Neural network wavefunctions for lattice gauge theory

Based on arXiv:*2509.12323*

Thomas Spriggs, QuTech and Kavli Institute of Nanoscience Delft

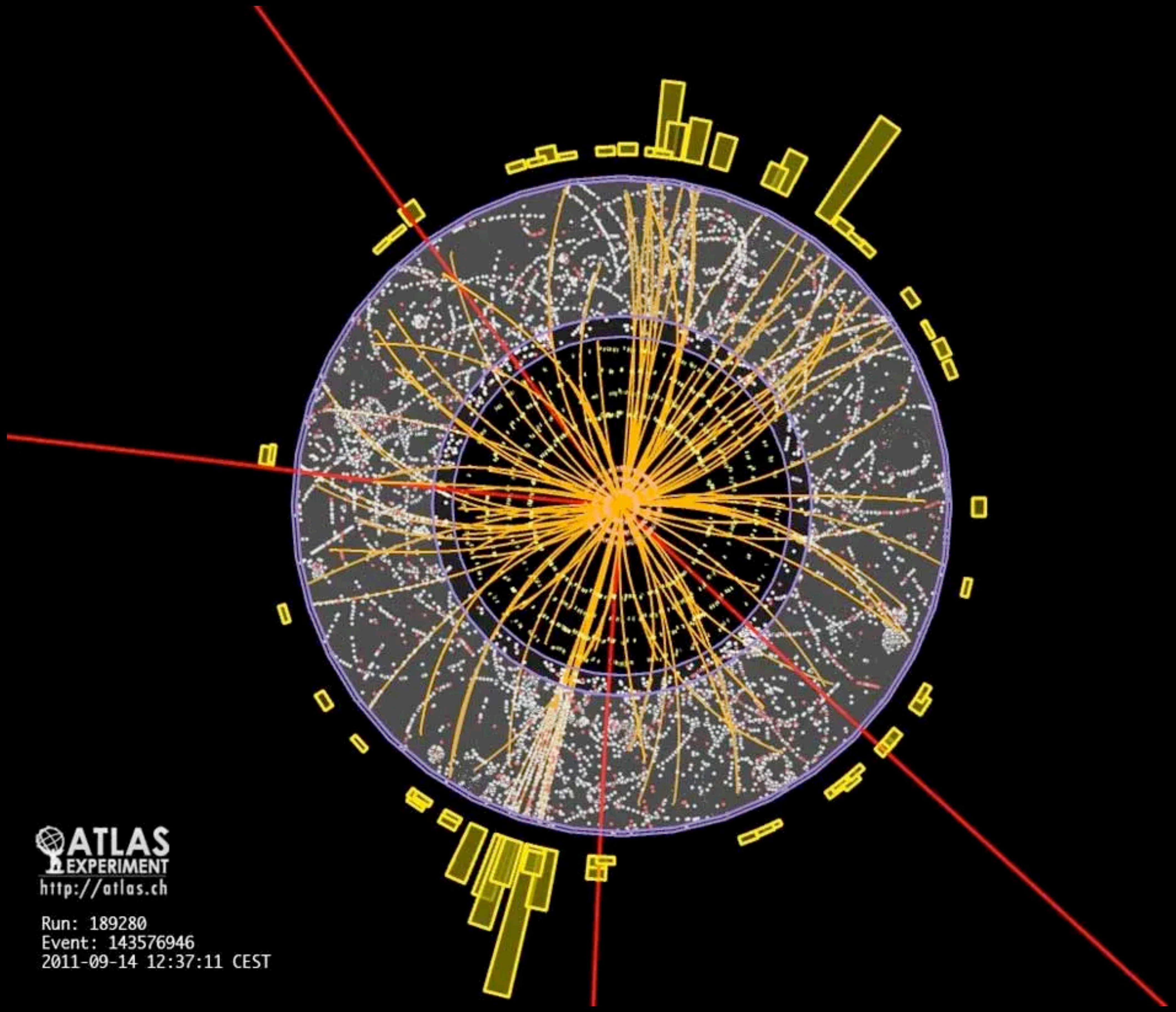
AI+HEP

In a nutshell

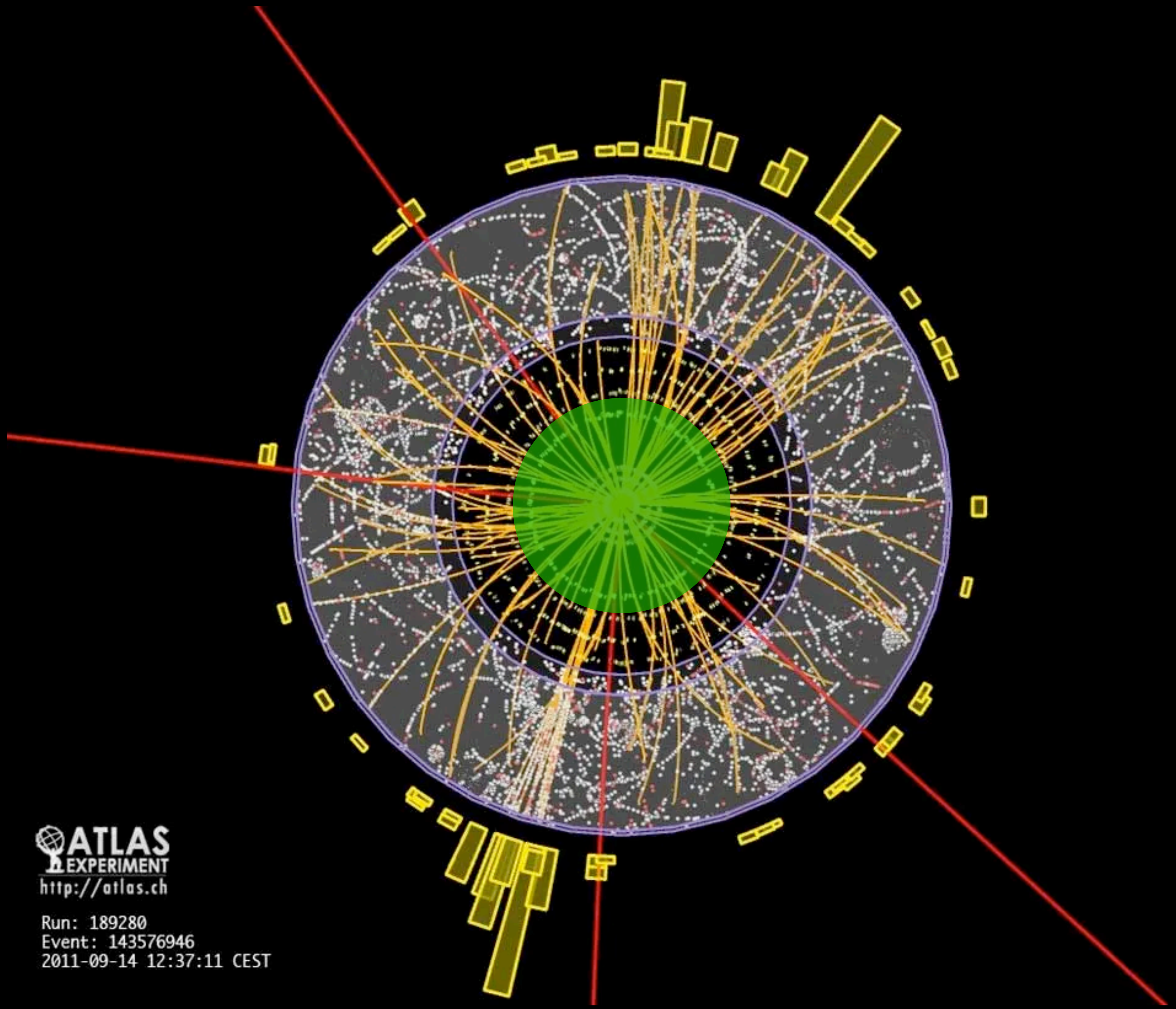
- New method for simulation non-Abelian lattice gauge theories in any* dimension
- I will show: ground state properties of SU(2) pure gauge
- Can be extended to: SU(N), time evolution, fermions

*GPU RAM dependent

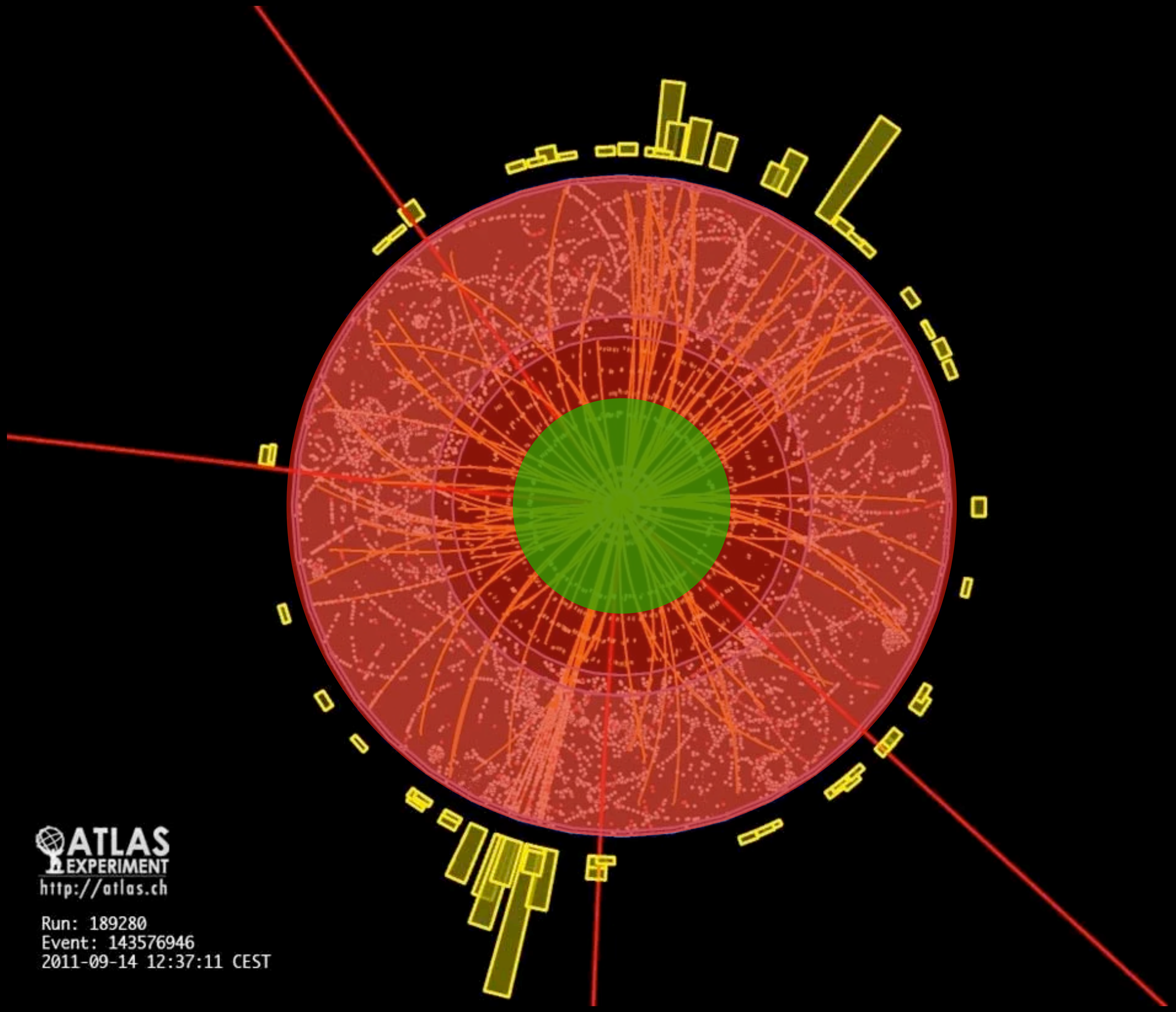
Why study the strong force?



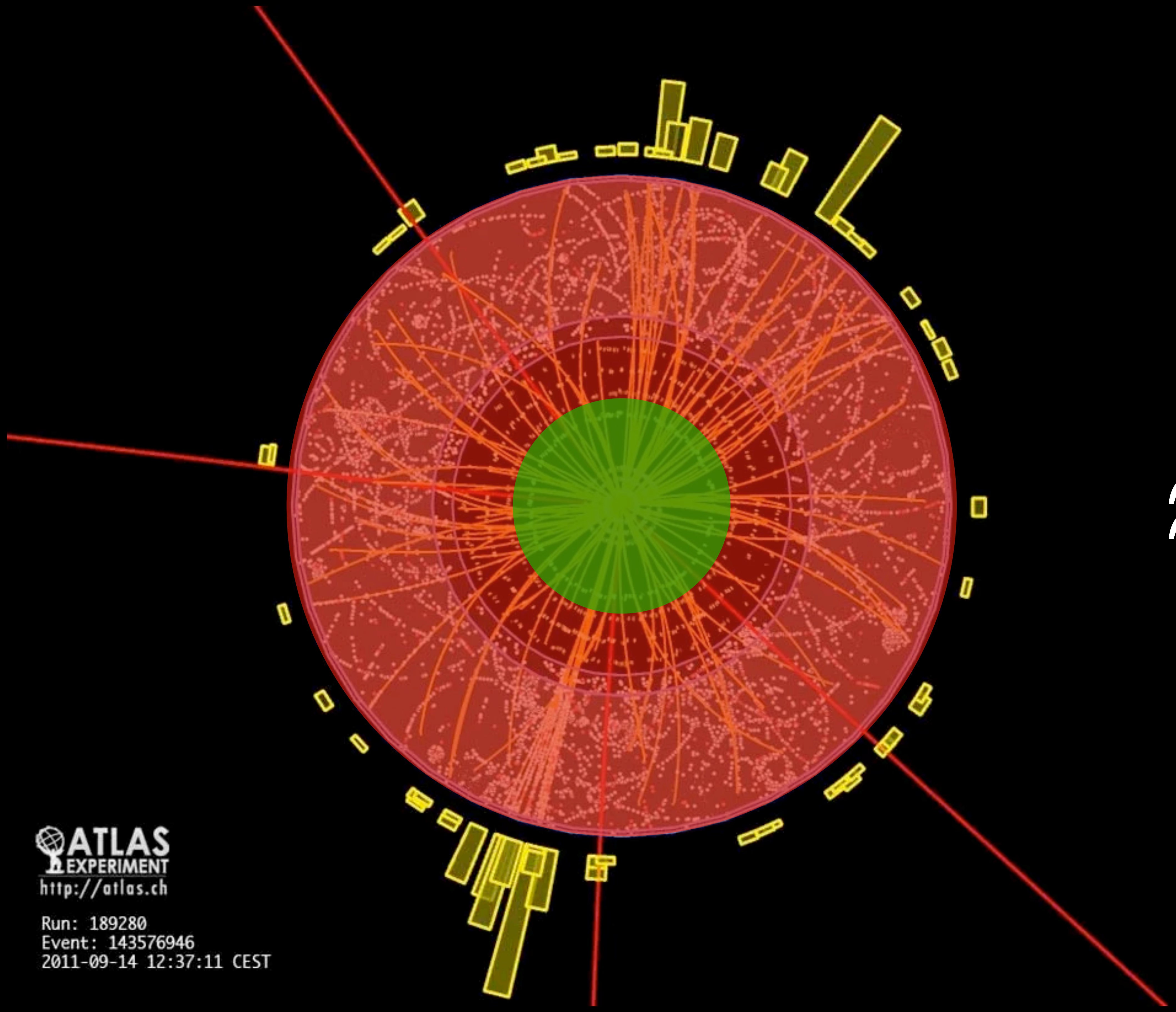
Why study the strong force?



