

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

Simone Romiti

University of Bern, Switzerland

Lattice 2026

u^b

^b
UNIVERSITÄT
BERN



Summary

Wavefunction ansatz

Neural networks represent $\psi[\mathcal{U}]$ directly; gauge invariance built in via L-CNN + MLP

Renormalization flow

Training moves along $1/g : 0 \rightarrow \infty$; network adapts to $H(g)$ without Hilbert-space truncation

Simone Romiti. “SU(N) lattice gauge theories with physics-informed neural networks”. In: *Phys. Rev. D* 113 (2026), p. 054511. DOI: [10.1103/mb67-9rkf](https://doi.org/10.1103/mb67-9rkf)

- ▶ U(1) and SU(2) single plaquette: exact agreement with exact diagonalization
- ▶ SU(3) in progress

Table of contents

Lattice QCD: Hamiltonian methods

Why Hamiltonians on the lattice

Kogut-Susskind Hamiltonian

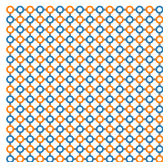
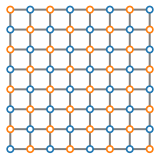
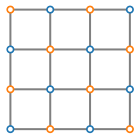
Physics-Informed Neural Networks

Training strategy

Loss function

Results

Lattice gauge theories

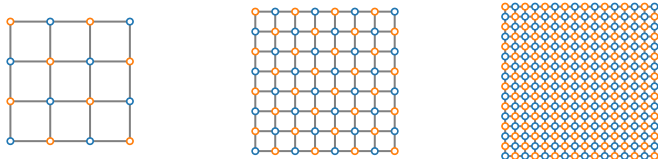


Expectation values (continuum limit)

$$Z_T = \text{Tr} [e^{-HT}] = \int \mathcal{D}\phi e^{-S_E[\phi]}$$

$$\langle O \rangle = \frac{1}{Z_T} \text{Tr} [e^{-HT} O] = \int \mathcal{D}\phi O[\phi] \frac{e^{-S_T[\phi]}}{Z_T}$$

Lattice gauge theories



Expectation values (continuum limit)

$$Z_T = \text{Tr} [e^{-HT}] = \int \mathcal{D}\phi e^{-S_E[\phi]}$$

$$\langle O \rangle = \frac{1}{Z_T} \text{Tr} [e^{-HT} O] = \int \mathcal{D}\phi O[\phi] \frac{e^{-S_T[\phi]}}{Z_T}$$

Physical predictions by extrapolation

Multiple simulations at different lattice spacings a and volumes V ,

- **Infinite volume:** $V \rightarrow \infty$
(increase N_s)
- **Continuum limit:** $a \rightarrow 0$
(renormalization flow)

Lagrangian VS Hamiltonian

Lagrangian = path integral + Monte Carlo

- ▶ Wick rotation $it \implies$ analytic continuation
- ▶ Sign problem, critical slowing down, topological freezing

Lagrangian VS Hamiltonian

Lagrangian = path integral + Monte Carlo

- ▶ Wick rotation $it \implies$ analytic continuation
- ▶ Sign problem, critical slowing down, topological freezing

Hamiltonian = Hilbert space + Schrödinger equation

- ▶ No sampling $\implies$ no MC limitations
- ▶ **Challenge:** infinite-dimensional Hilbert space

Lagrangian VS Hamiltonian

Lagrangian = path integral + Monte Carlo

- ▶ Wick rotation $it \implies$ analytic continuation
- ▶ Sign problem, critical slowing down, topological freezing

Hamiltonian = Hilbert space + Schrödinger equation

- ▶ No sampling $\implies$ no MC limitations
- ▶ **Challenge:** infinite-dimensional Hilbert space

Finite-difference methods (QC, TN)

- ▶ Choose a fixed basis at strong coupling
- ▶ Basis is rigid: adequate only in a region of the coupling
- ▶ Truncation becomes inadequate toward the continuum $g \rightarrow 0$

