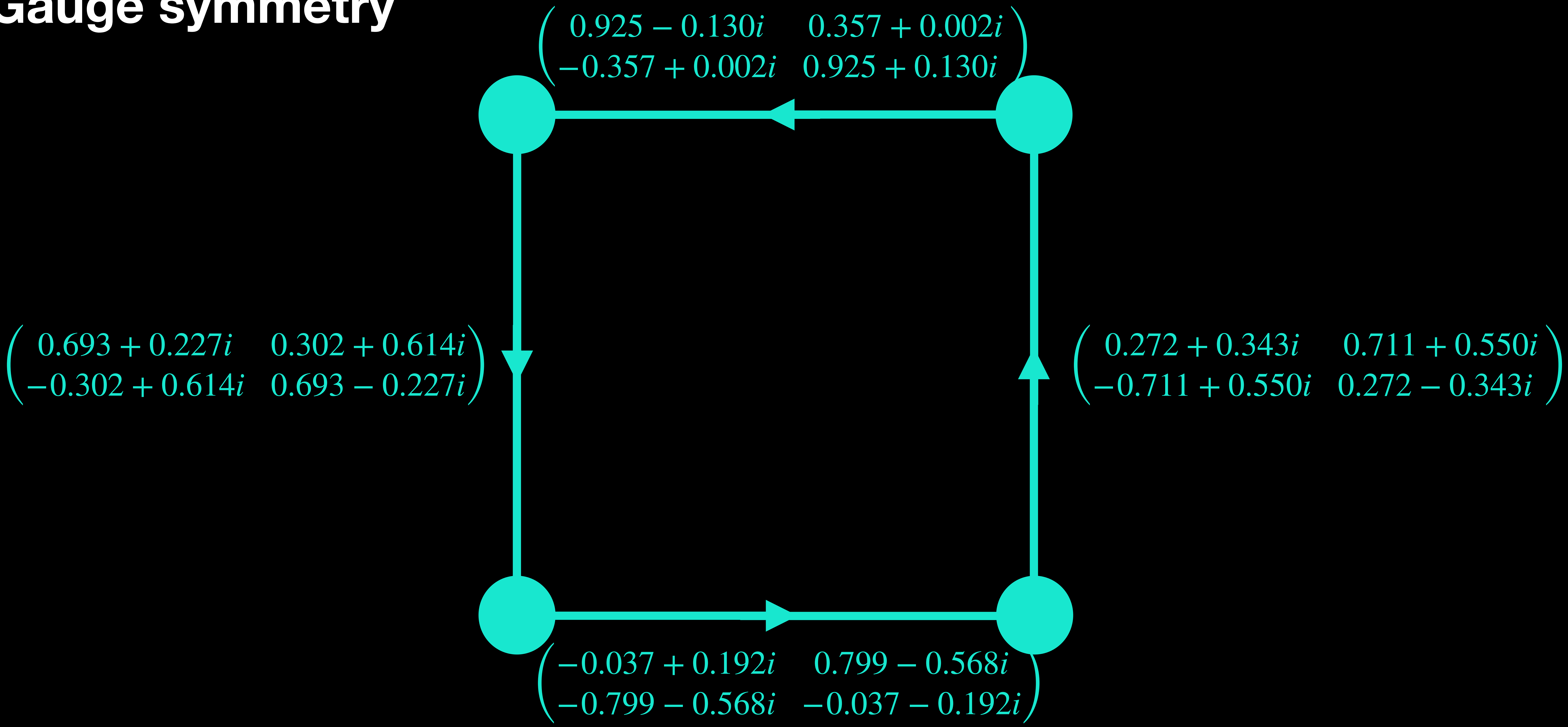
SU(2) Hamiltonian

Gauge symmetry



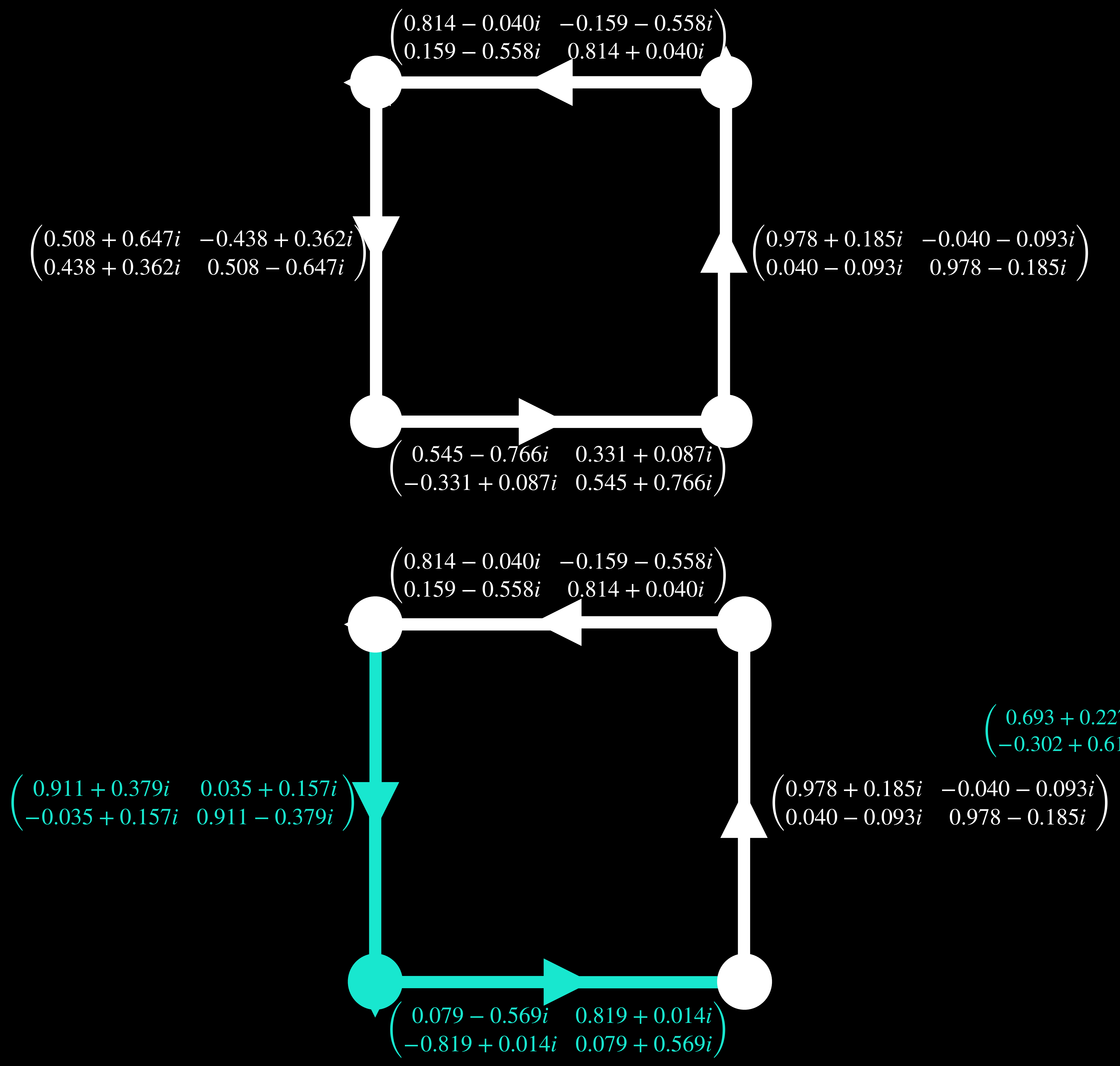
SU(2) Hamiltonian

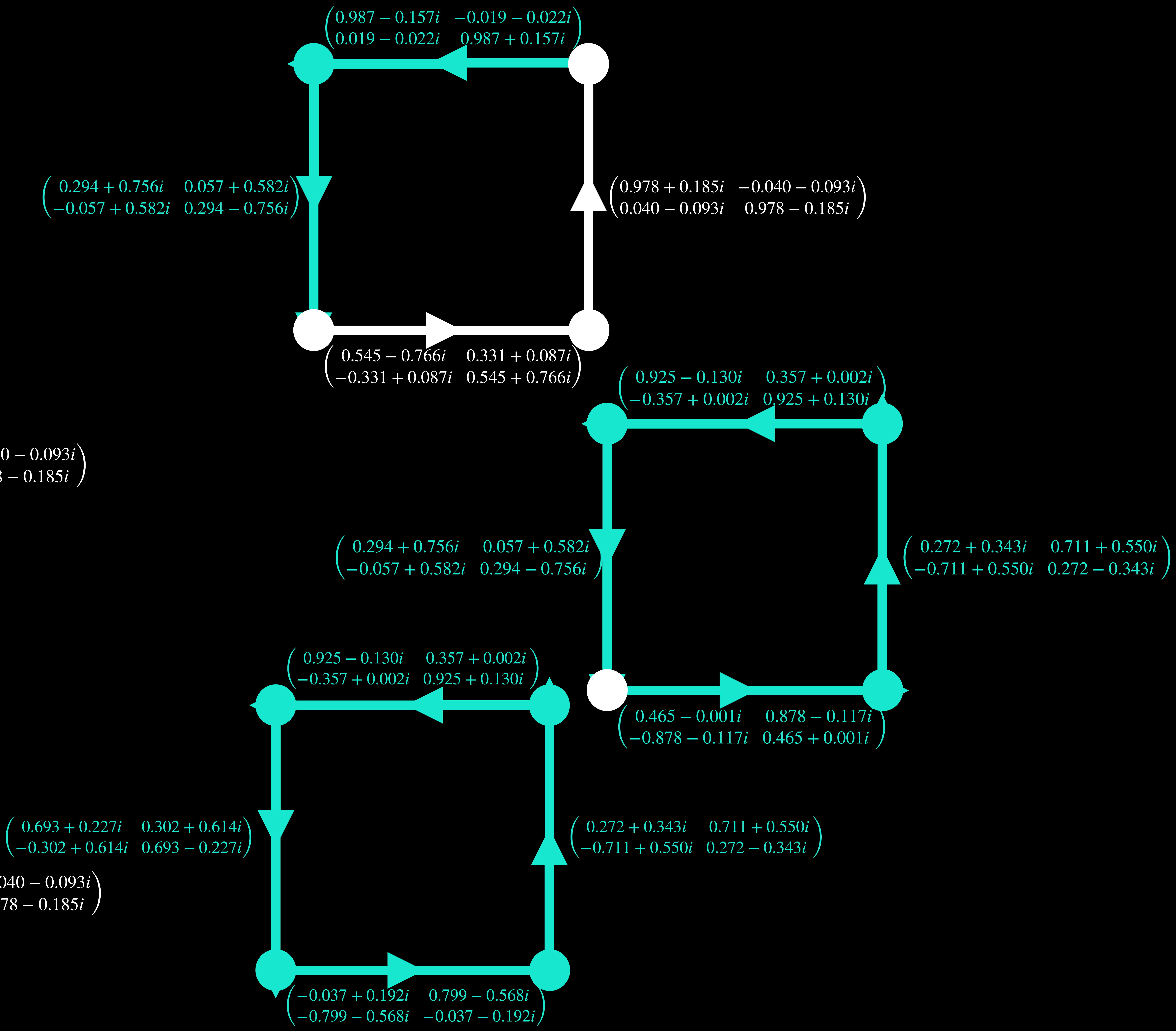
Gauge symmetry



SU(2) Hamiltonian

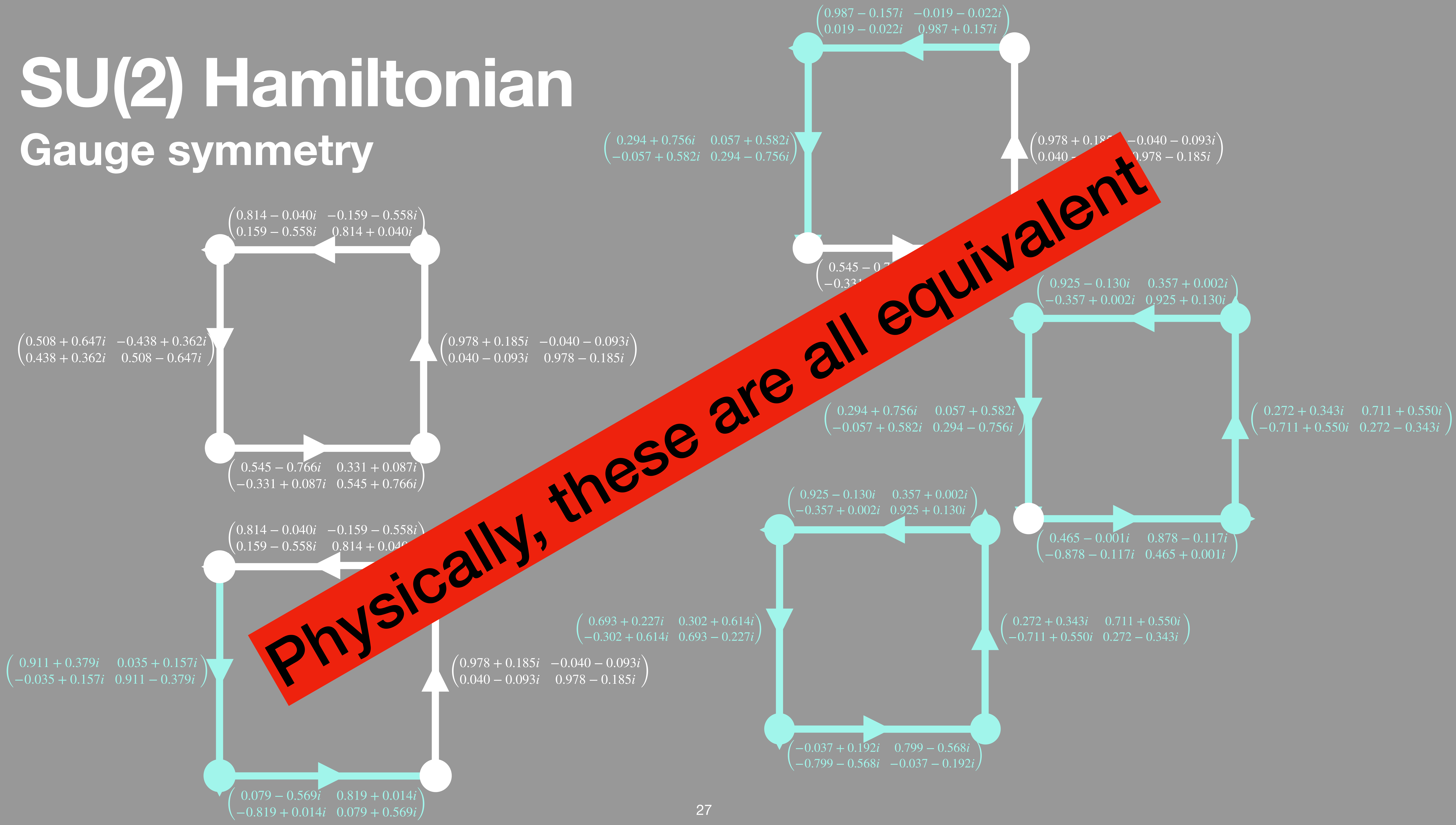
Gauge symmetry





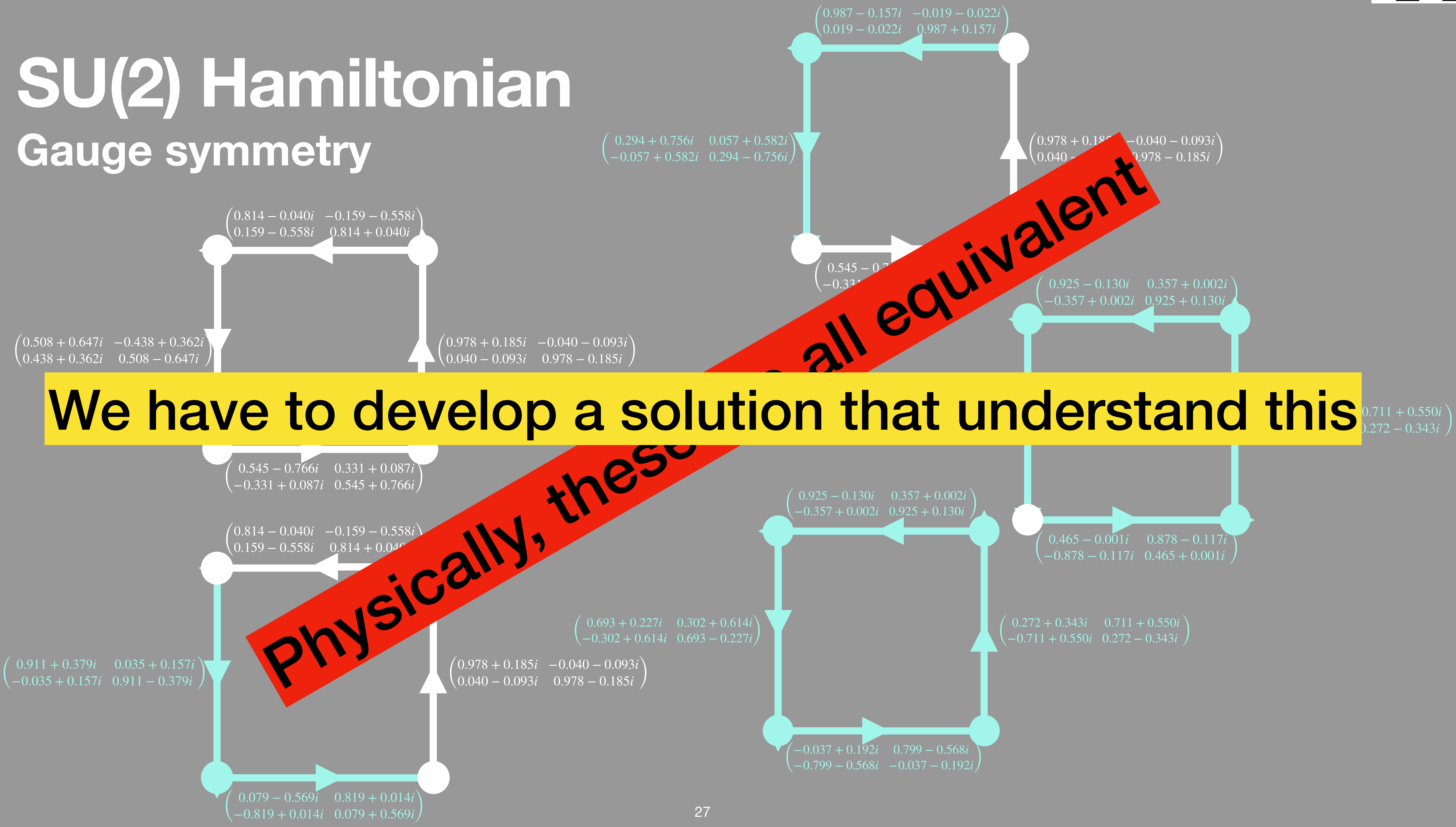
SU(2) Hamiltonian

Gauge symmetry



SU(2) Hamiltonian

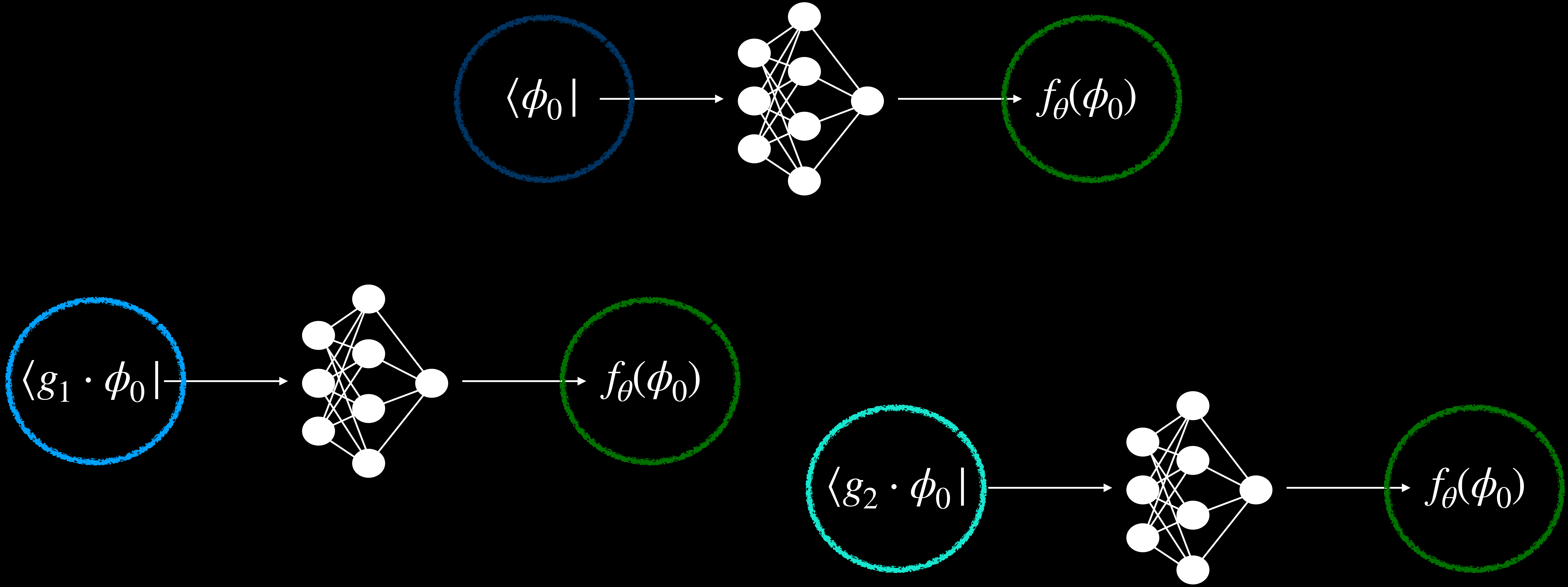
Gauge symmetry



Physically, these … all equivalent