Why study the strong force?



Why study the strong force?



$\approx$ Perturbation theory doesn't work

SU(2) pure gauge in the Hamiltonian formalism

Hamiltonian lattice gauge theories

- Continuum theory -> discrete spacial lattice
  - Regularises the theory
- Quantum many body problem
  - H - $n \times n$ matrix
  - Ψ - vector size n
  - Solve: $H|\Psi\rangle = E|\Psi\rangle$ for the lowest E

Hamiltonian lattice gauge theories

- Continuum theory -> discrete spacial lattice
  - Regularises the theory
- Quantum many body problem
  - H - $n \times n$ matrix
  - Ψ - vector size n
  - Solve: $H|\Psi\rangle = E|\Psi\rangle$ for the lowest E

Time is still real and continuous

Hamiltonian lattice gauge theories

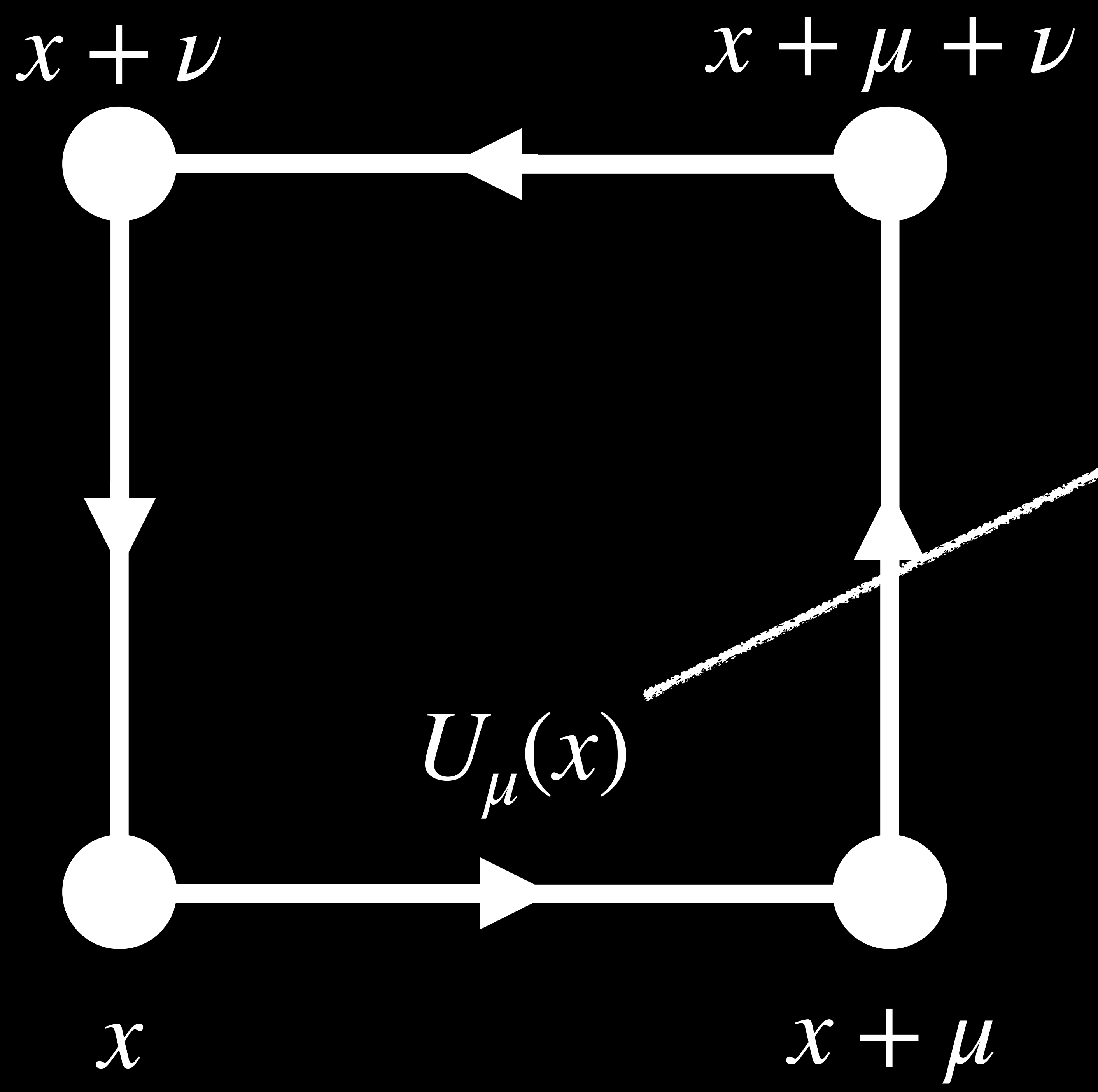
- Continuum theory -> discrete spacial lattice
  - Regularises the theory
- Quantum many body problem
  - H - $n \times n$ matrix
  - Ψ - vector size n
  - Solve: $H|\Psi\rangle = E|\Psi\rangle$ for the lowest E

Time is still real and continuous

n is huge, if not infinite

SU(2) Hamiltonian

Basic degrees of freedom

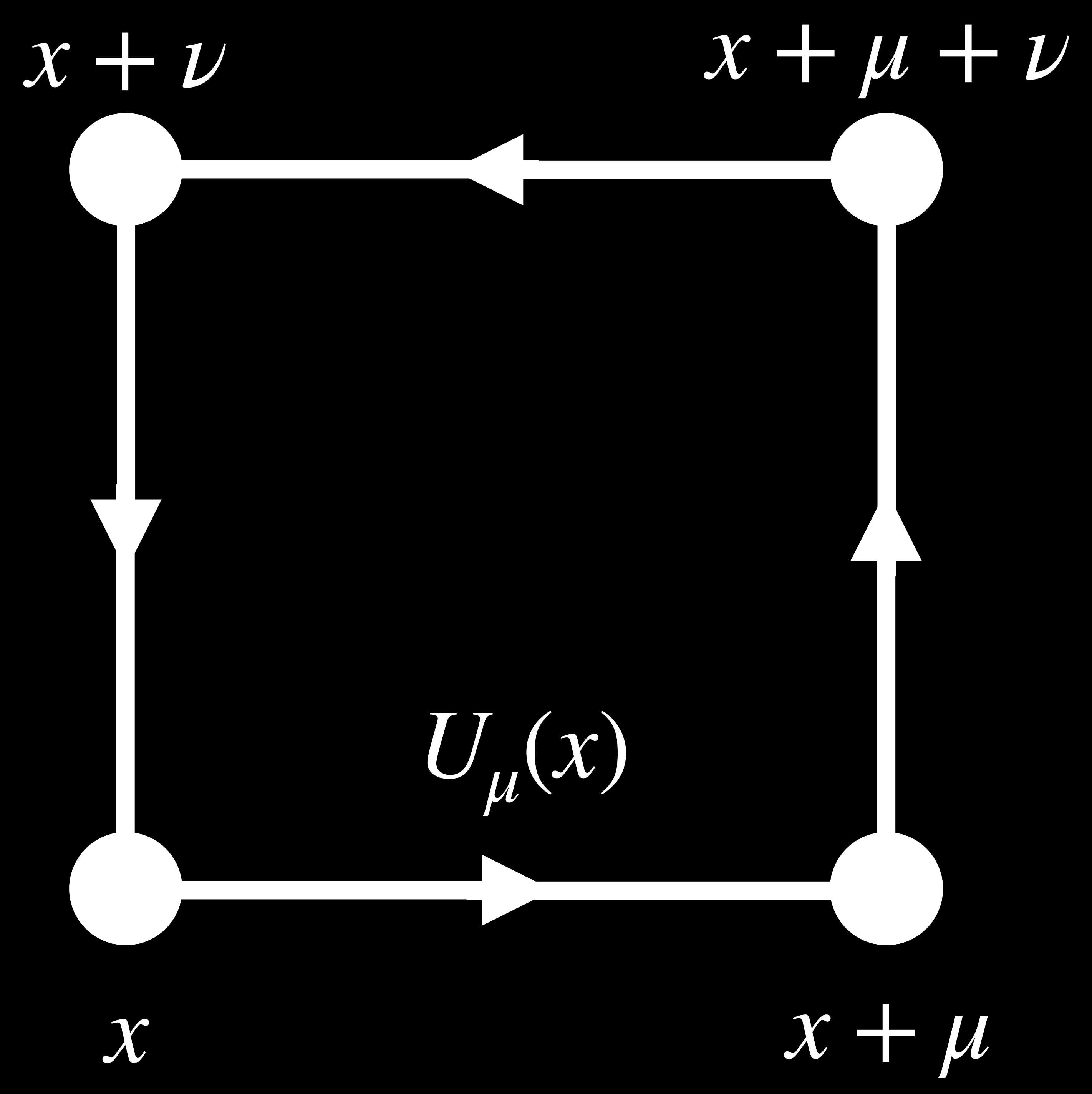


$U_\mu(x) \in SU(2)$

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



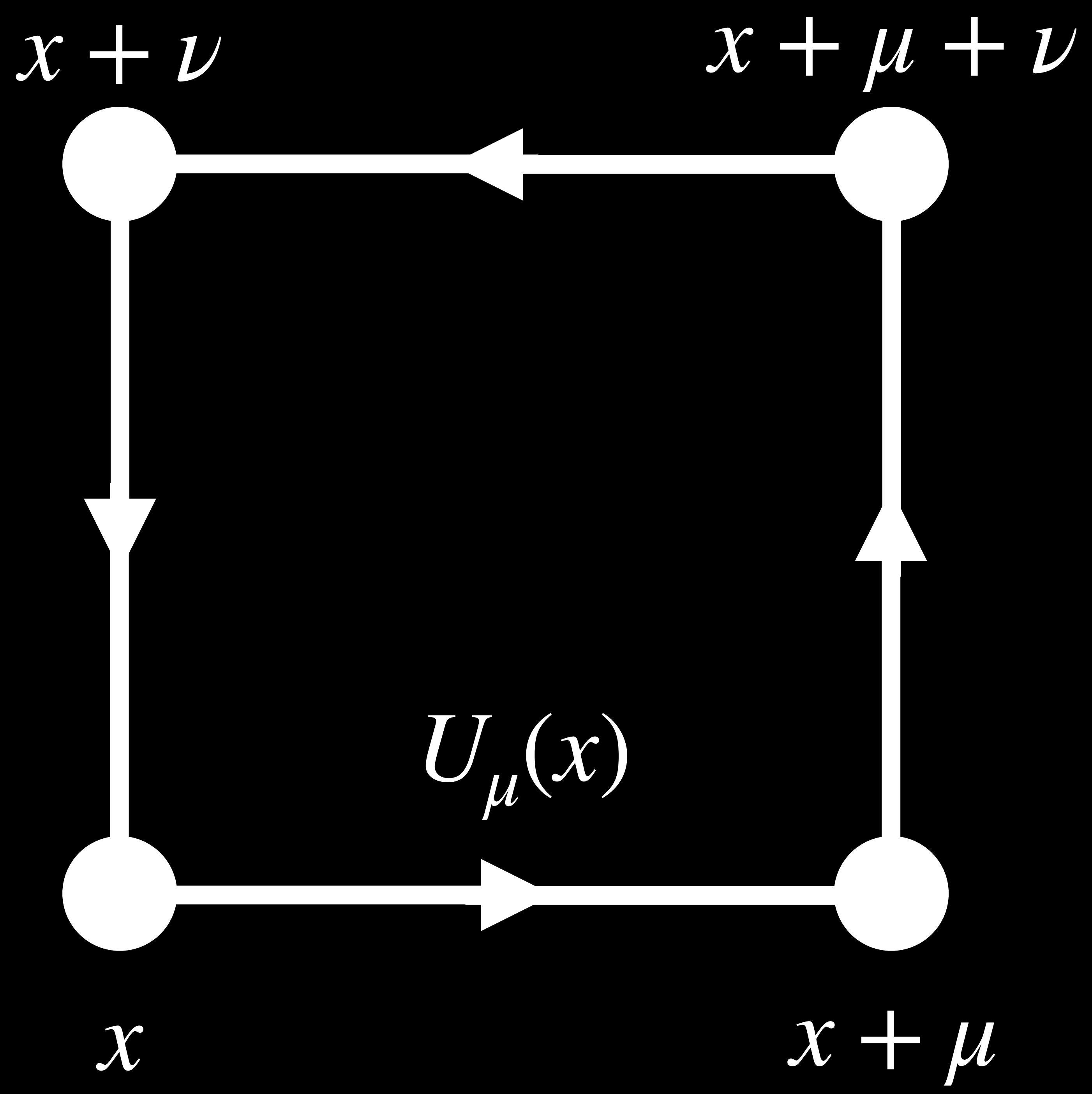
Hamiltonian:

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

kinetic $+\lambda$ potential

electric $+\lambda$ magnetic

SU(2) Hamiltonian



Hamiltonian:

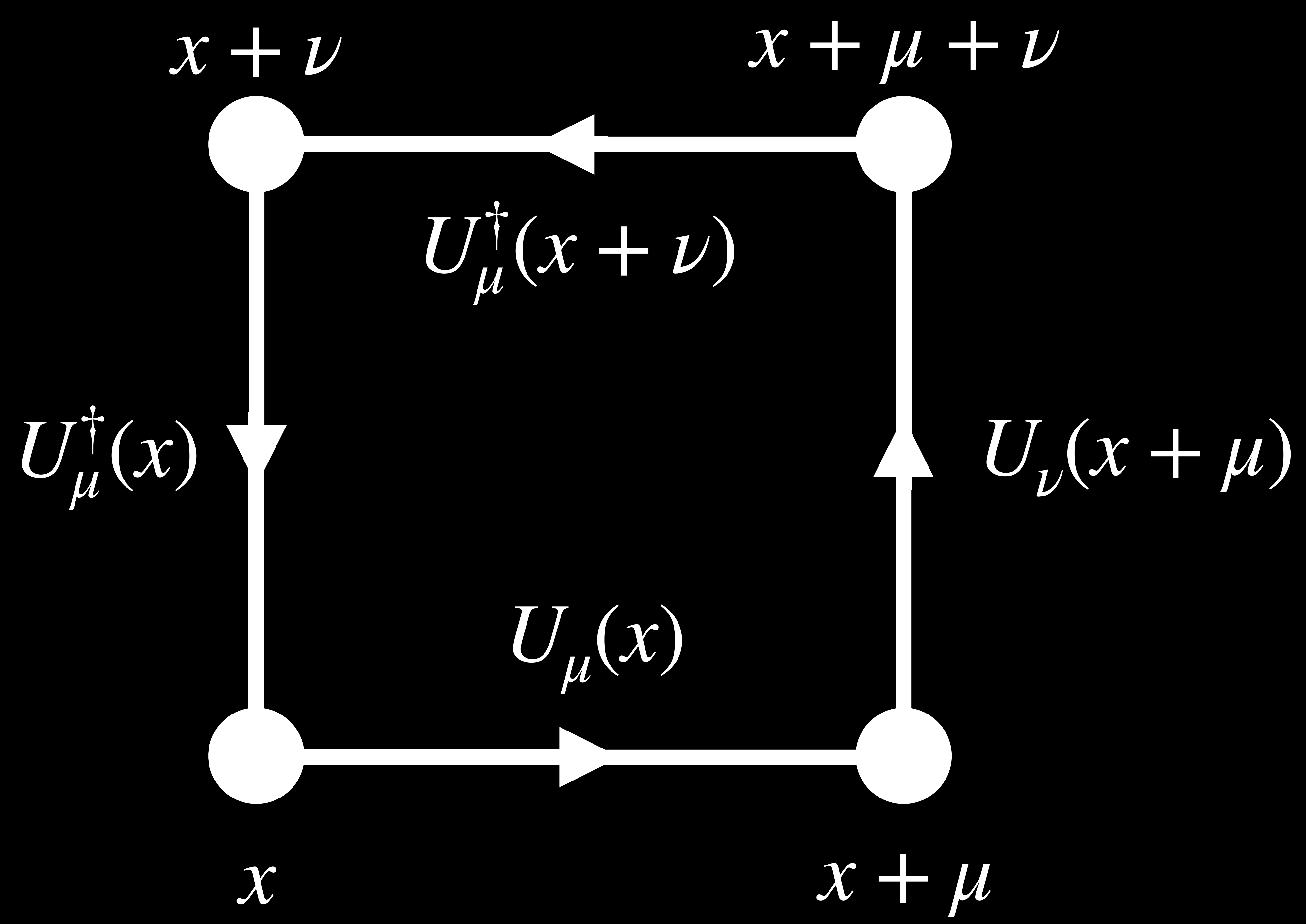
$$\mathcal{H} = -\frac{1}{2}\sum_l \nabla_l^2 + \lambda \sum_p \left(1 - \frac{1}{2}\mathrm{Tr}\left(P_p\right)\right)$$

kinetic $+\lambda$ potential

electric $+\lambda$ magnetic

$$\lambda = \frac{4}{g^4}$$

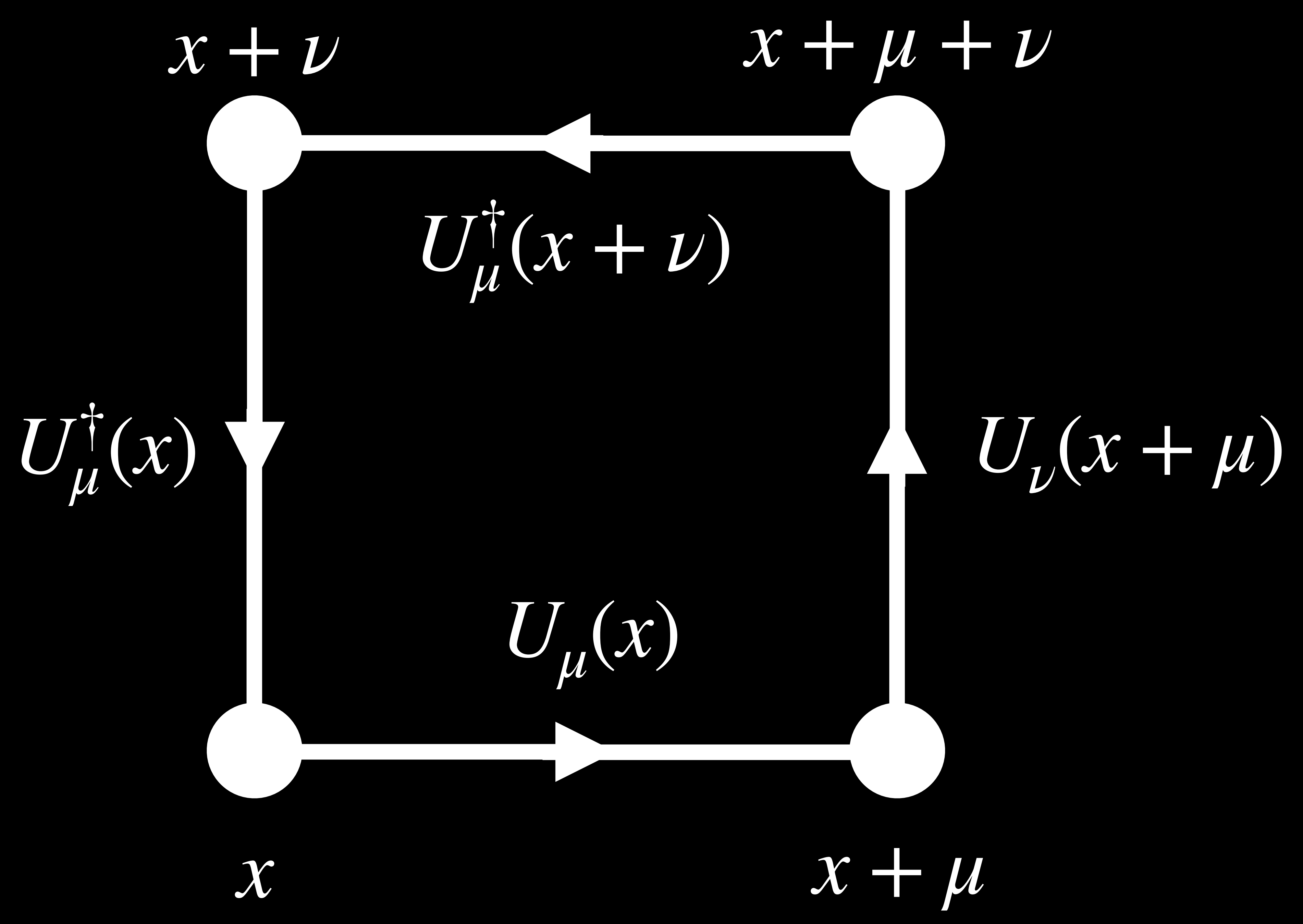
SU(2) Hamiltonian



Plaquette:

non-Abelian

SU(2) Hamiltonian

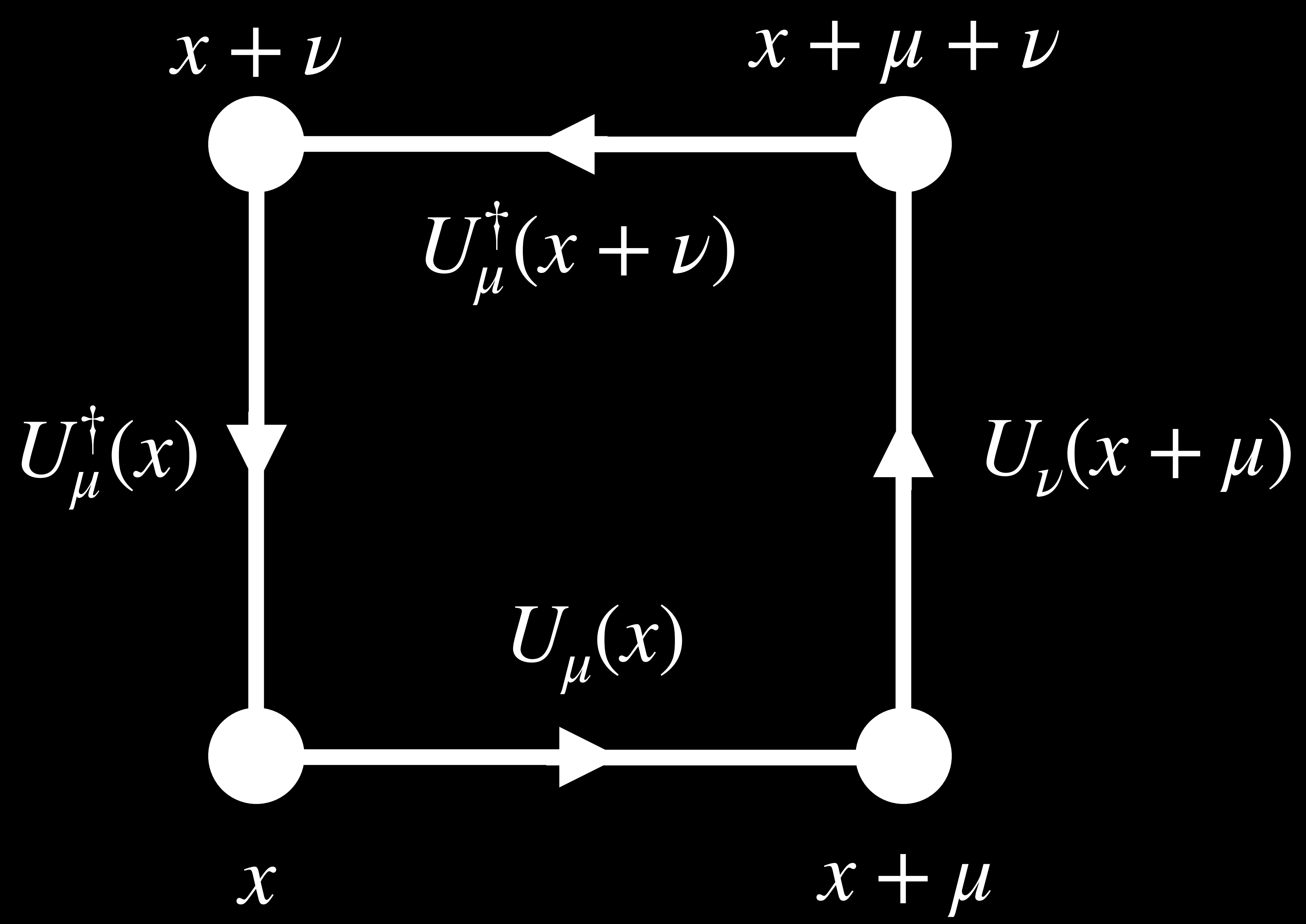


Plaquette:

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

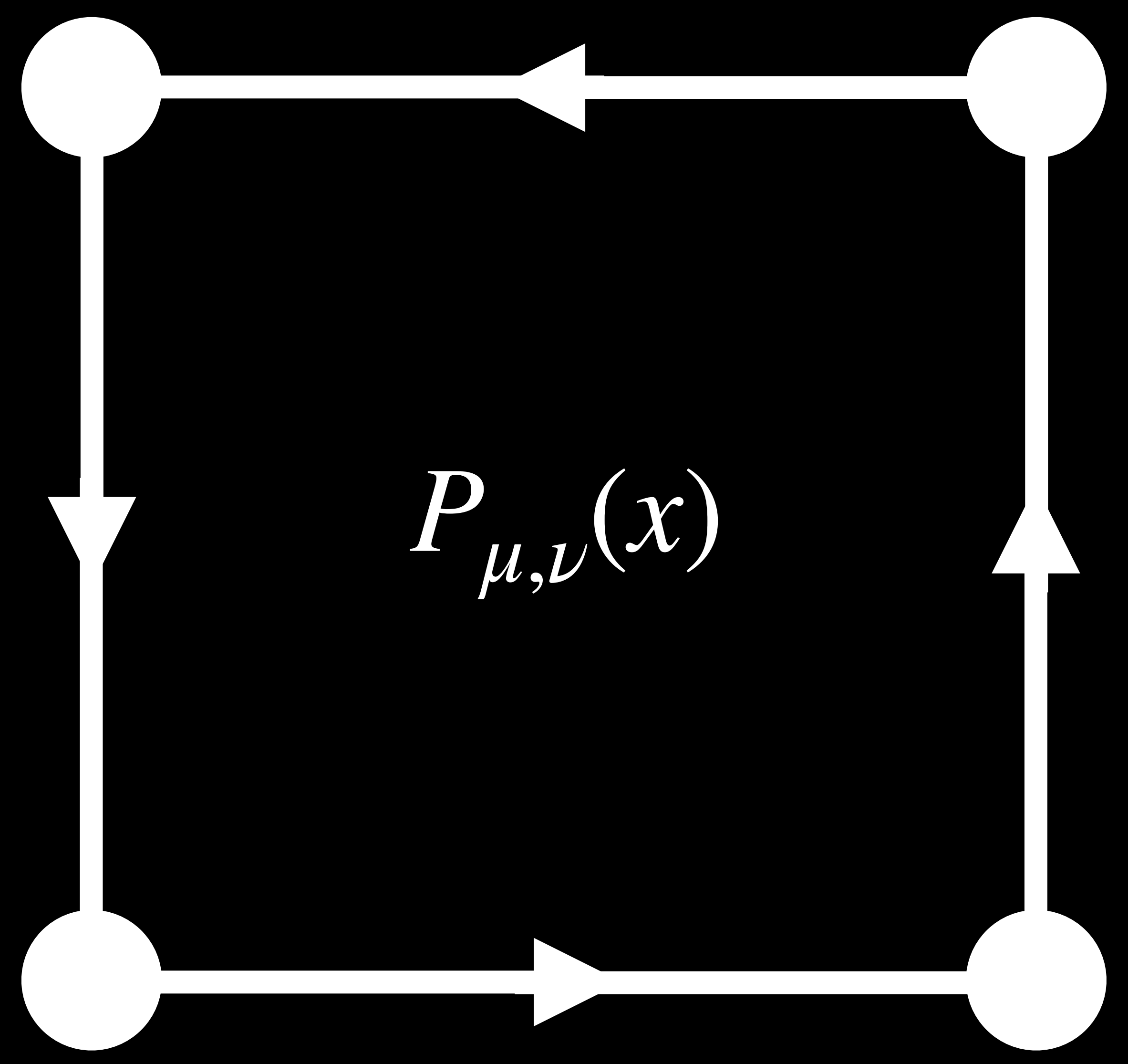
non-Abelian

SU(2) Hamiltonian



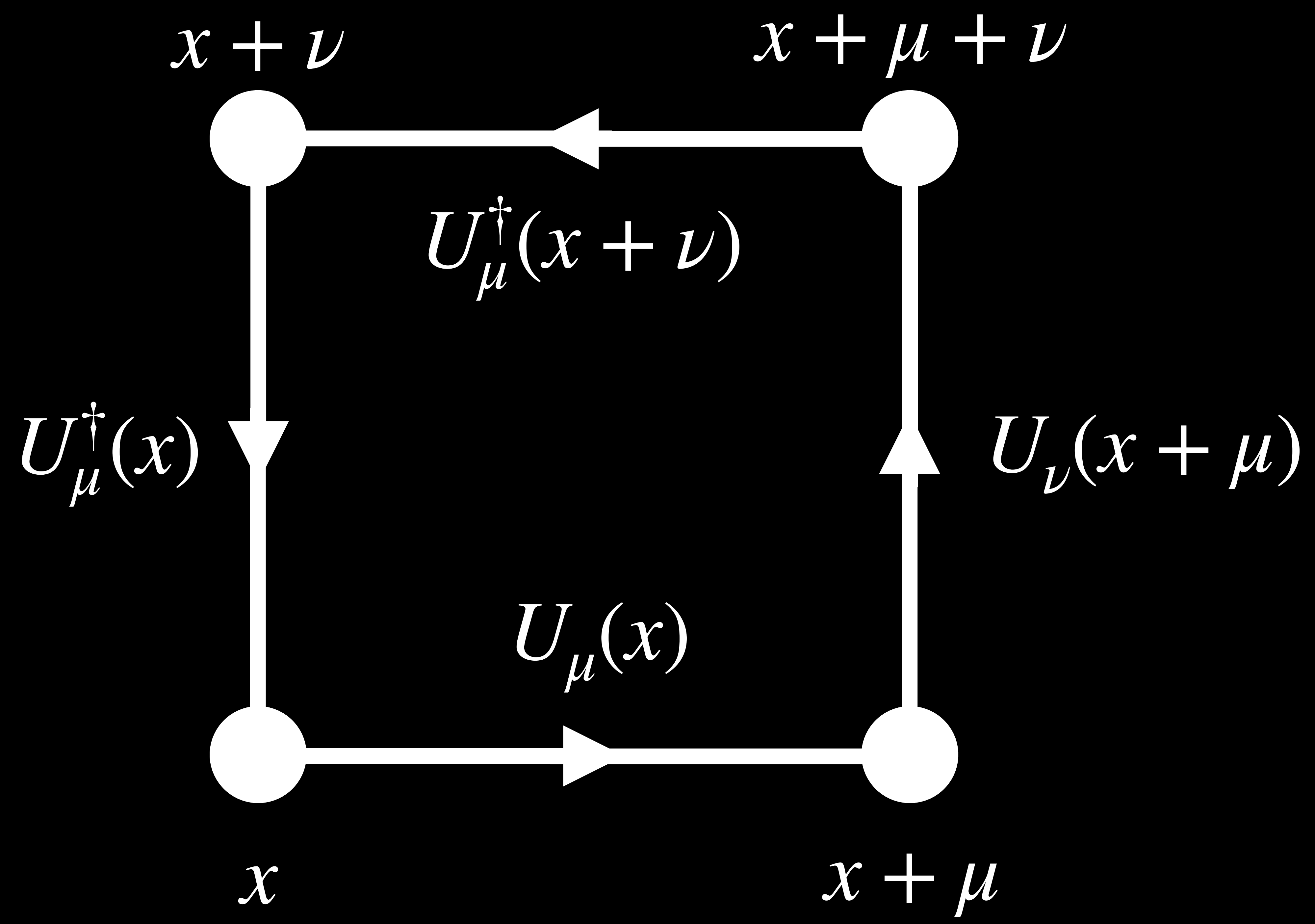
Plaquette:

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



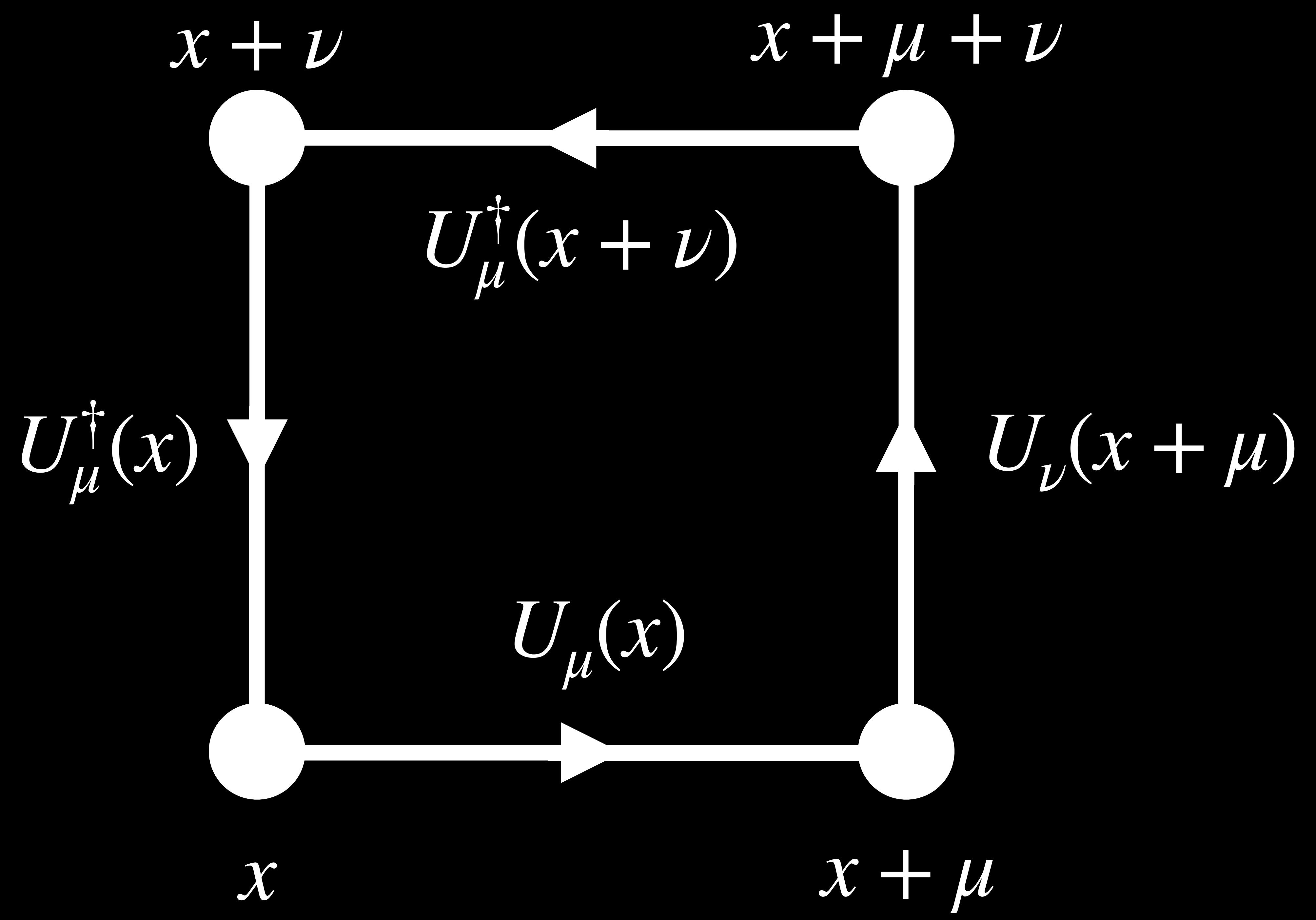
non-Abelian

SU(2) Hamiltonian



Derivative:

SU(2) Hamiltonian

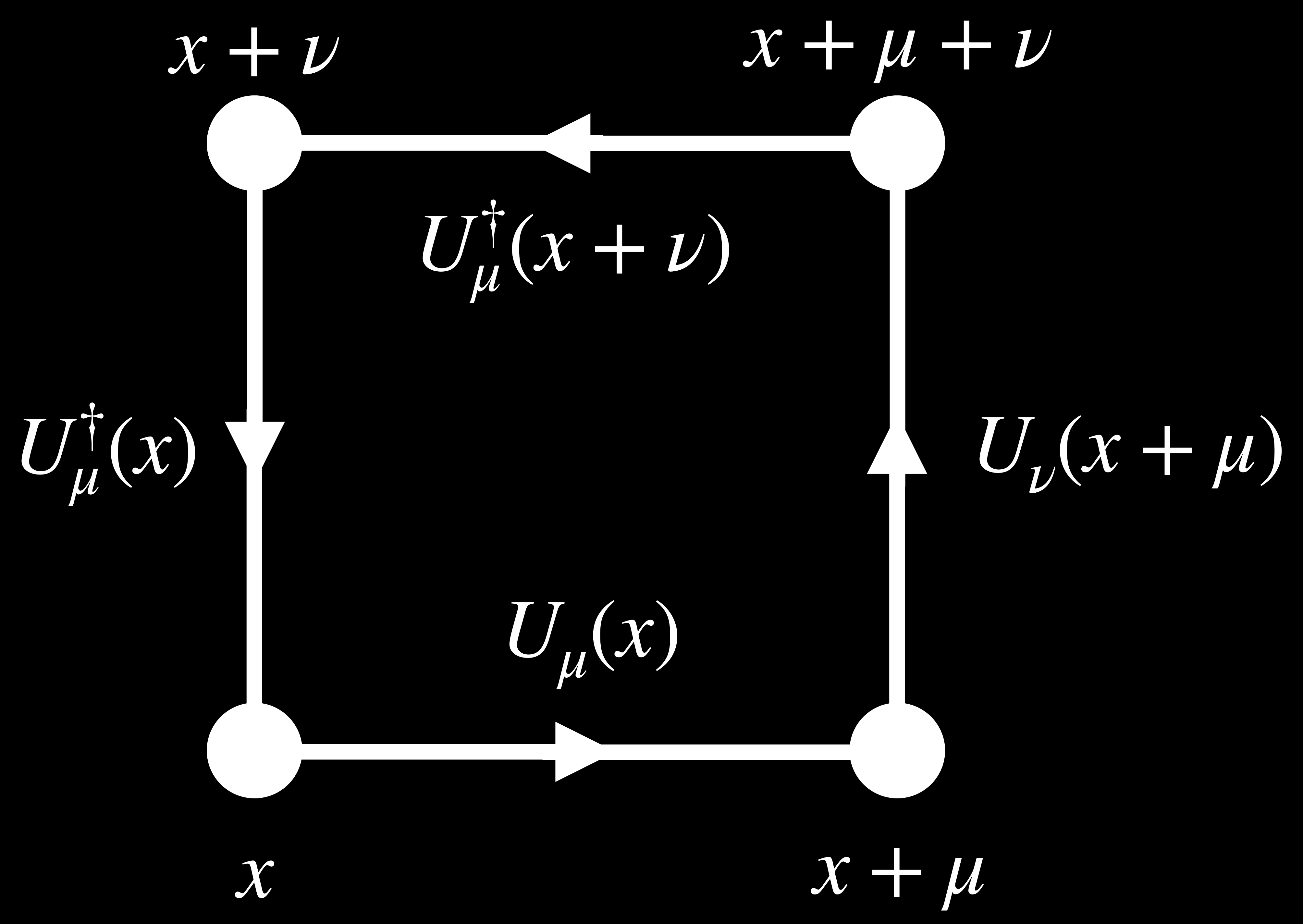


Derivative:

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

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

SU(2) Hamiltonian



Derivative:

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

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

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

Neural wavefunctions

Some definitions: wavefunction

Some definitions: wavefunction

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

Some definitions: wavefunction

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

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

Some definitions: variational wavefunction

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

Some definitions: variational wavefunction

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

$$\langle\psi_0\,|\,\Psi_\theta\rangle = \sum_i \langle\psi_0\,|\,f_\theta(\psi_i)\,|\,\psi_i\rangle = f_\theta(\psi_0)$$

Some definitions: variational wavefunction

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

$$\langle\psi_0|\Psi_\theta\rangle = \sum_i \langle\psi_0|f_\theta(\psi_i)|\psi_i\rangle = f_\theta(\psi_0)$$

"Configuration in, amplitude out"

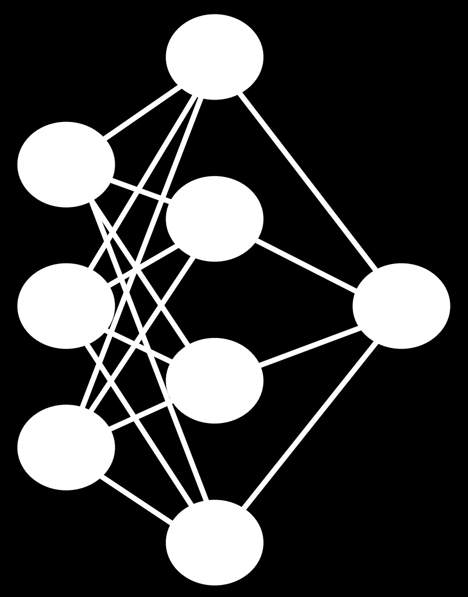
Some definitions: variational wavefunction

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

$$\langle \psi_0 | \Psi_\theta \rangle = \sum_i \langle \psi_0 | f_\theta(\psi_i) | \psi_i \rangle = f_\theta(\psi_0)$$

“Configuration in, amplitude out”