Wavefunctions on the lattice

State vector

$$|\psi\rangle = \int \mathcal{D}U \, \psi(\mathcal{U}) |\mathcal{U}\rangle, \quad \psi(\mathcal{U}) = \langle \mathcal{U} | \psi \rangle$$

Inner product & expectation

$$\langle \phi | O | \psi \rangle = \int \mathcal{D}U \, \phi^*(\mathcal{U}) O(\mathcal{U}) \psi(\mathcal{U})$$

$$\langle \psi | O | \psi \rangle = \int \mathcal{D}U \, |\psi(\mathcal{U})|^2 O(\mathcal{U})$$

No sign problem

$|\psi|^2 \geq 0 \implies$ importance sampling directly from the wavefunction

What does it mean to “solve” the theory?

Hamiltonian approach

$$H |n, \alpha\rangle = E_{n,\alpha} |n, \alpha\rangle$$

$E_{n,\alpha}$ and $\psi_{n,\alpha}$ are **direct outputs**

$$C(t) = \sum_n |\langle \Omega | O | n \rangle|^2 e^{-E_n t}$$

- ▶ Features extracted *a posteriori* from t dependence $\implies$ signal-to-noise ratio hindrance
- ▶ Sign problem for observables with imaginary action

Table of contents

Lattice QCD: Hamiltonian methods

Why Hamiltonians on the lattice

Kogut-Susskind Hamiltonian

Physics-Informed Neural Networks

Training strategy

Loss function

Results

Kogut-Susskind Hamiltonian

Hamiltonian

$$H_{\text{KS}} = \underbrace{\frac{g^2}{2} \sum_{\vec{x}, \mu, a} (L_a)_\mu^2(\vec{x})}_{H_E} - \underbrace{\frac{1}{g^2} \sum_{\vec{x}, \mu > \nu} \text{Tr} [U_{\mu\nu} + U_{\mu\nu}^\dagger]}_{H_B}(\vec{x})$$

H_E : chromoelectric term, H_B : chromomagnetic term

Analogy

- ▶ N interacting particles (one per link)
- ▶ Plaquette term $\leftrightarrow$ many-body potential
- ▶ $H_{\text{KS}} \xrightarrow{a \rightarrow 0} \int d^3x (\vec{E}^2 + \vec{B}^2)/2$

Canonical momenta

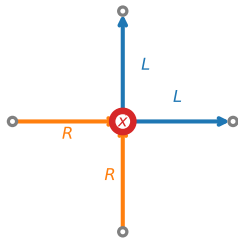
Left and right generators

Left and right gauge transformations for a single link are independent.

$$(L_a)_\mu \psi(\mathcal{U}) = -i \frac{d}{d\omega} \psi(e^{-i\omega\tau_a} \mathcal{U}_\mu) \Big|_{\omega=0}, \quad (R_a)_\mu \psi(\mathcal{U}) = -i \frac{d}{d\omega} \psi(\mathcal{U}_\mu e^{i\omega\tau_a}) \Big|_{\omega=0}$$

Gauss's law

$$G_a(\vec{x}) = \sum_\mu (L_a)_\mu(\vec{x}) + (R_a)_\mu(\vec{x} - \hat{\mu}), \quad G_a |\psi\rangle_{\text{phys}} = 0$$



Hamiltonian methods: revisiting the truncation problem

Key principle

- ▶ Monte Carlo: importance sampling
- ▶ Hamiltonian: effective truncation of the Hilbert space

Fixed basis (QC/TN)

Chosen at strong coupling, inadequate toward $g \rightarrow 0$

Network advantage

The representation *adapts* to $H(g)$ and follows the RG flow. Action of L_a , R_a is **exact** (autodifferentiation); cost is **coupling-independent**.