We have to develop a solution that understand this

“Configuration in, amplitude out”



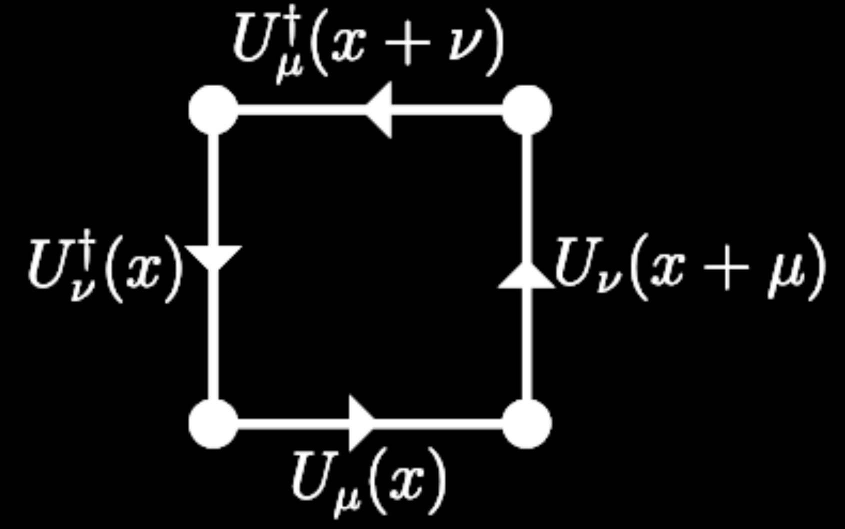
Gauge invariant ansatze

Simple SU(2) gauge invariant ansatz

- Trace of the plaquette is gauge invariant
  - Build our ansatz from this

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

$$\prod \mathrm{Tr}$$

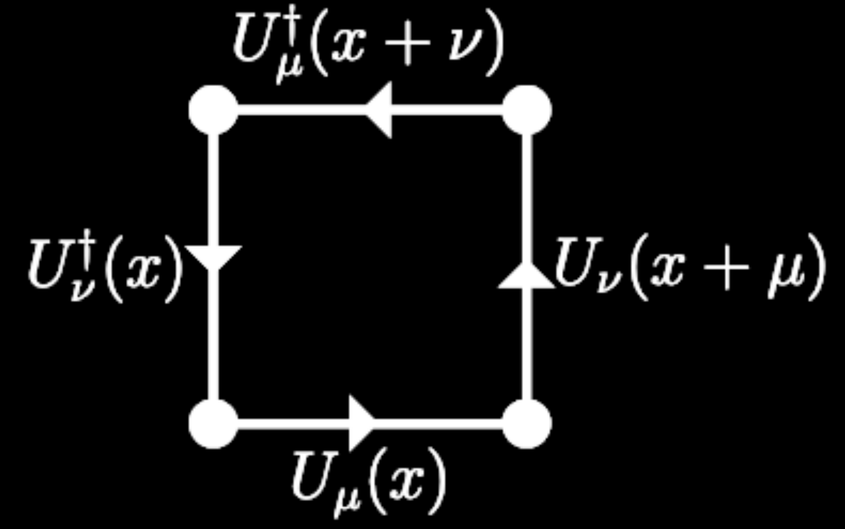


(2) $$f(\mathbf{U}) = \prod_{p} e^{\alpha \frac{1}{2}\mathrm{Tr}\left(P_p\right)} + \prod_{p,q} e^{\frac{1}{4}\mathrm{Tr}\left(P_p\right)\beta_{p,q}\mathrm{Tr}\left(P_q\right)}$$

Simple SU(2) gauge invariant ansatz

- Trace of the plaquette is gauge invariant
    - Build our ansatz from this

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

$\prod \mathrm{Tr}$ 

(2) $$f(\mathbf{U}) = \prod_{p} e^{\alpha \frac{1}{2} \mathrm{Tr}\left(P_p\right)} + \prod_{p,q} e^{\frac{1}{4} \mathrm{Tr}\left(P_p\right) \beta_{p,q} \mathrm{Tr}\left(P_q\right)}$$