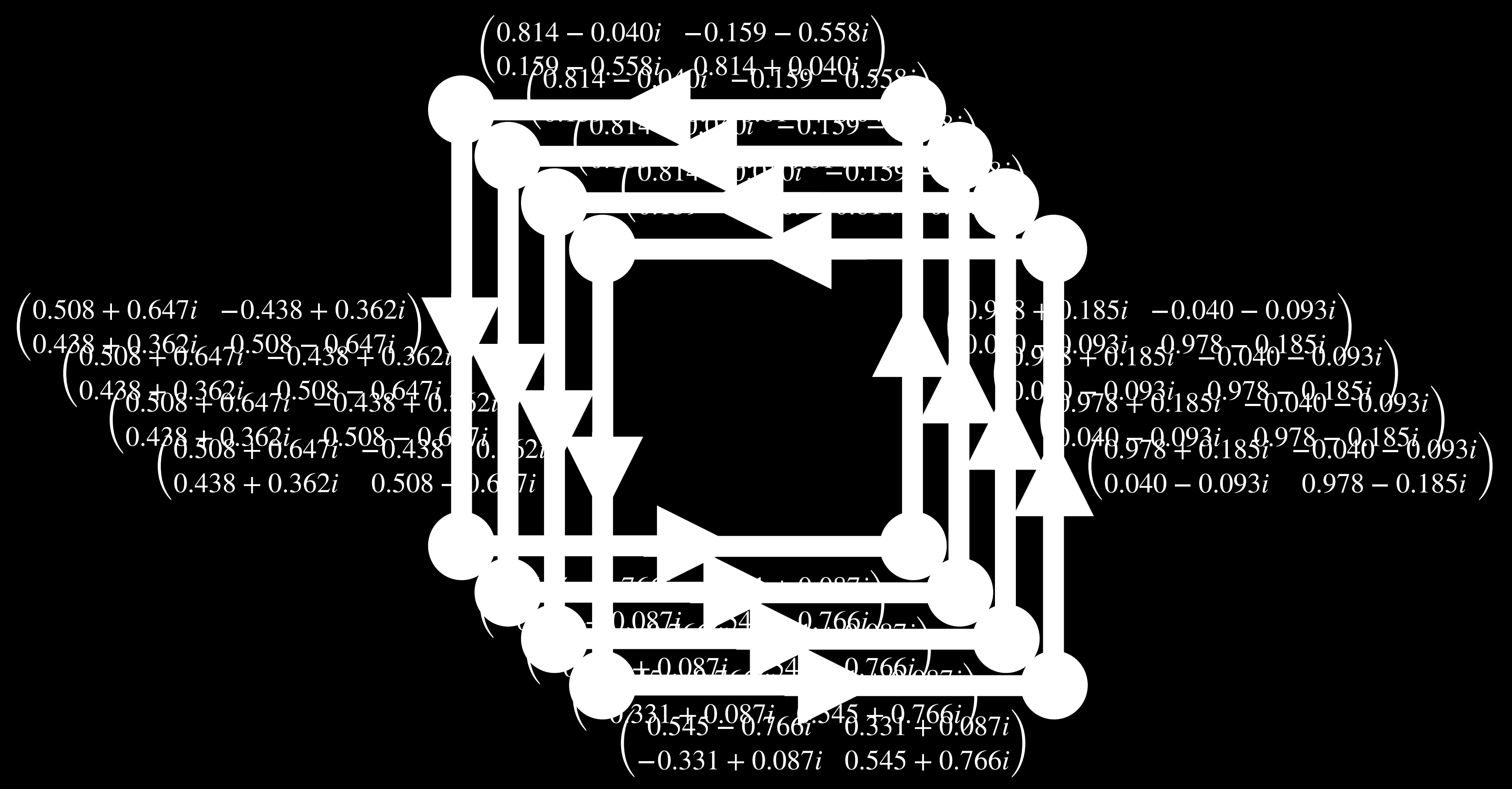
$\langle \psi_0 |$ $\longrightarrow$



$\longrightarrow$ $f_\theta(\psi_0)$

Neural wavefunctions for LGTs

$$\{\mathrm{U}_i\} \sim |\Psi_\theta\rangle$$

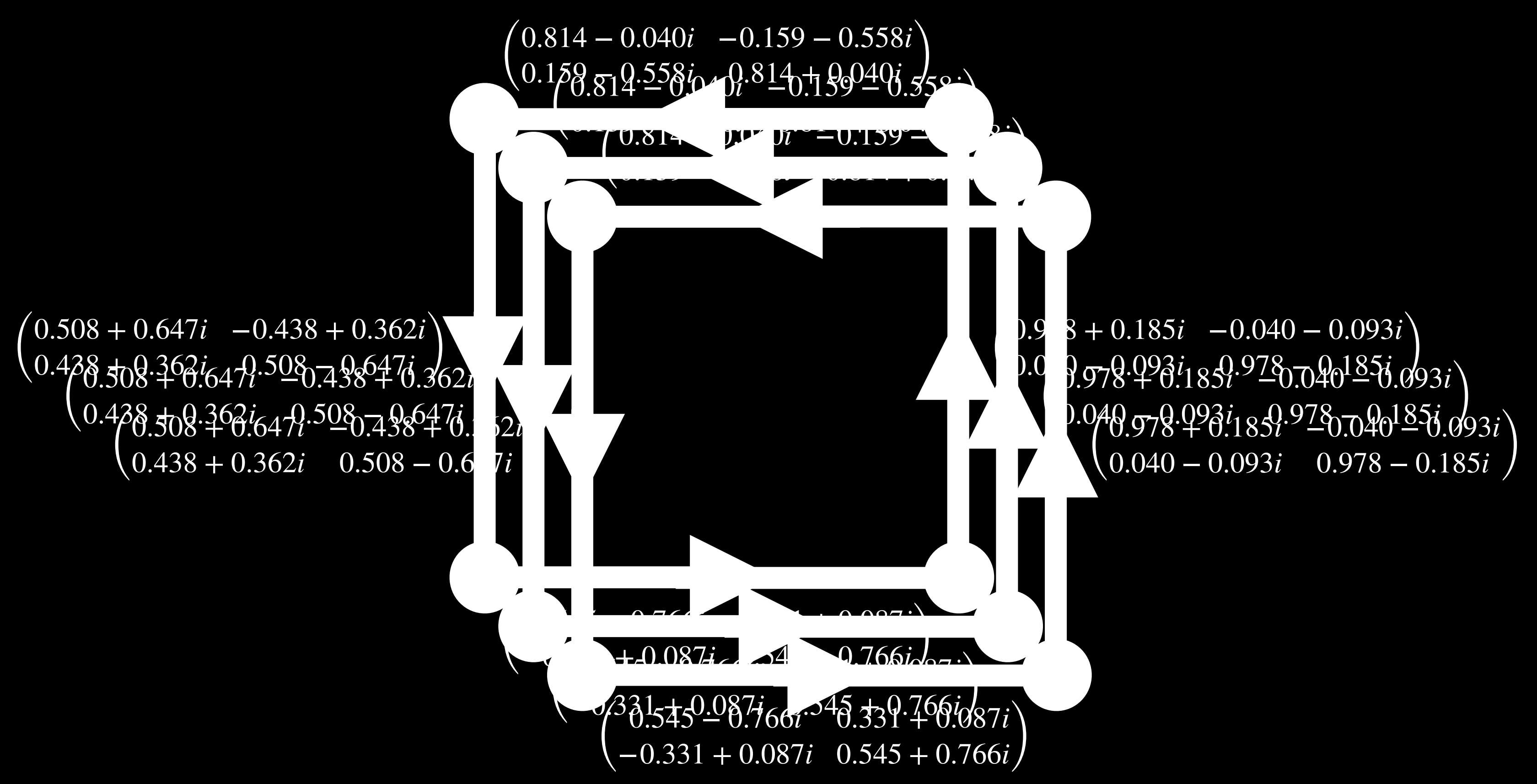


NN

θ

Neural wavefunctions for LGTs

$\{\mathrm{U}_i\} \sim |\Psi_\theta\rangle$



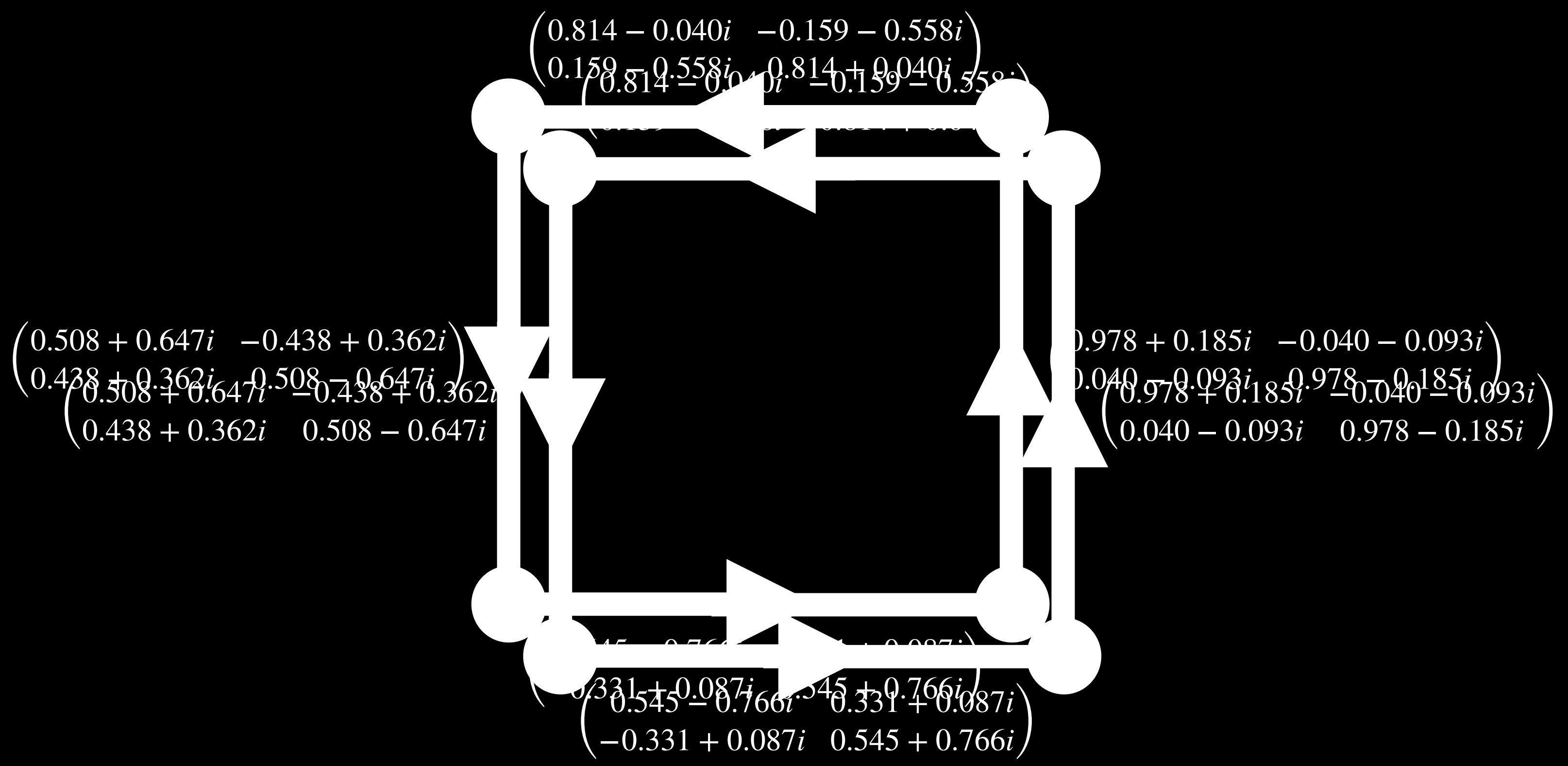
NN

θ

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

Neural wavefunctions for LGTs

$\{\mathrm{U}_i\} \sim |\Psi_\theta\rangle$



NN
θ

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

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

Neural wavefunctions for LGTs

NN
θ

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

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

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

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

Neural wavefunctions for LGTs

NN
θ

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

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

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

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

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

Neural wavefunctions for LGTs

NN
θ

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

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

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

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