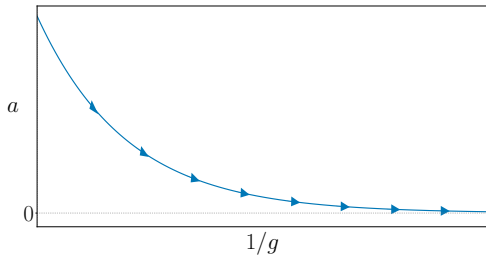
$$\hat{H}(g) \psi_n(g) = E_n(g) \psi_n(g)$$

Finding the ansatz for $\psi[\mathcal{U}]$

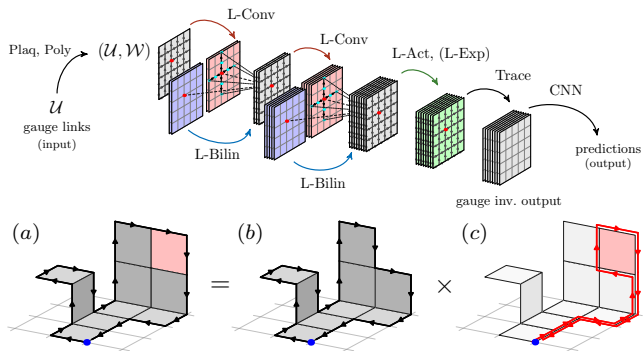
L-CNN[2] + MLP ansatz

- ▶ Built from **gauge-covariant** building blocks
- ▶ **Automatic gauge invariance** — Gauss law by construction
- ▶ Gauge-invariant trace + MLP $\rightarrow$ scalar $\psi[\mathcal{U}] \in \mathbb{C}$
- ▶ Learn ψ_n, E_n from Schrödinger equation
- ▶ Evolve along RG flow $1/g : 0 \rightarrow \infty$

$$\hat{H} = \hat{H}(g); \text{continuum at } 1/g \rightarrow \infty$$



How does the L-CNN work?



Three layers[3]

1. L-Conv:

$$W' = \sum_{\mu,k} \omega_{\mu k} U_{\mu} W U_{\mu}^{\dagger}$$

2. L-Bilin: $W'' = \sum_{ij} \alpha_{ij} W'_i W'_j$

3. L-Act: $W''' = f(\text{Tr } W'') W''$

All layers are locally gauge covariant; Tr at the output projects to invariants.

L-CNN + MLP architecture



Output

$$\psi(U) = \text{MLP}(\text{Tr}[W''']) \in \mathbb{C}$$

- ▶ L-CNN $\rightarrow$ gauge **covariant**
- ▶ Tr $\rightarrow$ gauge **invariant**
- ▶ MLP $\rightarrow$ function of invariants

Gauss' law by construction (gauge invariance)

$$G_a(\vec{x}) \psi = 0 \quad \forall \vec{x}, a$$

Table of contents

Lattice QCD: Hamiltonian methods

Why Hamiltonians on the lattice

Kogut-Susskind Hamiltonian

Physics-Informed Neural Networks

Training strategy

Loss function

Results

Physics-Informed Neural Networks (PINNs)

Key idea (example)

Neural network $f_\theta(x)$ solves the DE via automatic differentiation

$$\begin{cases} \frac{df}{dx} = -xf, \\ f(0) = 1 \end{cases} \implies \mathcal{L} = \sum_i \left| \frac{df_\theta}{dx_i} + x_i f_\theta(x_i) \right|^2 + |f_\theta(0) - 1|^2$$

No training data needed

The DE itself is the training signal

Interpolation property

Solution quality depends on where points are sampled; network interpolates between training configurations

PINNs for the Schrödinger equation

Eigenvalue problem

Network represents ψ :