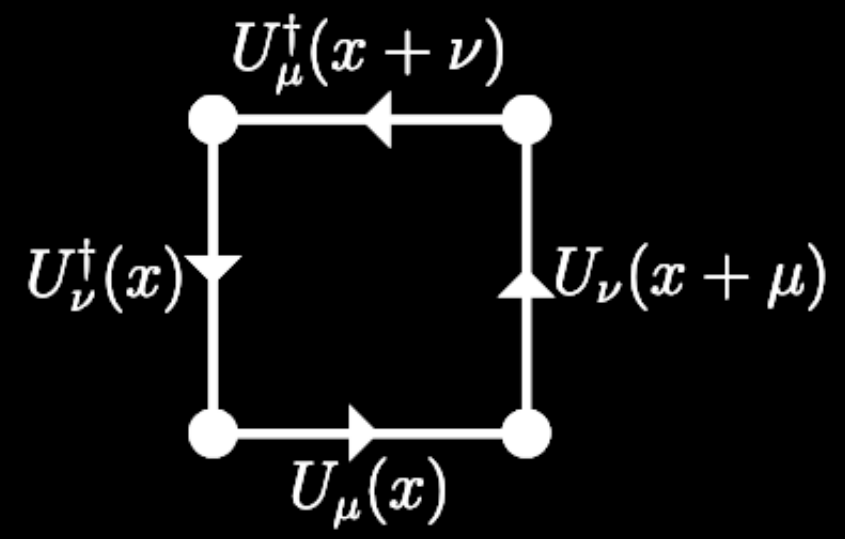
Only involves local*, scalar quantities

Simple SU(2) gauge invariant ansatz

- Trace of the plaquette is gauge invariant
  - Build our ansatz from this

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

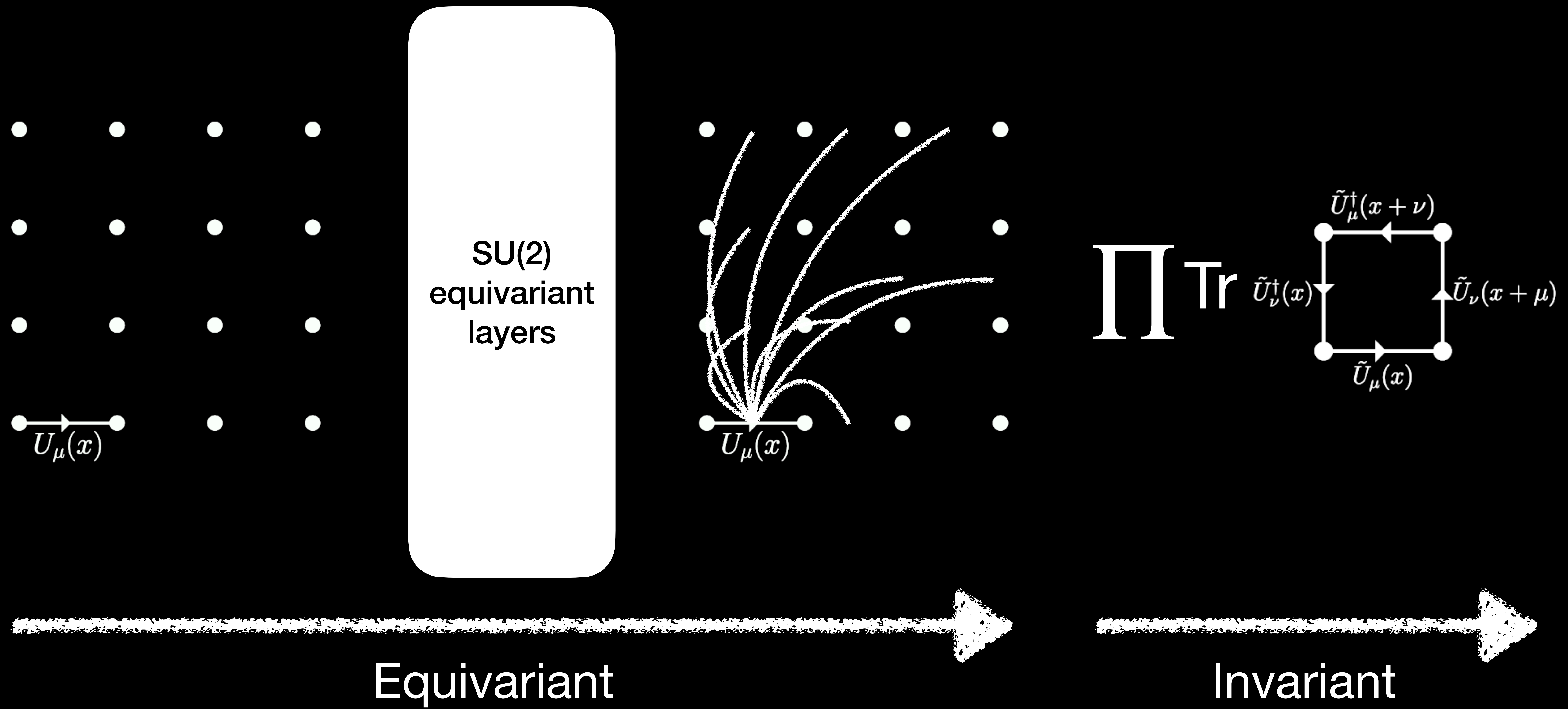
$$\prod \mathrm{Tr}$$



(2) $$f(\mathbf{U}) = \prod_{p} e^{\alpha \frac{1}{2} \mathrm{Tr}\left(P_p\right)} + \prod_{p,q} e^{\frac{1}{4} \mathrm{Tr}\left(P_p\right) \beta_{p,q} \mathrm{Tr}\left(P_q\right)}$$