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

Neural wavefunctions for LGTs

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

NN
θ

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

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

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

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

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

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

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

Neural wavefunctions 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

Neural wavefunctions for LGTs

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

NN
θ

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

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

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

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

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

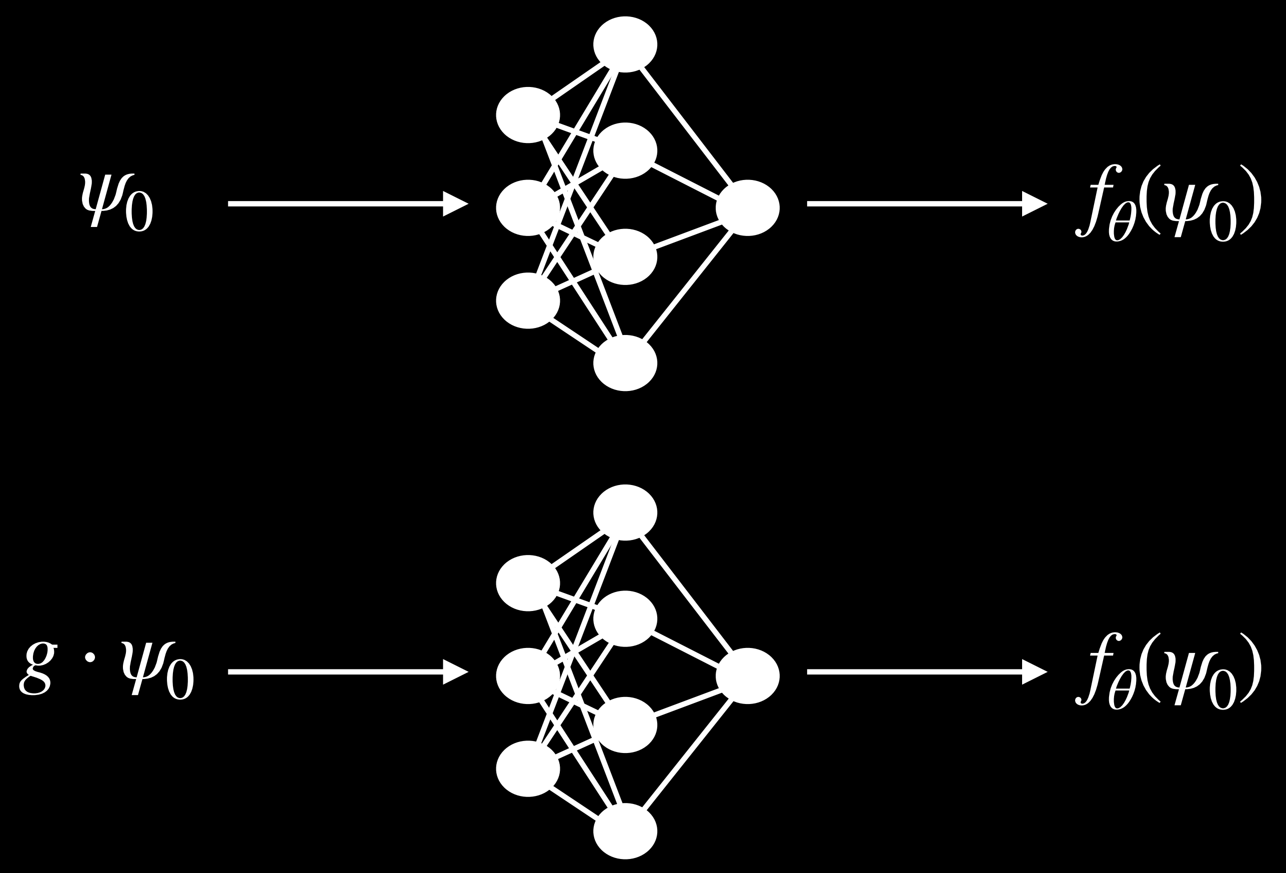
JAX

MCMC

Gauge invariance

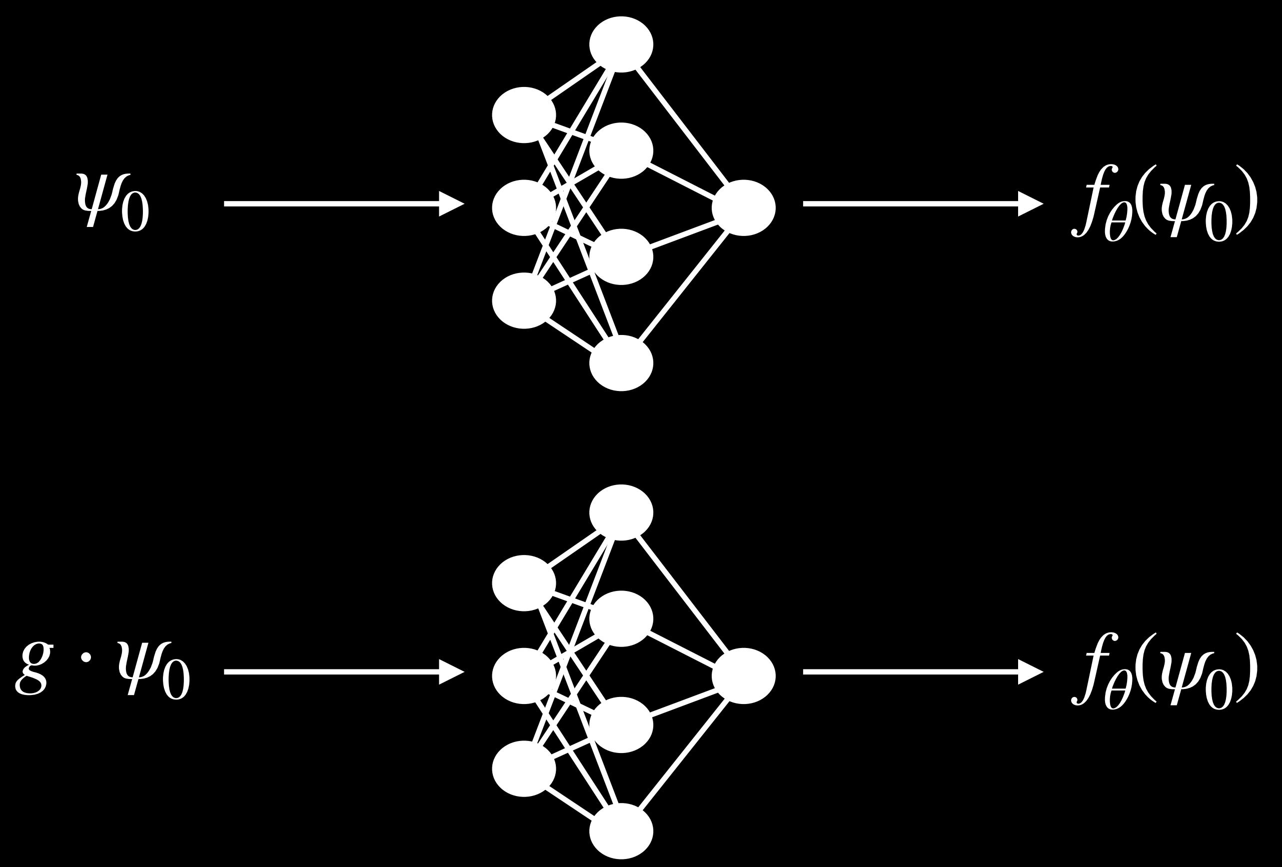
Gauge invariance

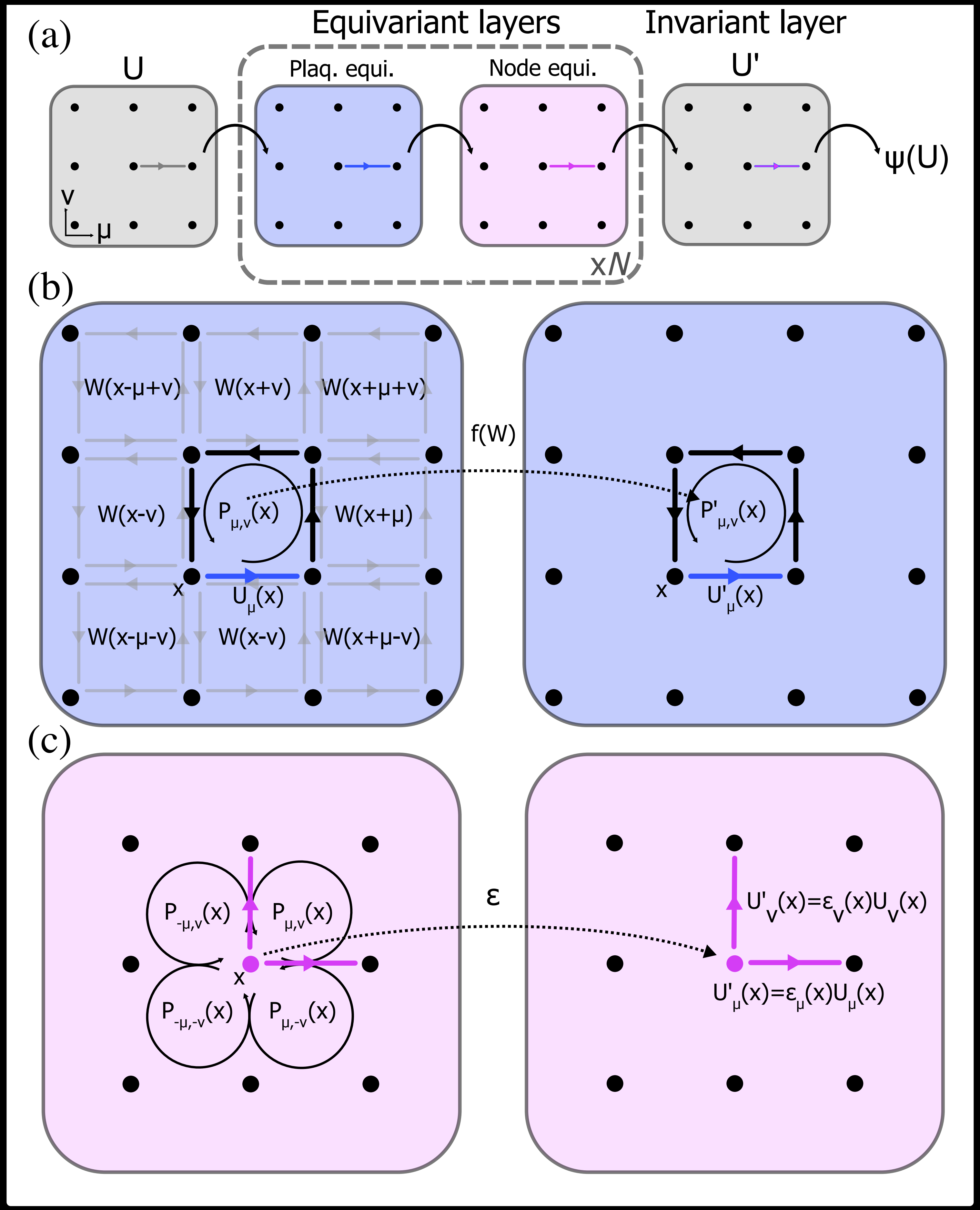
“Configuration in, amplitude out”



Gauge invariance

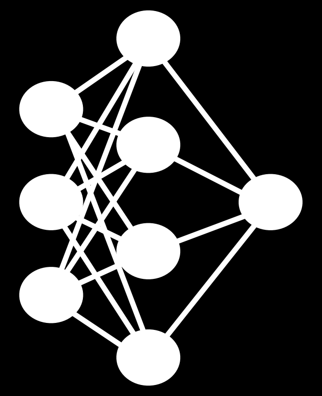
“Configuration in, amplitude out”

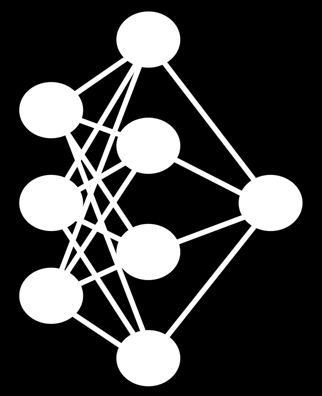


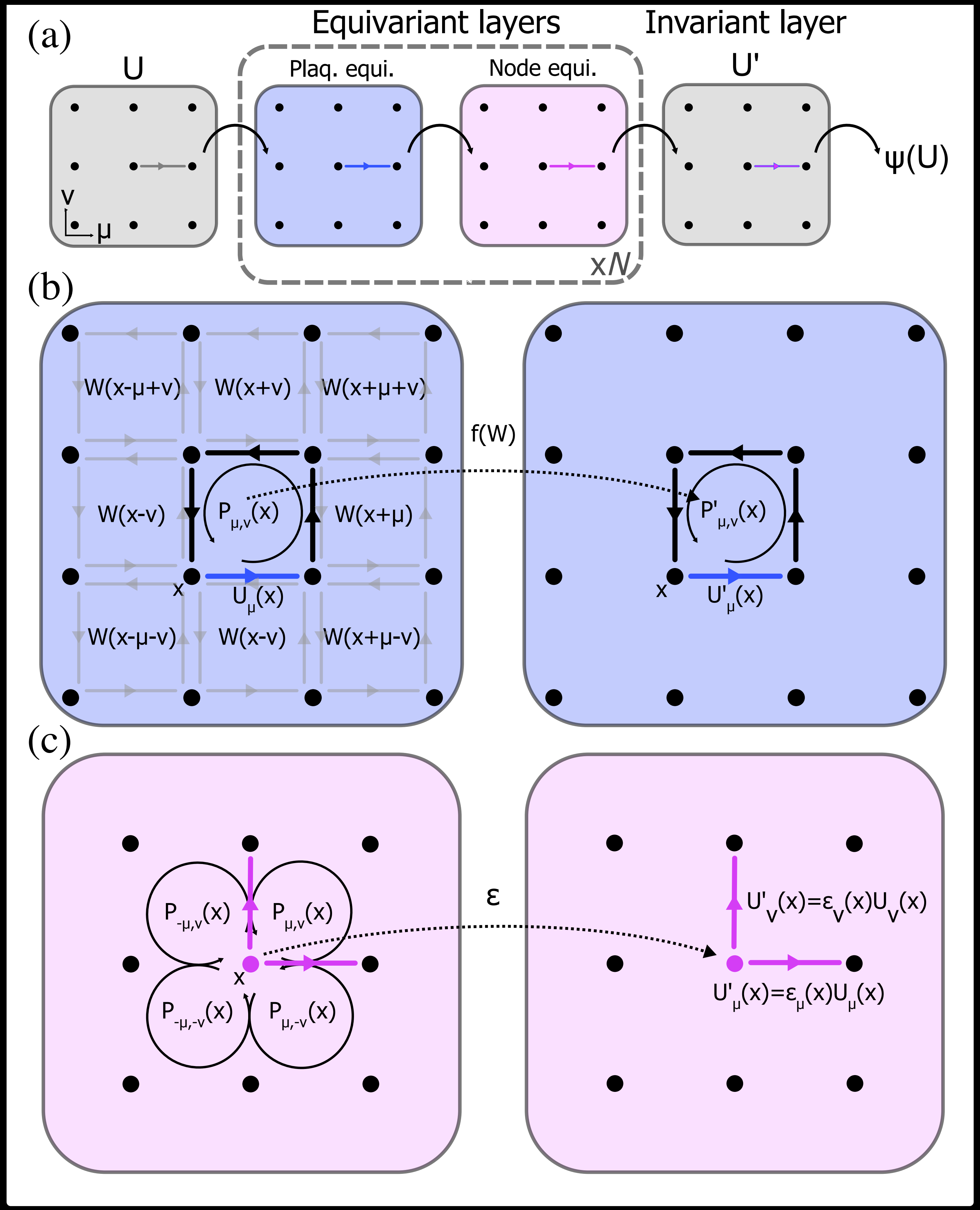


Gauge invariance

“Configuration in, amplitude out”