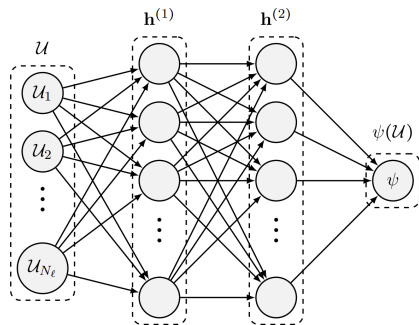
$$H\psi = E\psi$$

E learned as an additional parameter

Normalization (bound states)

$$\int \mathcal{D}U |\psi|^2 = 1$$

Condition for confined hadrons

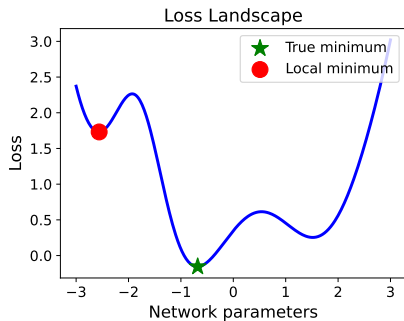


$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_{\theta} \mathcal{L}(\theta)$$

$\theta = \{\text{weights}, E\}$ tuned via
gradient descent

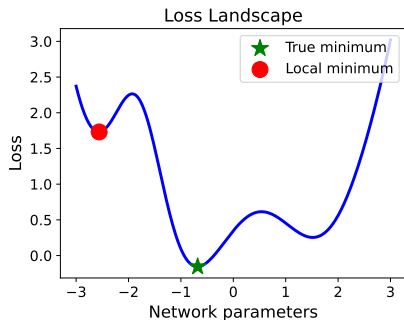
Loss function: minimization challenges

The loss function is in general hard to minimize...



Loss function: minimization challenges

The loss function is in general hard to minimize...

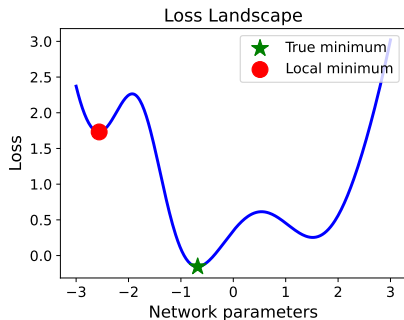


How do we formulate the problem?

- ▶ We do not want to accidentally end up in the minimum corresponding to another eigenstate
- ▶ We want to access different eigenstates, not just the ground state (cf. variational methods)
- ▶ Once we train at some coupling, how do we move to another coupling?

Loss function: minimization challenges

The loss function is in general hard to minimize...



How do we formulate the problem?

- ▶ We do not want to accidentally end up in the minimum corresponding to another eigenstate
- ▶ We want to access different eigenstates, not just the ground state (cf. variational methods)
- ▶ Once we train at some coupling, how do we move to another coupling?

Solution: adiabatic flow

Start at $1/g = 0$ where eigenstates are known analytically; slowly vary g to track the state along the RG flow

Table of contents

Lattice QCD: Hamiltonian methods

Why Hamiltonians on the lattice

Kogut-Susskind Hamiltonian

Physics-Informed Neural Networks

Training strategy

Loss function

Results

Exciting eigenstates from the vacuum

Strong-coupling eigenstates ϕ_n

$$\phi_0[\mathcal{U}] = 1, \quad \phi_n[\mathcal{U}] = \text{known analytically}$$

Neural-network eigenstates

$$\psi_n(g)[\mathcal{U}] = \text{NN}[\mathcal{U}], \quad \psi_n(\infty) = \phi_n$$

At finite g , $\psi_n(g)$ is the same physical state deformed by the RG flow; the NN learns the deformation.

Why this matters

- ▶ Without analytic ϕ_n : variational method $\Rightarrow$ ground state only
- ▶ With ϕ_n : pick any sector at $1/g = 0$, track it along $g : \infty \rightarrow 0$

Solution: adiabatic learning

Adiabatic evolution

If at $t = 0$ we start from an eigenstate $|n_0\rangle$ ($c_m(0) = \delta_{mn_0}$), and vary $g(t)$:

$$|\psi(t)\rangle = \sum_i c_i(t) |n(t)\rangle$$

$$\dot{c}_n + i\dot{\theta}_{\text{tot}}c_n = \sum_{m \neq n} \frac{\langle n | \dot{H} | m \rangle}{E_n - E_m} c_m$$