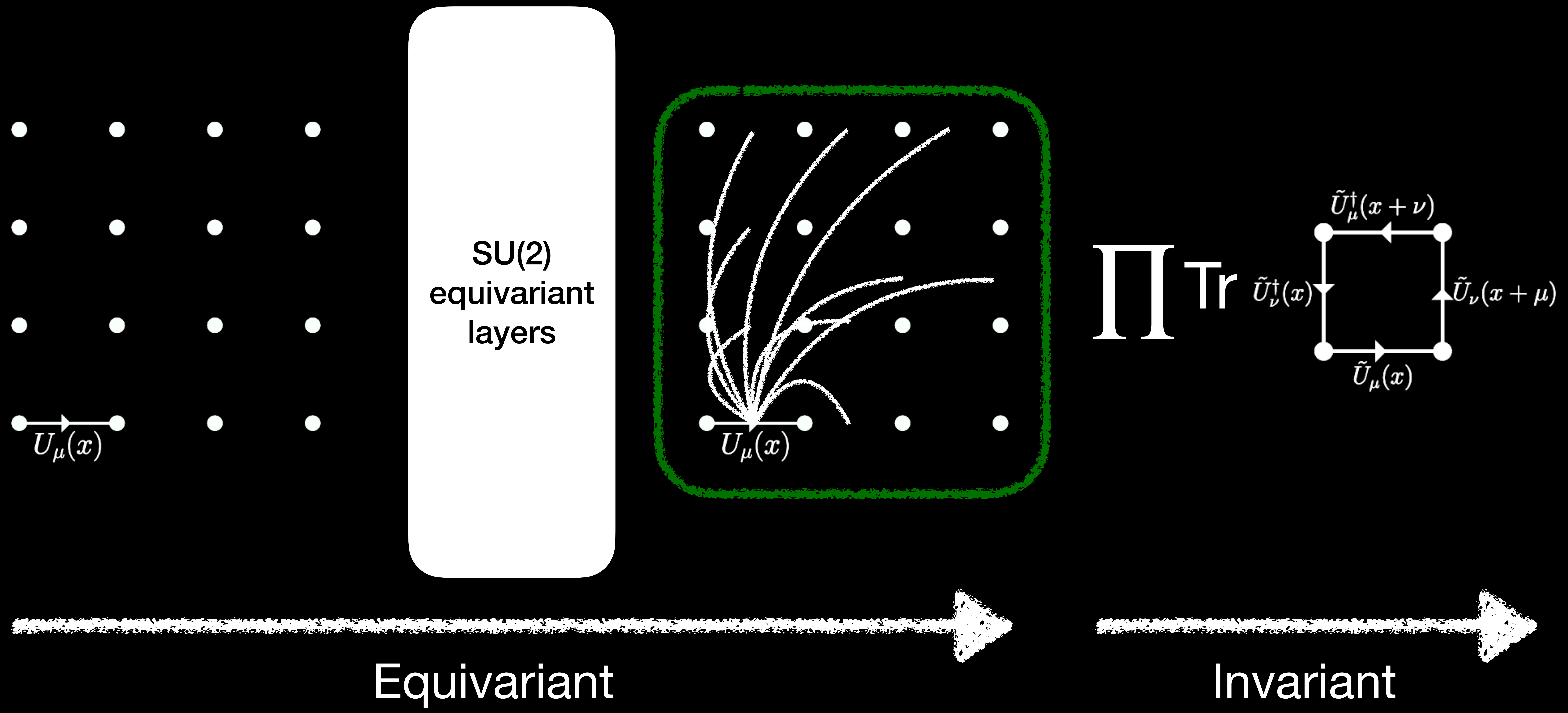
Cannot even learn clover action

Our SU(2) invariant ansatz

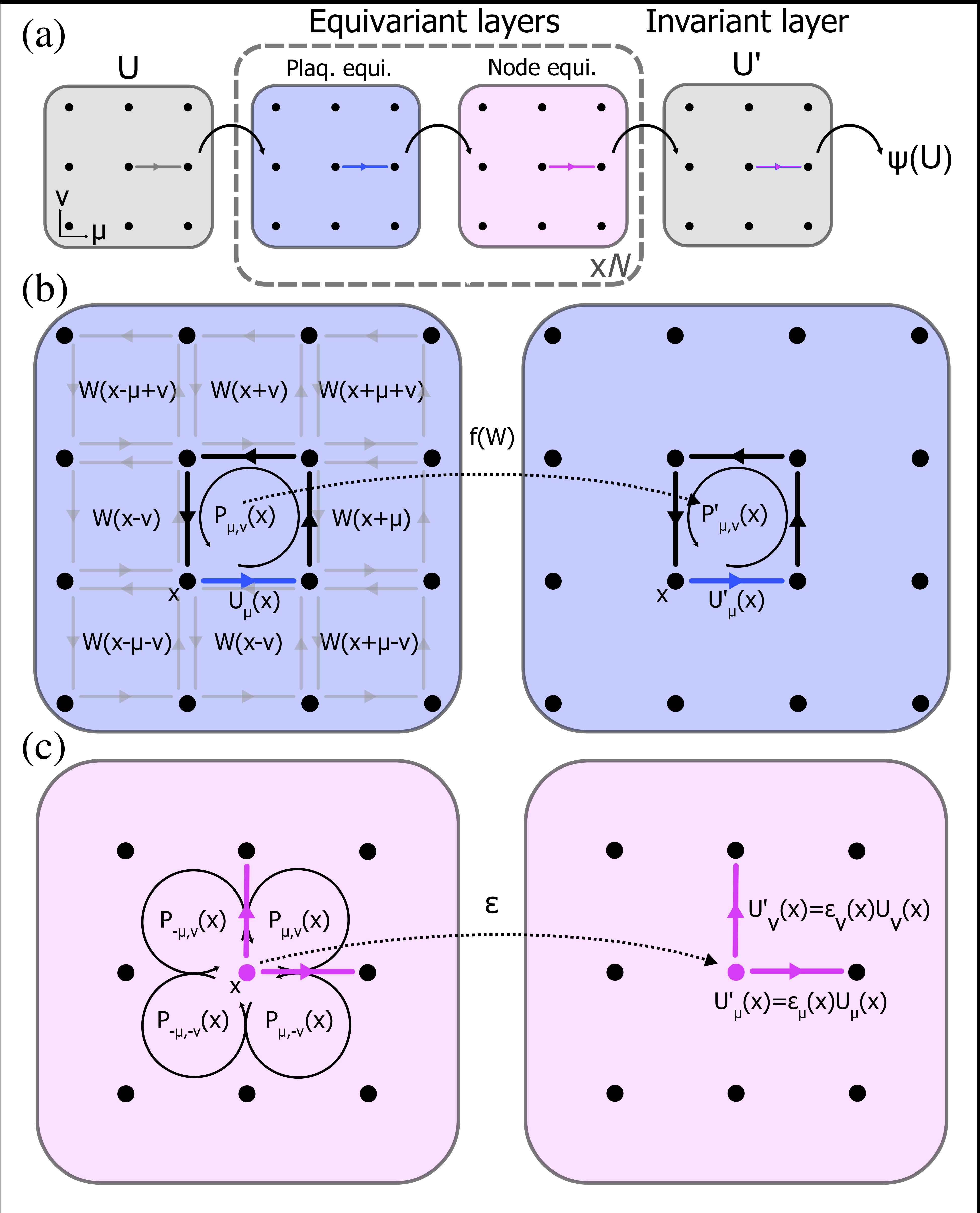


Equivariant layers from:
Kanwar, G. *et al.* Phys. Rev. Lett. **125**, 121601 & Boyda, D. L. *et* al. Phys. Rev. D **103**, 074504
Favoni, M. *et al.* Phys. Rev. Lett. **128**, 032003

Our SU(2) invariant ansatz



Equivariant layers from:
Kanwar, G. *et al.* Phys. Rev. Lett. **125**, 121601 & Boyda, D. L. *et* al. Phys. Rev. D **103**, 074504
Favoni, M. *et al.* Phys. Rev. Lett. **128**, 032003


(a)
Equivariant layers
Invariant layer
U
Plaq. equi.
Node equi.
U'
ψ(U)
ν
μ
xN
(b)
W(x-μ+ν)
W(x+ν)
W(x+μ+ν)
f(W)
W(x-ν)
W(x+μ)
x
W(x-μ-ν)
W(x-ν)
W(x+μ-ν)
x
(c)
ε
x

Recap

- Want to learn the distribution of configurations (i.e. learn the ground state wavefunction), $|\psi_\theta\rangle$
    - Gauge invariant learned distribution
    - Sample from this distribution $\{\mathbf{U}\} \sim |\psi_\theta\rangle$
- Perform measurements on these configurations

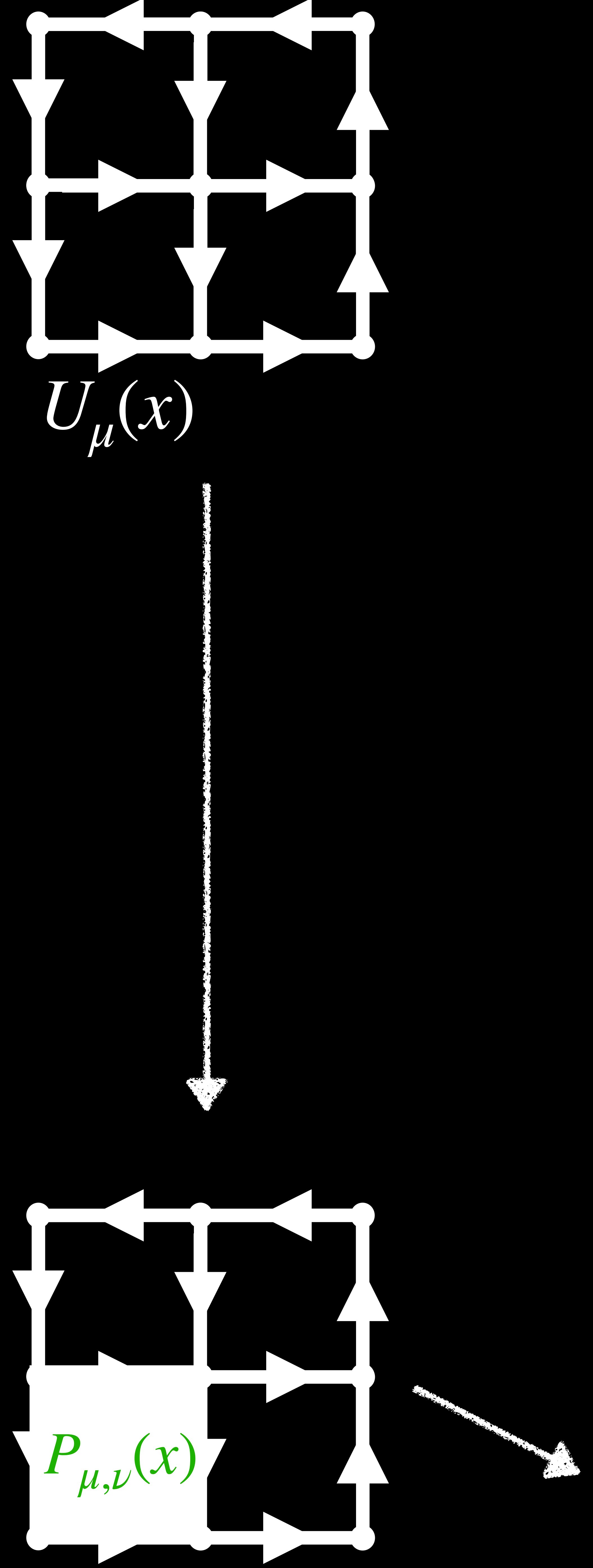
Simulations

Simulation 1

$$\mathscr{H} = -\frac{1}{2}\sum_{l}\nabla_l^2 + \lambda\sum_{p}\left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

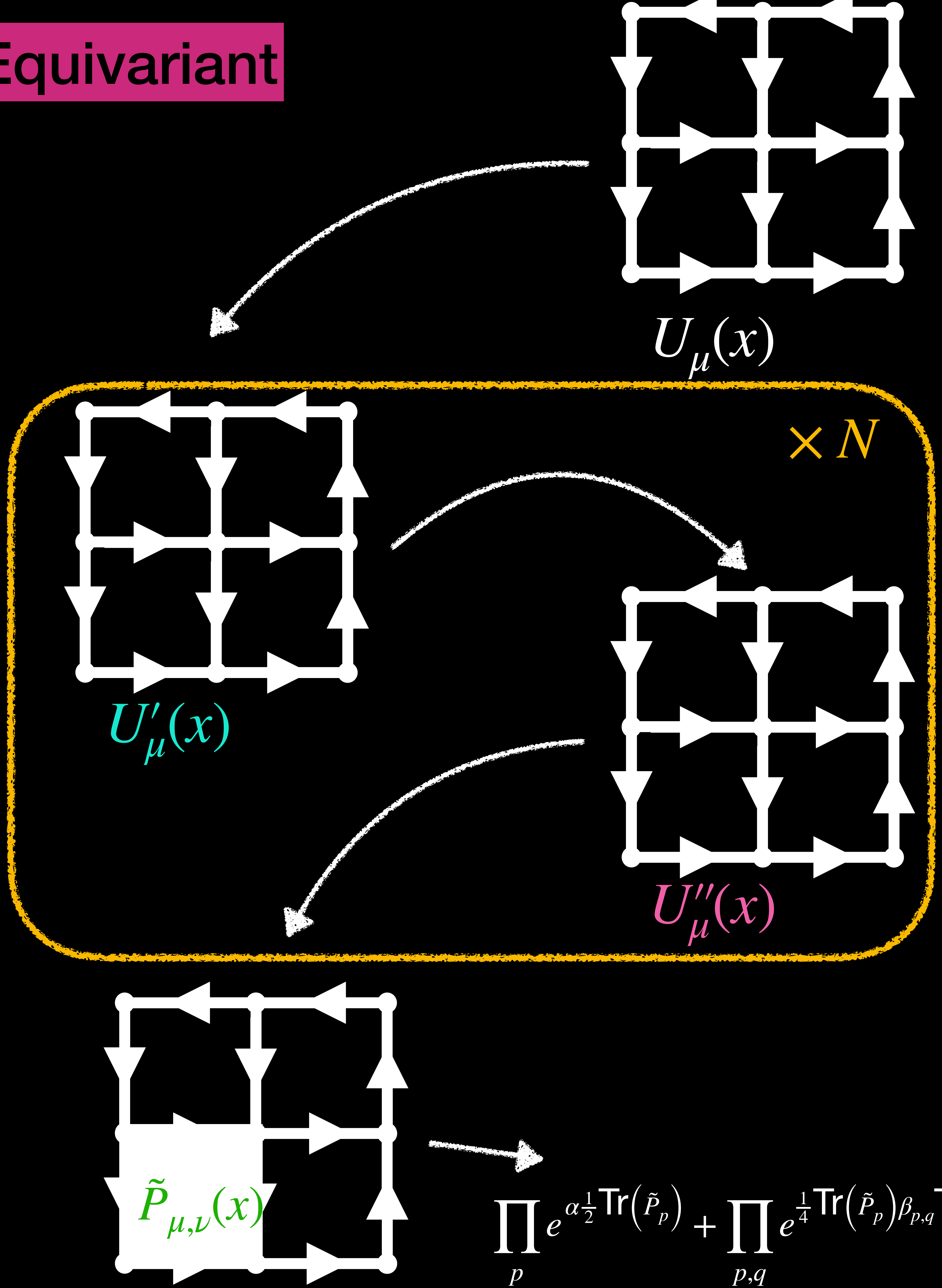
- Lattice of size 12x12
- Goal: find the lowest ground state energy at various λs
- Ansätze:
  - simple, invariant-only, ansatz -> **Jastrow**
  - multi-layer equivariant+invariant ansatz -> **Equivariant**
- Relative energy decrease coming from new ansatz:
  - $$\delta E = \frac{E_{Equivariant} - E_{Jastrow}}{E_{Jastrow}}$$