$\psi_0 \longrightarrow$  $\longrightarrow f_\theta(\psi_0)$

$g \cdot \psi_0 \longrightarrow$  $\longrightarrow f_\theta(\psi_0)$



Continuous basis

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

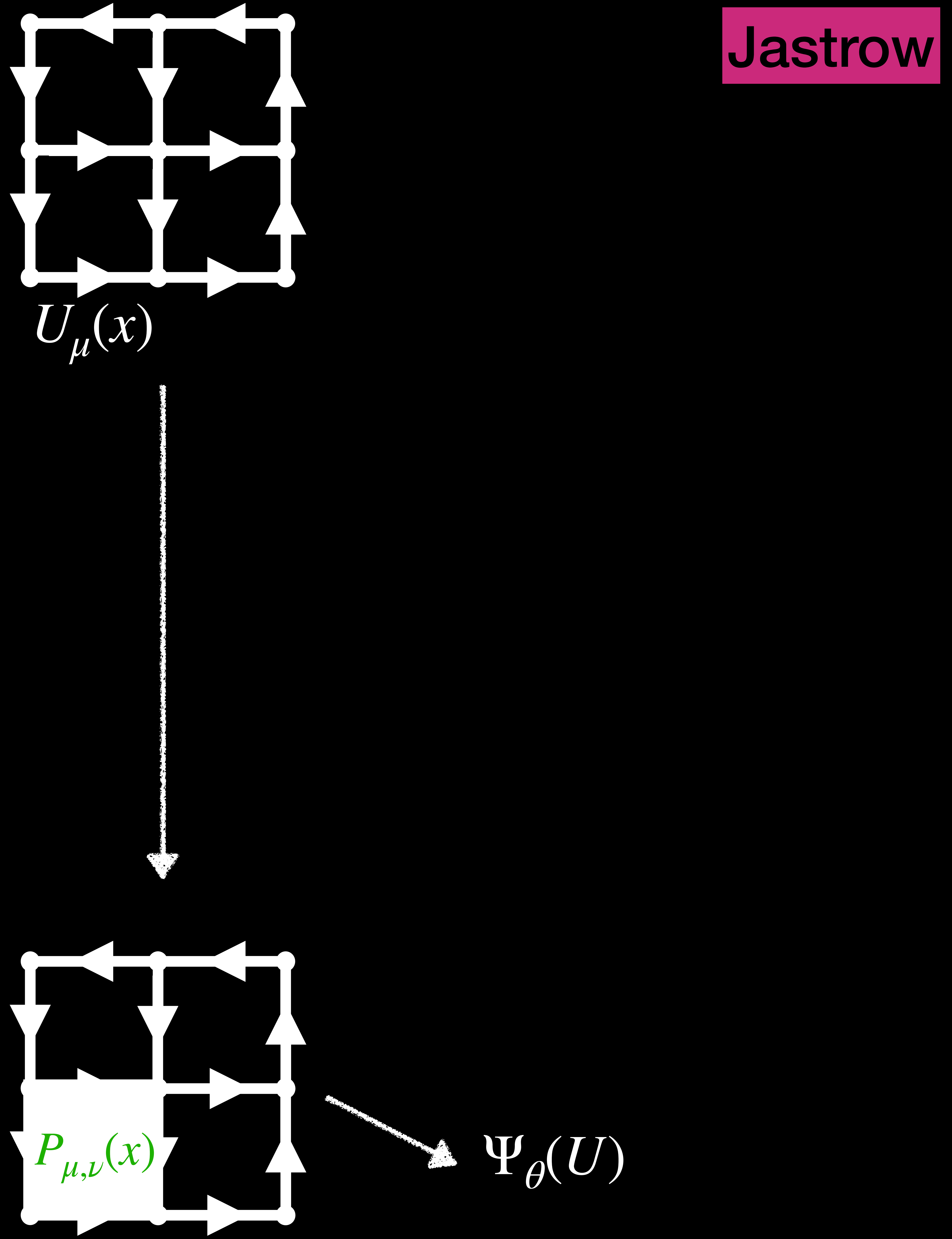
Gauge invariant

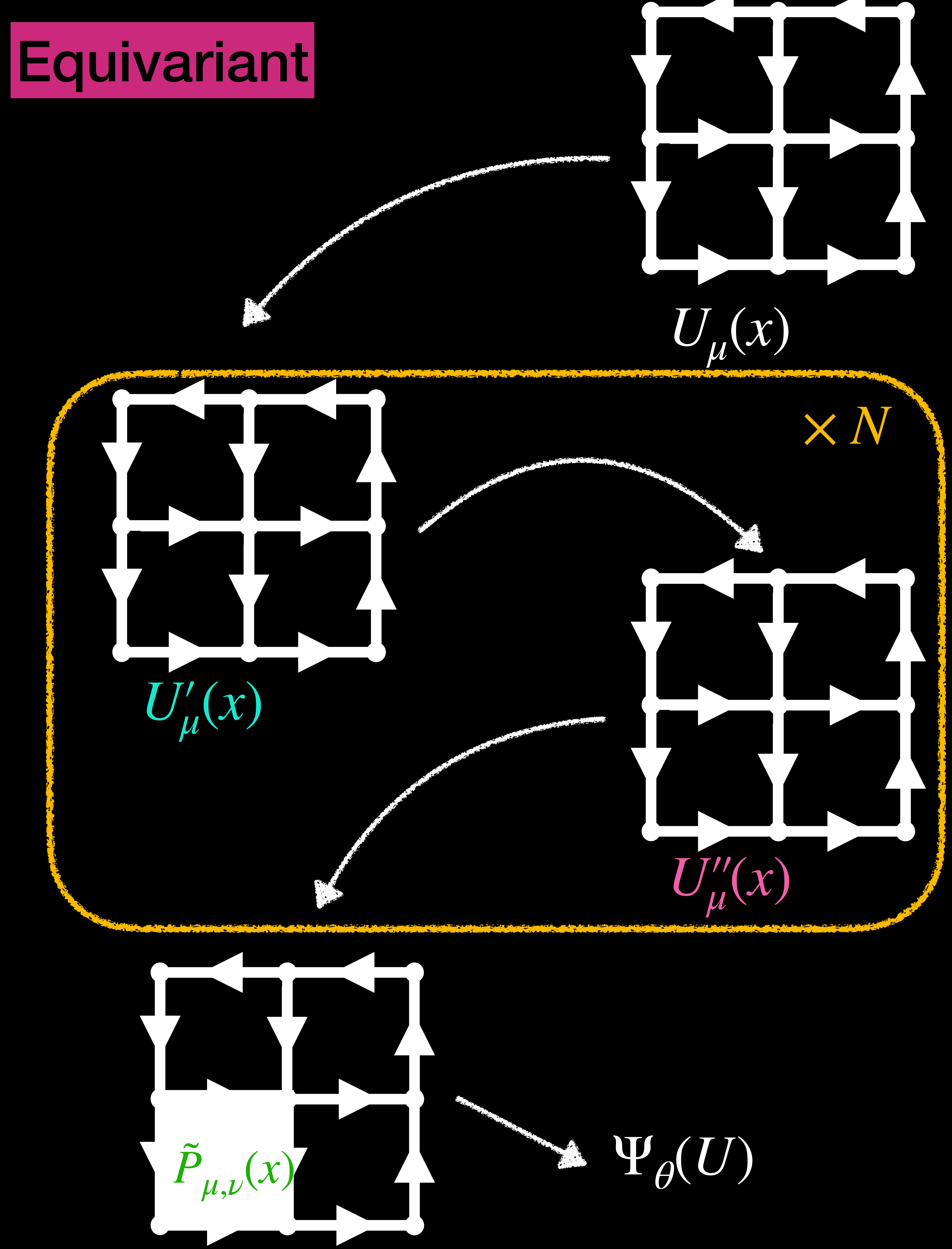
Benchmark simulations

Simulation 1

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

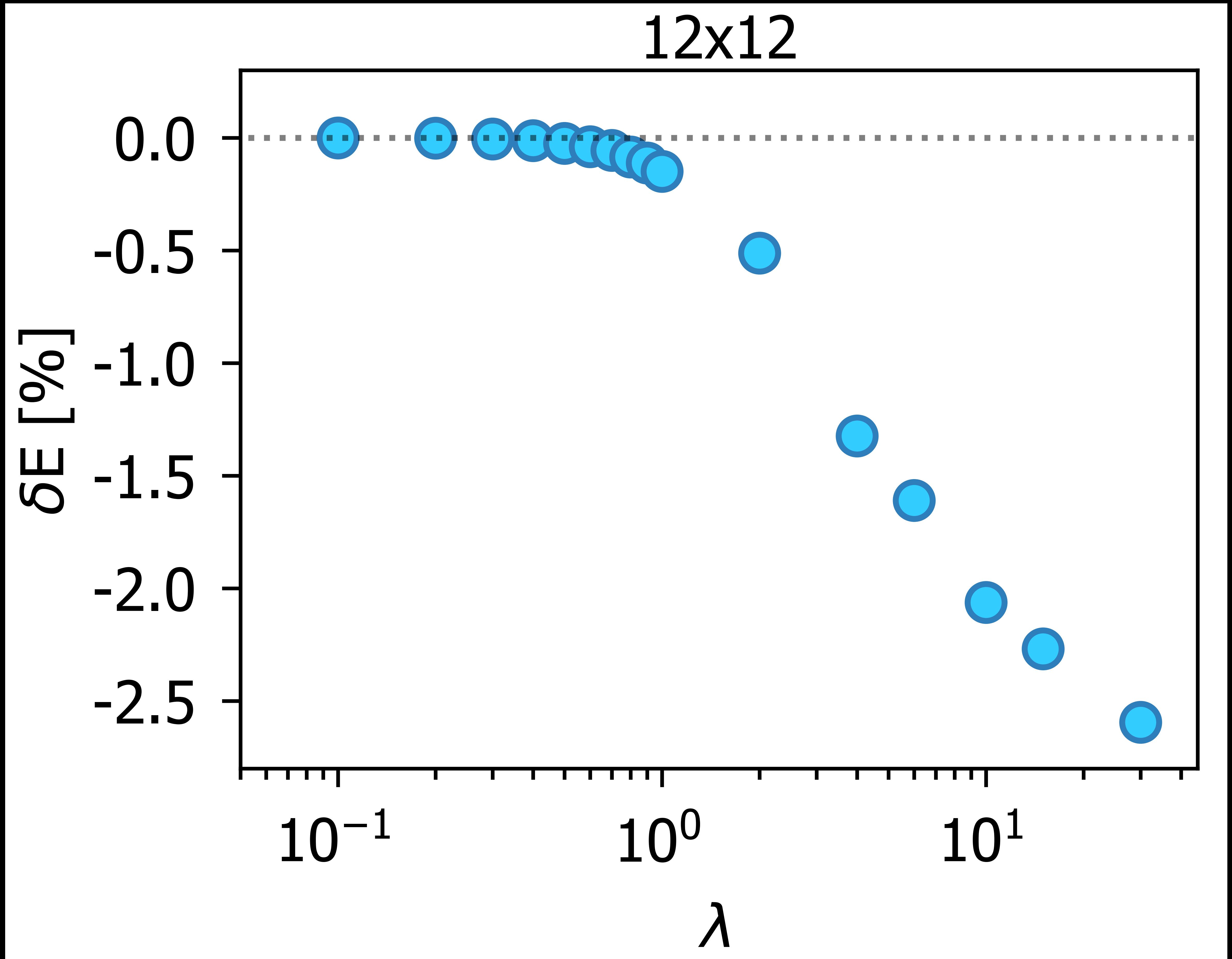
- Lattice of size 12x12
- Goal: find the lowest ground state energy at various λs
- Ansätze:
  - simple, invariant-only, ansatz -> **Jastrow**
  - multi-layer equivariant+invariant ansatz -> **Equivariant**
- Relative energy decrease coming from new ansatz:
  - $$\delta E = \frac{E_{Equivariant} - E_{Jastrow}}{E_{Jastrow}}$$


Jastrow
$U_\mu(x)$
$P_{\mu,\nu}(x)$
$\Psi_\theta(U)$


Equivariant
$U_\mu(x)$
$\times N$
$U'_\mu(x)$
$U''_\mu(x)$
$\tilde{P}_{\mu,\nu}(x)$
$\Psi_\theta(U)$