Adiabatic theorem

Small $\dot{H} \implies$ stays in instantaneous eigenstate

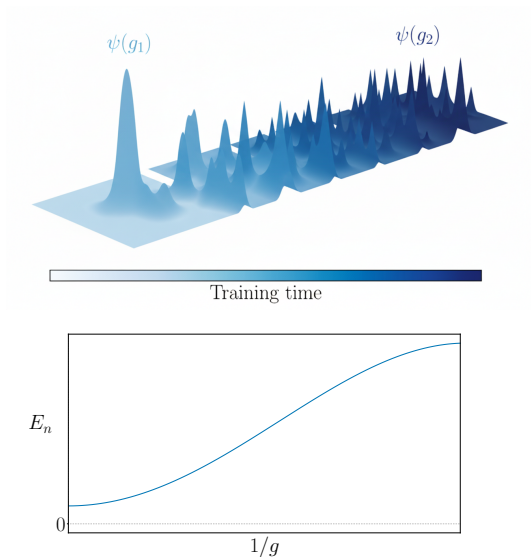


Table of contents

Lattice QCD: Hamiltonian methods

Why Hamiltonians on the lattice

Kogut-Susskind Hamiltonian

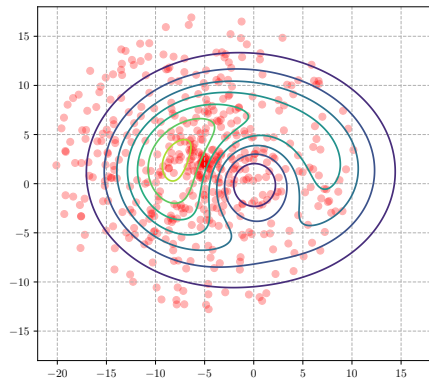
Physics-Informed Neural Networks

Training strategy

Loss function

Results

Loss at Strong Coupling ($1/g \rightarrow 0$)



Supervised training

$$\mathcal{L}^{(1/g \rightarrow 0)} = \frac{1}{N_{\text{pt}}} \sum_{i=1}^{N_{\text{pt}}} |\psi(\mathcal{U}_i) - \phi_n(\mathcal{U}_i)|^2$$

Points sampled via Langevin dynamics of ϕ_n

ϕ_n is an eigenstate: $H_E \phi_n = E_n^{1/g=0} \phi_n$

Loss at Finite Coupling ($1/g \neq 0$)

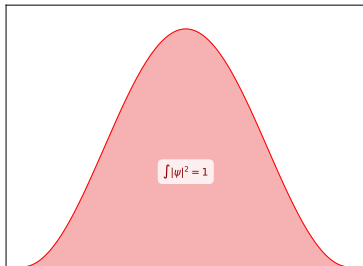
Total Loss

$$\mathcal{L} = \omega_{\text{PDE}} \mathcal{L}_{\text{PDE}} + \omega_{\text{norm}} \mathcal{L}_{\text{norm}}$$

Individual terms

$$\mathcal{L}_{\text{PDE}} = \frac{1}{N_{\text{pt}}} \sum_i |(H - E)\psi(\mathcal{U}_i)|^2$$

$$\mathcal{L}_{\text{norm}} = \left(1 - \sqrt{\int d\mathcal{U} |\psi(\mathcal{U})|^2}\right)^2$$



$$\int d\mathcal{U} |\psi(\mathcal{U})|^2 \approx \frac{1}{N_{\text{pt}}} \sum_i \frac{|\psi(\mathcal{U}_i)|^2}{q(\mathcal{U}_i)}$$

Monte Carlo estimate

$\mathcal{U}_i \sim q(\mathcal{U})$ (reference distribution)

Sampling from $|\psi|^2$ via Langevin dynamics

Stochastic flow

$$dx_t = g_t(x) dt + \frac{\sigma^2}{2} \nabla \log(\sigma^2 p_t(x)) dt + \sigma dW$$

$g_t(x) = 0 \implies$ Langevin dynamics $\implies$ convergence to $p_\infty(x)$.