Jastrow



$$\prod_p e^{\alpha \frac{1}{2}\mathrm{Tr}\left(P_p\right)} + \prod_{p,q} e^{\frac{1}{4}\mathrm{Tr}\left(P_p\right)\beta_{p,q}\mathrm{Tr}\left(P_q\right)}$$

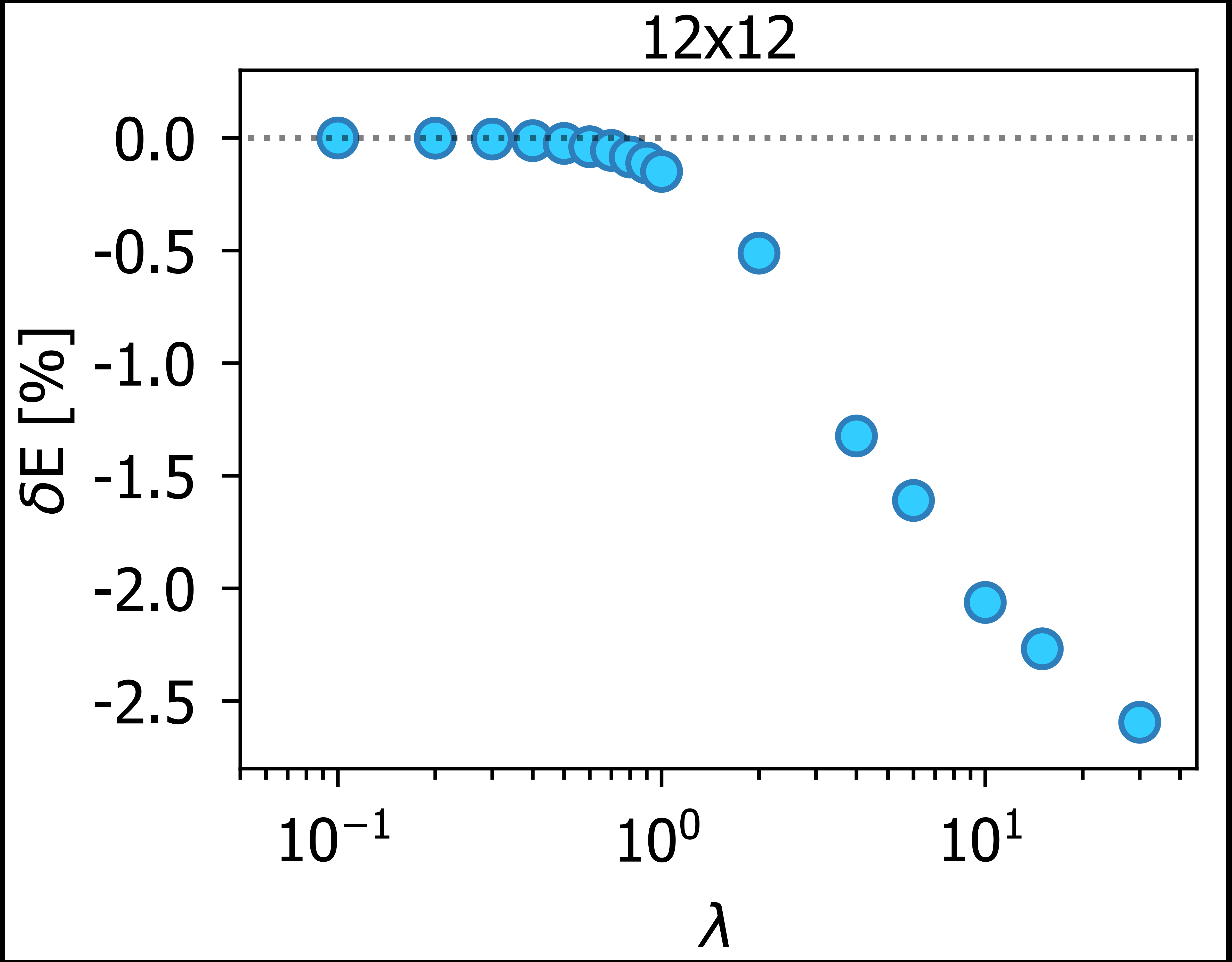
Equivariant



$$\prod_p e^{\alpha \frac{1}{2}\mathrm{Tr}\left(\tilde{P}_p\right)} + \prod_{p,q} e^{\frac{1}{4}\mathrm{Tr}\left(\tilde{P}_p\right)\beta_{p,q}\mathrm{Tr}\left(\tilde{P}_q\right)}$$

$$\mathcal{H} = -\frac{1}{2}\sum_l \nabla_l^2 + \lambda \sum_p \left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

Simulation 1



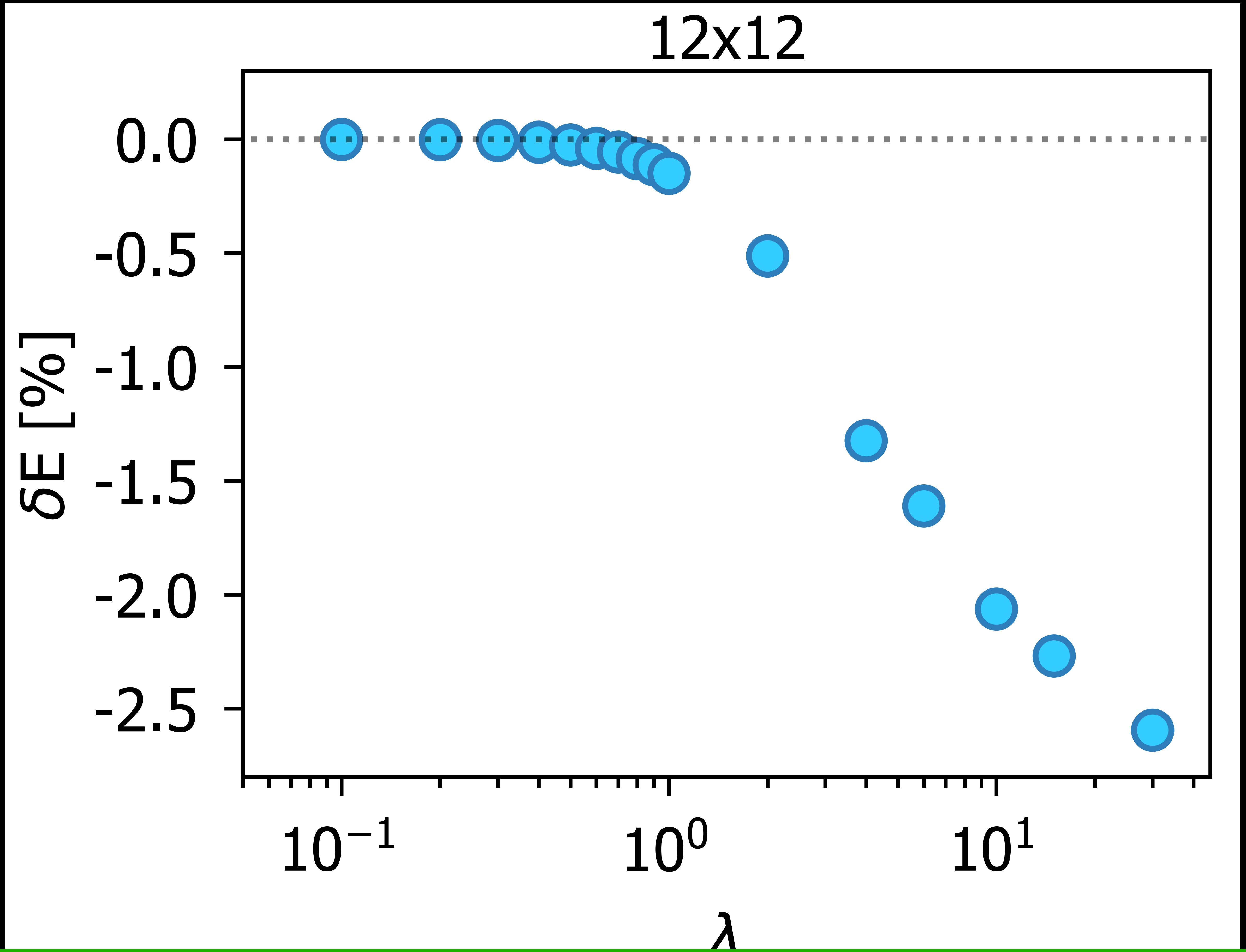
g range ~ 2.5 - 0.6

$$\delta E = \frac{E_{Equivariant} - E_{Jastrow}}{E_{Jastrow}}$$

Simulation 1

$$\mathcal{H} = -\frac{1}{2}\sum_l \nabla_l^2 + \lambda \sum_p \left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