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

Simulation 1

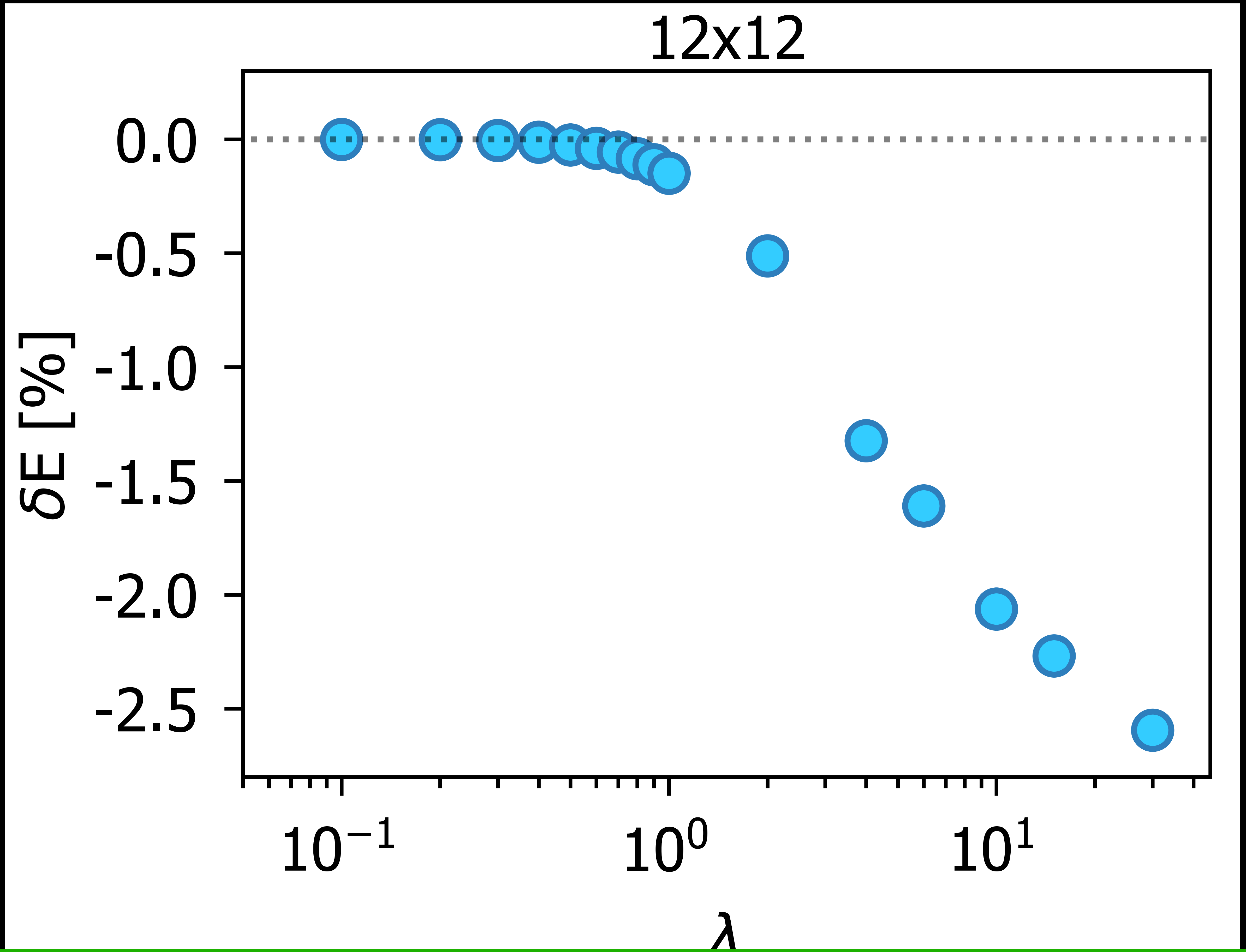


g range ~ 2.5 - 0.6

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

Simulation 1

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



g range ~ 2.5 - 0.6

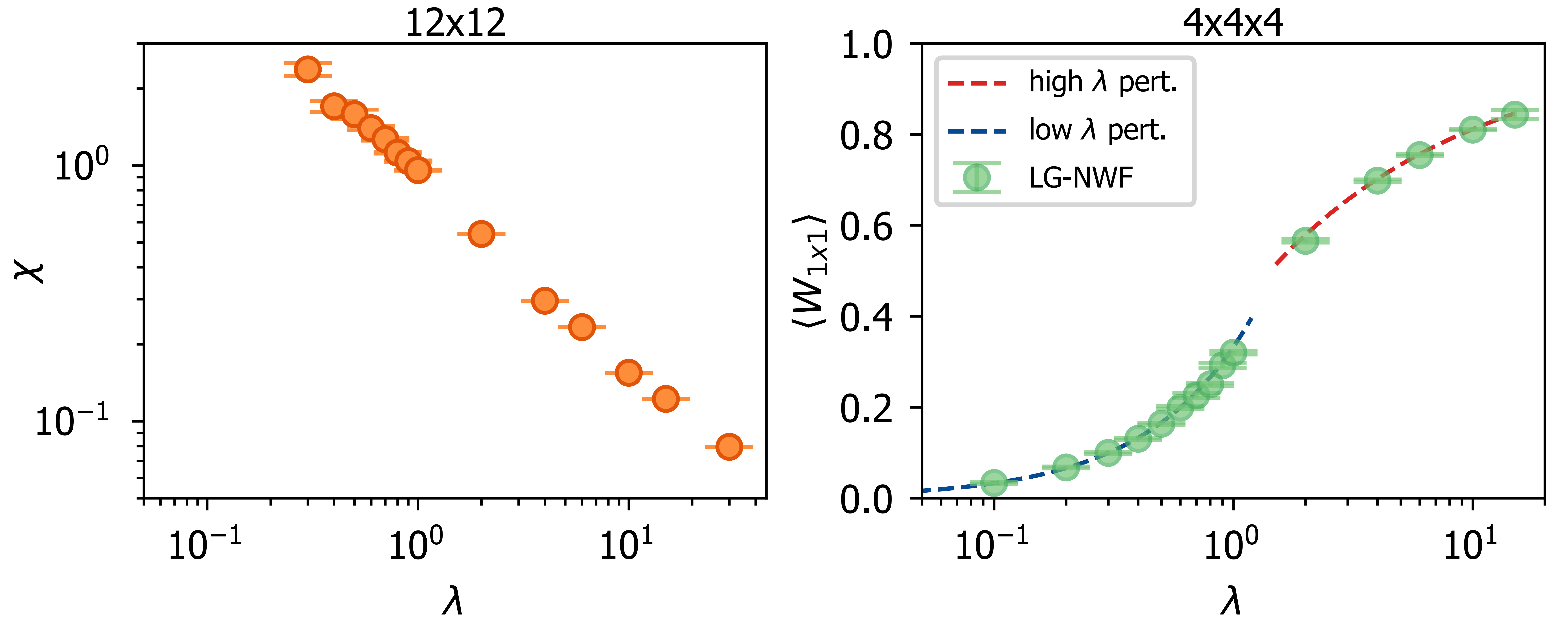
Equivariant layers needed to learn a better representation of the ground state

Simulation 2

- Lattice of size 12x12 and 4x4x4
- Goal: measure physical observables
- Ansätze:
  - Equivariant
- Average 1x1 Wilson loop $\langle W^{1\times1}\rangle$
- Creutz ratio $\chi = -\log\left(\frac{\langle W^{1\times1}\rangle\langle W^{2\times2}\rangle}{\langle W^{1\times2}\rangle\langle W^{2\times1}\rangle}\right)$

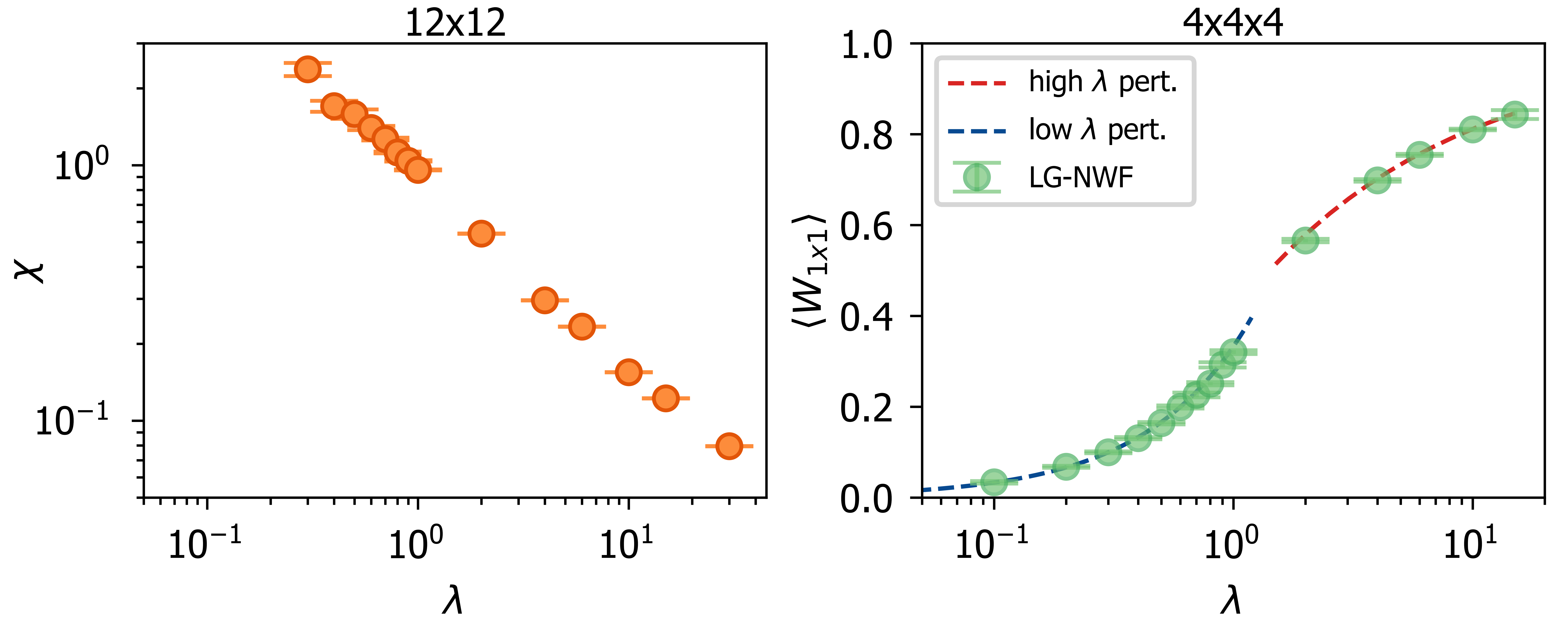
g range ~ 2.5 - 0.6

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

In closing

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(2) with fermions
  - SU(3)
  - Open to suggestions/collaborations (our code is public, see *arXiv:2509.12323*)

[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

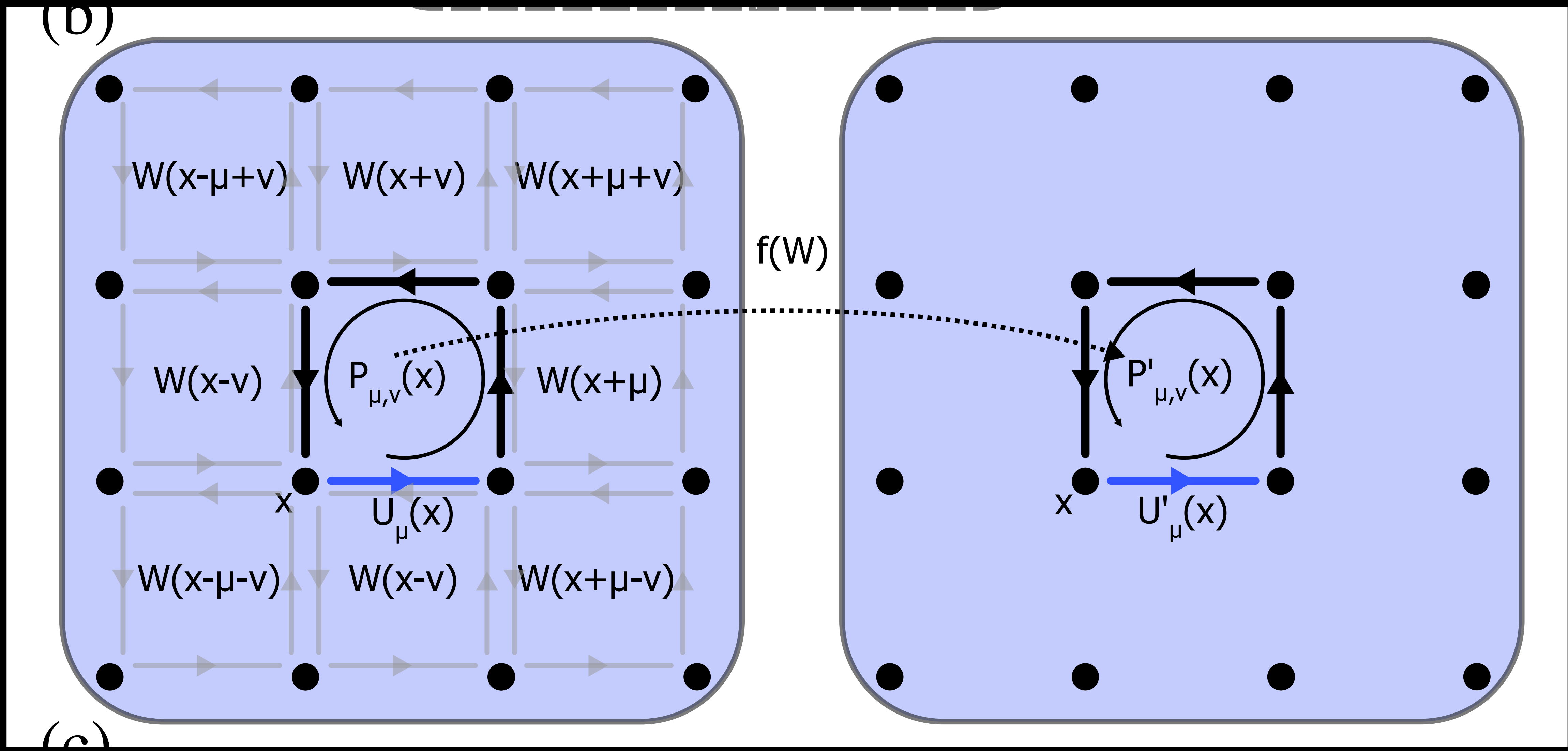
$$U'_\mu(x) = P'_{\mu,\nu}(x) * (U_\nu(x+\mu)U^\dagger_\mu(x+\nu)U^\dagger_\nu(x))^{-1}$$

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

CNN

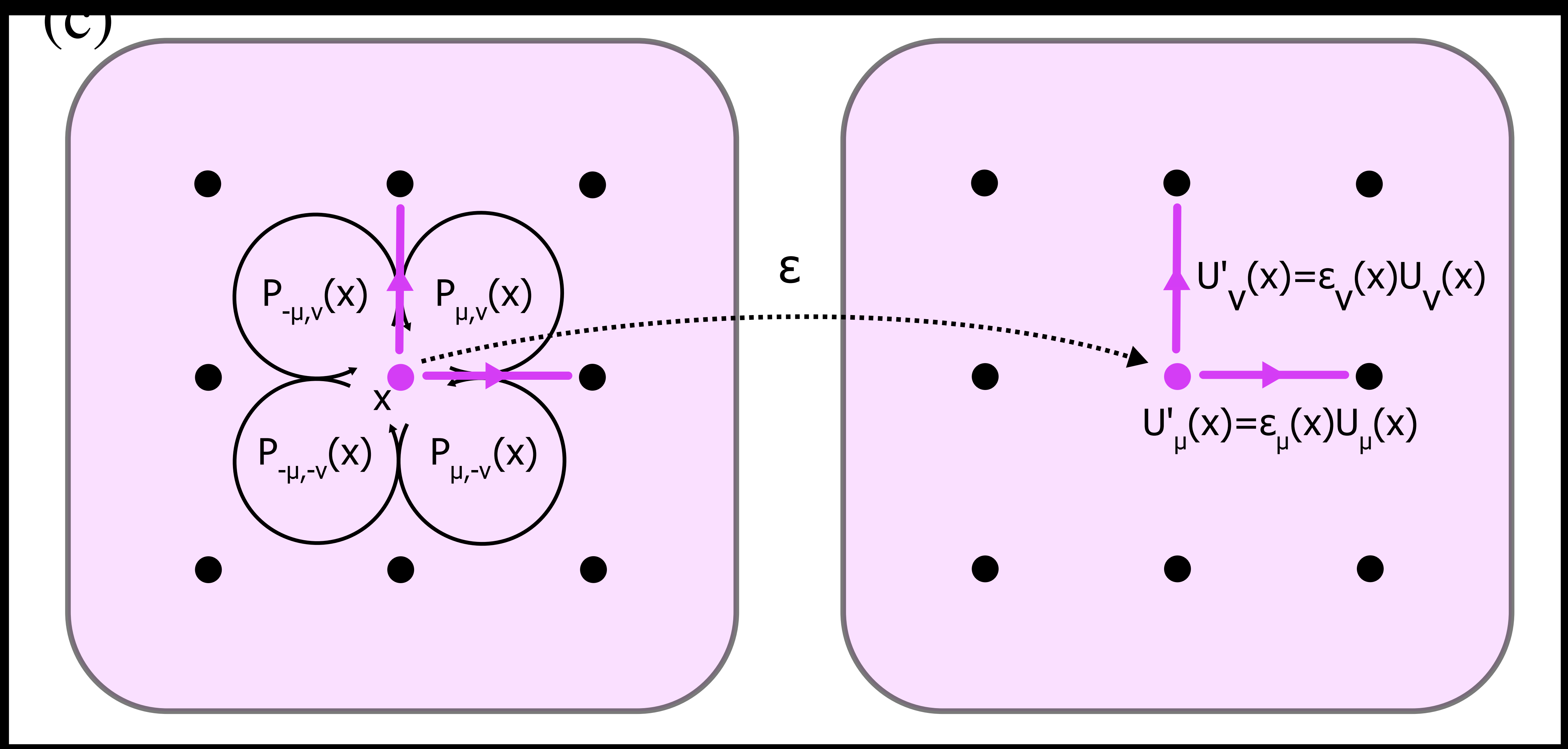
Eigenvalue decomposition

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



Eigenvalue recomposition

Node equivariant layer



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

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

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

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

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

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

Gauge symmetry

Our new SU(2) invariant ansatz

Equivariant

Invariant

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

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

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