g range ~ 2.5 - 0.6



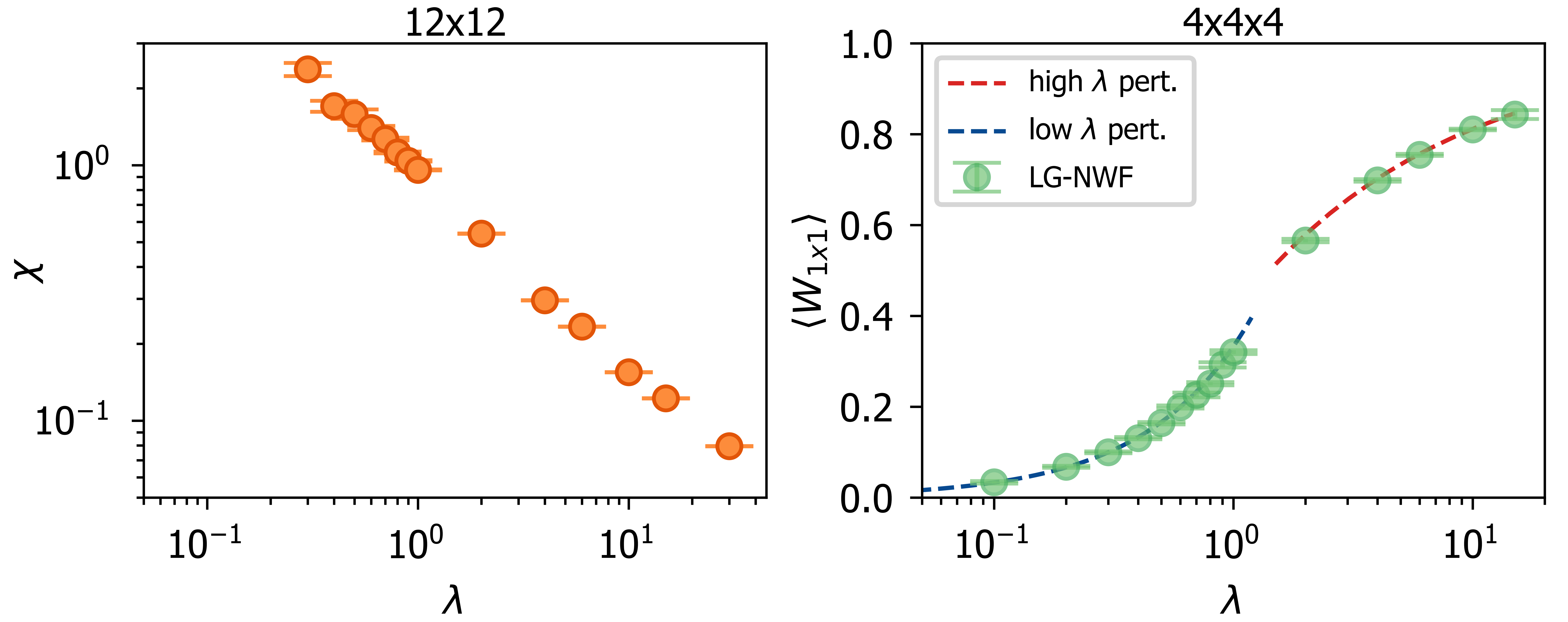
Equivariant layers needed to learn a better representation of the ground state

Simulation 2

- Lattice of size 12x12 and 4x4x4
- Goal: measure physical observables
- Ansätze:
  - Equivariant
- Average 1x1 Wilson loop $\langle W^{1\times 1}\rangle$
- Creutz ratio $\chi = -\log\left(\frac{\langle W^{1\times 1}\rangle\langle W^{2\times 2}\rangle}{\langle W^{1\times 2}\rangle\langle W^{2\times 1}\rangle}\right)$

g range ~ 2.5 - 0.6

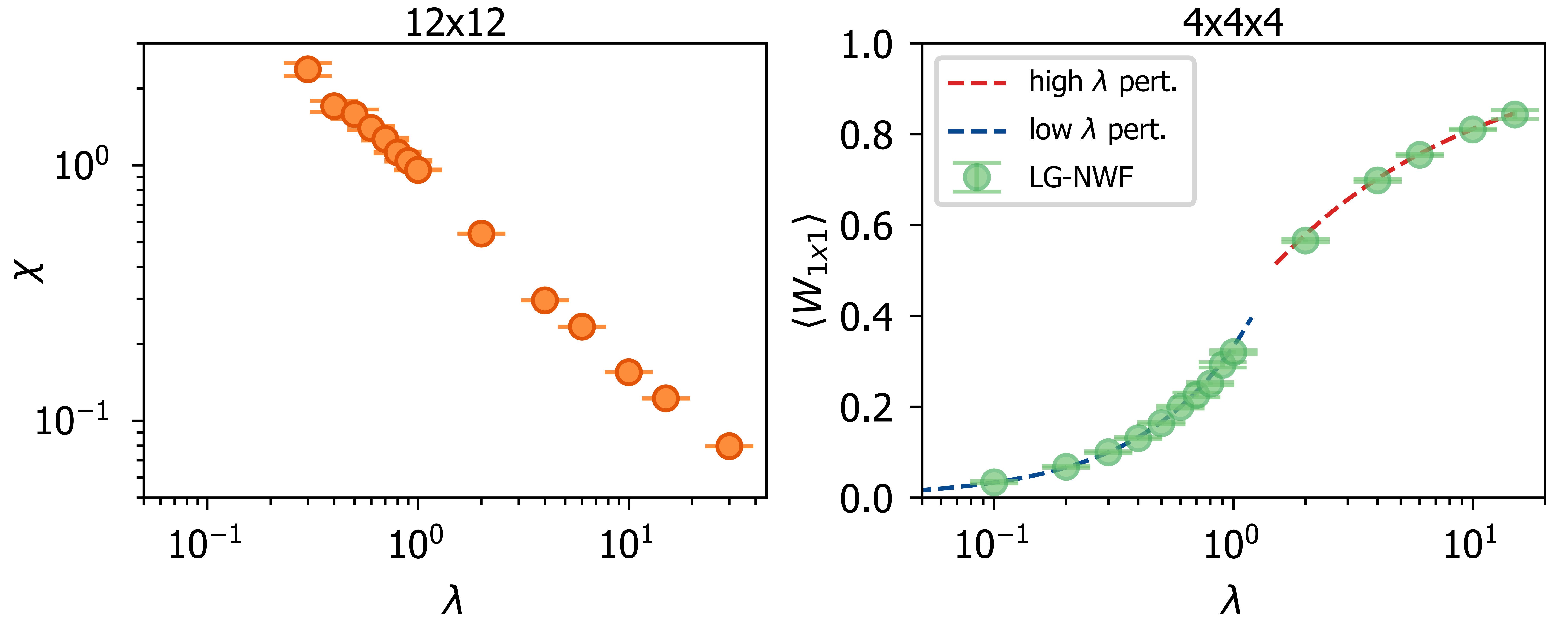
Simulation 2



Perturbative estimates from Chin, S. A. *et al.* (1985) *Phys Rev D* **31**, 3201

g range ~ 2.5 - 0.6

Simulation 2



Perturbative estimates from Chin, S. A. *et al.* (1985) *Phys Rev D* **31**, 3201

Ground state representation physically agrees with theoretical predictions

What now?

LGTs and neural wavefunctions

- What has been done:
  - Abelian theories in 2D
- What we have just done
  - Non-Abelian theories in 2 and 3D
- What can now be done
  - Time evolution[1] and quantum information entropies[2]
  - SU(3) without and then with fermions

[1] Carleo, G. Troyer, M. Science 355, Schmitt, M. & Heyl, M., Phys. Rev. Lett. **125** and many more…
[2] TS *et al. Mach. Learn. Sci. Tech.* **6** 015042, Sinibaldi, A. *et al. arXiv:2502.09725*

Eliska Greplova

TUDelft

ETH zürich

Juan Carrasquilla

Jannes Nys

Thank you for listening

Backup slides

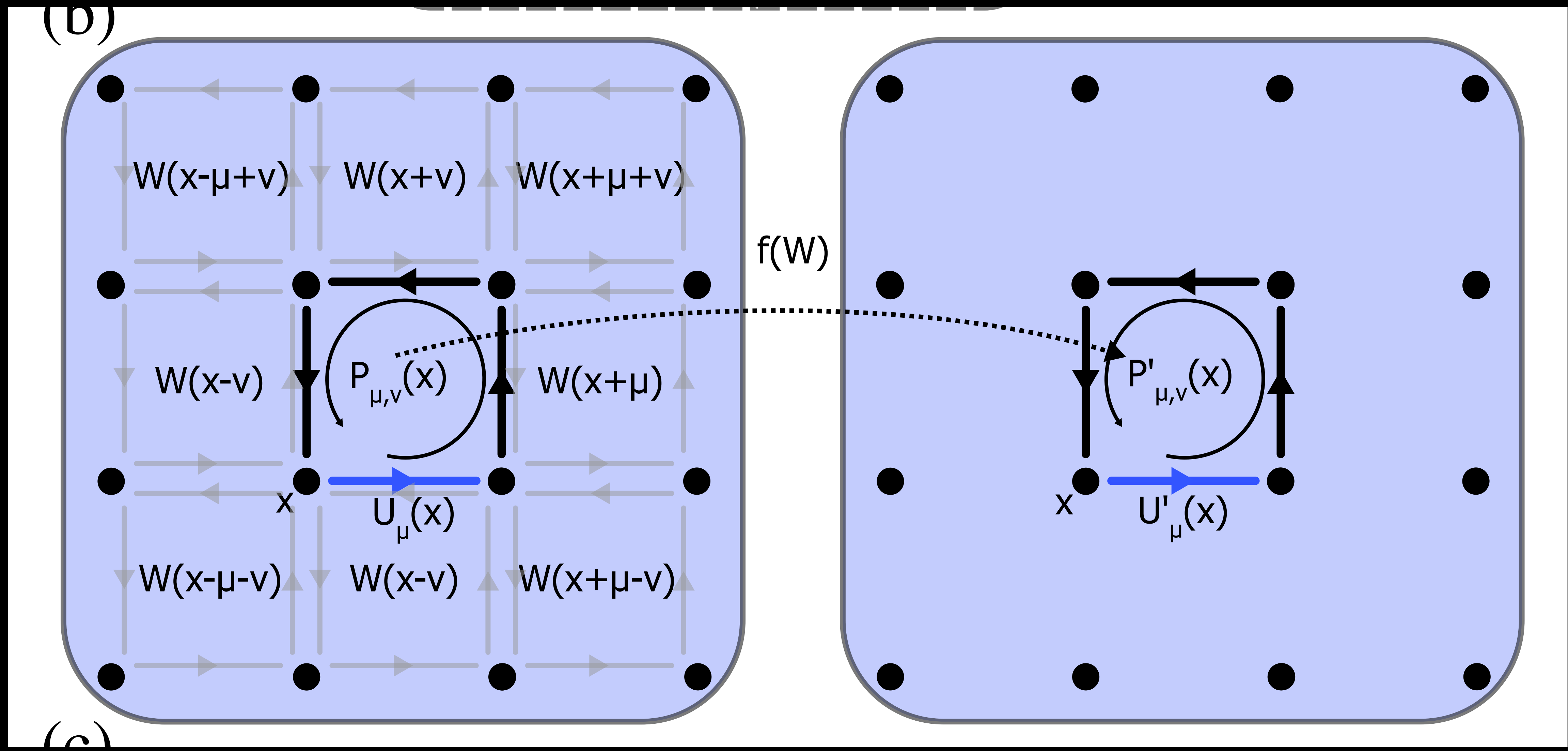
Plaquette equivariant layer

$$U'_\mu(x) = P'_{\mu,\nu}(x) * (U_\nu(x+\mu)U^\dagger_\mu(x+\nu)U^\dagger_\nu(x))^{-1}$$

$$U_\mu(x) \rightarrow P_{\mu,\nu}(x) \rightarrow \lambda_P \rightarrow \lambda'_P \rightarrow P'_{\mu,\nu}(x) \rightarrow U'_\mu(x)$$

CNN

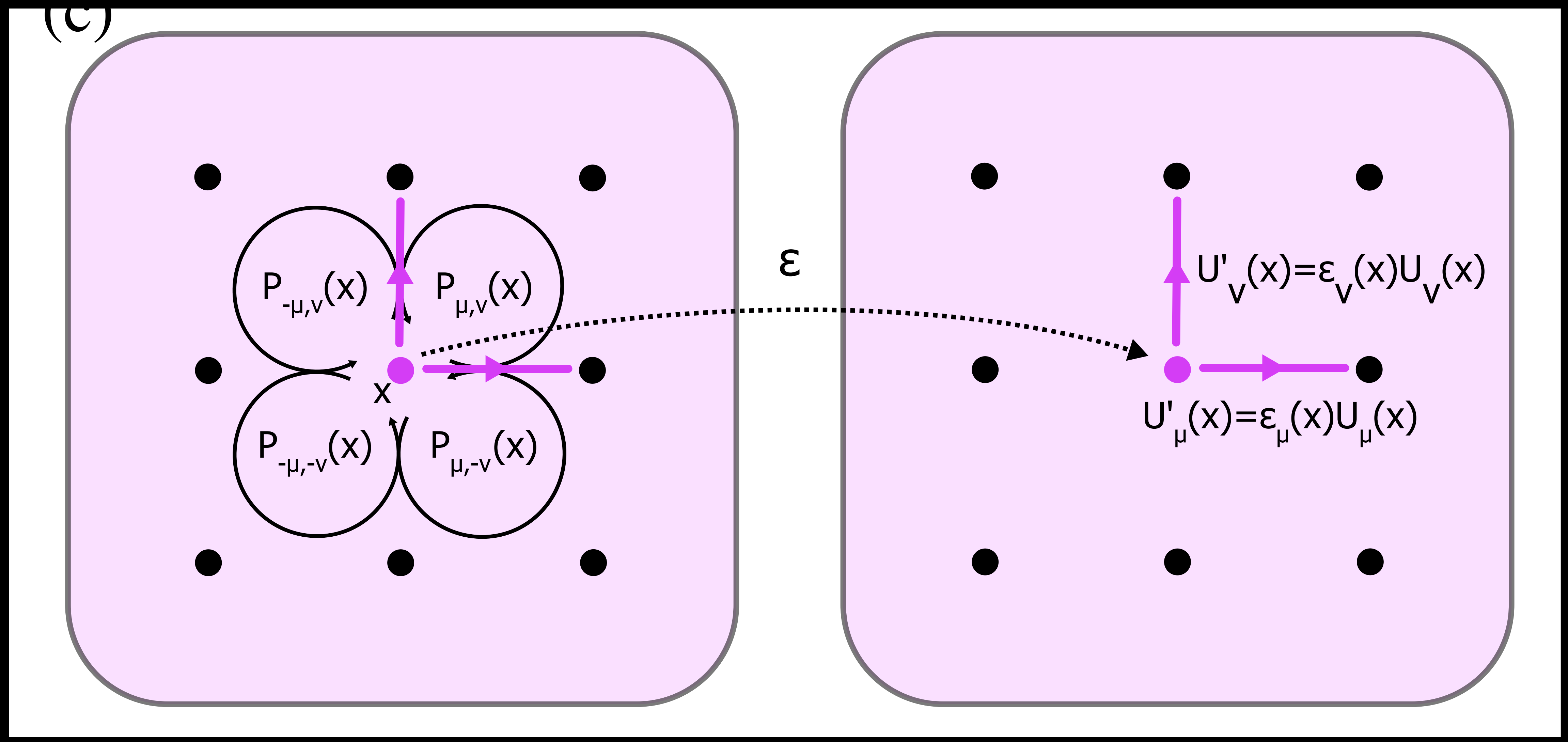
Eigenvalue decomposition



$$P_{\mu,\nu}(x) = U_\mu(x)U_\nu(x+\mu)U^\dagger_\mu(x+\nu)U^\dagger_\nu(x)$$

Eigenvalue recomposition

Node equivariant layer



$$U_\mu(x) \rightarrow C^i_{\mu,\nu}(x) \rightarrow \epsilon_{\mu,\nu}(x) \rightarrow \epsilon_{\mu,\nu}(x) U_\mu(x)$$

$$\epsilon_\mu(x) = e^{i \sum_i \gamma_{\mu,i} [C^i_{\mu,\nu}(x)]_{aH}}$$

$$C^i_{\mu,\nu}(x) = \{P_{\mu,\nu}(x), P_{\mu,-\nu}(x), P_{-\mu,\nu}(x), P_{-\mu,-\nu}(x)\}^i$$

Gauge symmetry

Our new SU(2) invariant ansatz

Equivariant

Invariant

$U_\mu(x)$ $f(U_\mu(x))$ $g(f(U_\mu(x)))$ $h(g(f(U_\mu(x))))$

Gauge symmetry

Our new SU(2) invariant ansatz

Equivariant

Invariant

$U_\mu(x)$ $f(U_\mu(x))$ $g(f(U_\mu(x)))$ $h(g(f(U_\mu(x))))$

$\Omega(x)U_\mu(x)$ $f(\Omega(x)U_\mu(x))$ $g(f(\Omega(x)U_\mu(x)))$ $h(g(f(\Omega(x)U_\mu(x))))$

Gauge symmetry

Our new SU(2) invariant ansatz

Equivariant

Invariant

$U_\mu(x)$ $f(U_\mu(x))$ $g(f(U_\mu(x)))$ $h(g(f(U_\mu(x))))$

$\Omega(x)U_\mu(x)$ $f(\Omega(x)U_\mu(x))$ $g(f(\Omega(x)U_\mu(x)))$ $h(g(f(\Omega(x)U_\mu(x))))$

$h(\Omega(x)g(f(U_\mu(x))))$