Sampling from $|\psi|^2$ via Langevin dynamics

Stochastic flow

$$dx_t = g_t(x) dt + \frac{\sigma^2}{2} \nabla \log(\sigma^2 p_t(x)) dt + \sigma dW$$

$g_t(x) = 0 \implies$ Langevin dynamics $\implies$ convergence to $p_\infty(x)$.

Langevin dynamics on $SU(N)$

$$U \rightarrow e^{i\theta^a \tau_a} U$$

$$\theta^a = -\varepsilon \nabla_a \log p(U) + \sqrt{2\varepsilon} \eta^a$$

Sampling from $|\psi|^2$

$$\theta^a = -\varepsilon L_a[\log |\psi|^2] + \sqrt{2\varepsilon} \eta^a$$

Stationary solution: $p_\infty(U) \propto |\psi(U)|^2$

Sampling from $|\psi|^2$ via Langevin dynamics

Stochastic flow

$$dx_t = g_t(x) dt + \frac{\sigma^2}{2} \nabla \log(\sigma^2 p_t(x)) dt + \sigma dW$$

$$g_t(x) = 0 \implies \text{Langevin dynamics} \implies \text{convergence to } p_\infty(x).$$

Langevin dynamics on $SU(N)$

$$U \rightarrow e^{i\theta^a \tau_a} U$$

$$\theta^a = -\varepsilon \nabla_a \log p(U) + \sqrt{2\varepsilon} \eta^a$$

Normalization at the next coupling

$$\int d\mathcal{U} |\psi_{g_{i+1}}|^2 = \left\langle \frac{|\psi_{g_{i+1}}(\mathcal{U})|^2}{|\psi_{g_i}(\mathcal{U})|^2} \right\rangle_{|\psi_{g_i}|^2}$$

Sampling from $|\psi|^2$

$$\theta^a = -\varepsilon L_a[\log |\psi|^2] + \sqrt{2\varepsilon} \eta^a$$

Stationary solution: $p_\infty(U) \propto |\psi(U)|^2$

Table of contents

Lattice QCD: Hamiltonian methods

Why Hamiltonians on the lattice

Kogut-Susskind Hamiltonian

Physics-Informed Neural Networks

Training strategy

Loss function

Results

U(1): single plaquette

$$\mathcal{U} = e^{i\phi}, \quad L = -R = i \frac{\partial}{\partial \phi}.$$

Hamiltonian

$$H = -\frac{g^2}{2} \sum_k \frac{\partial^2}{\partial \phi_k^2} - \frac{2}{g^2} \sum_{\square} \cos \phi_{\square}.$$

Schrödinger equation after gauge fixing

$$\left(-2g^2 \frac{d^2}{d\omega^2} - \frac{2}{g^2} \cos \omega \right) \psi(\omega) = E \psi(\omega).$$

U(1): single plaquette

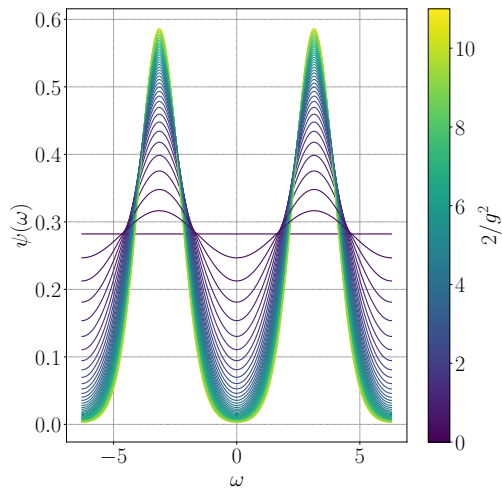
$$\mathcal{U} = e^{i\phi}, \quad L = -R = i \frac{\partial}{\partial \phi}.$$

Hamiltonian

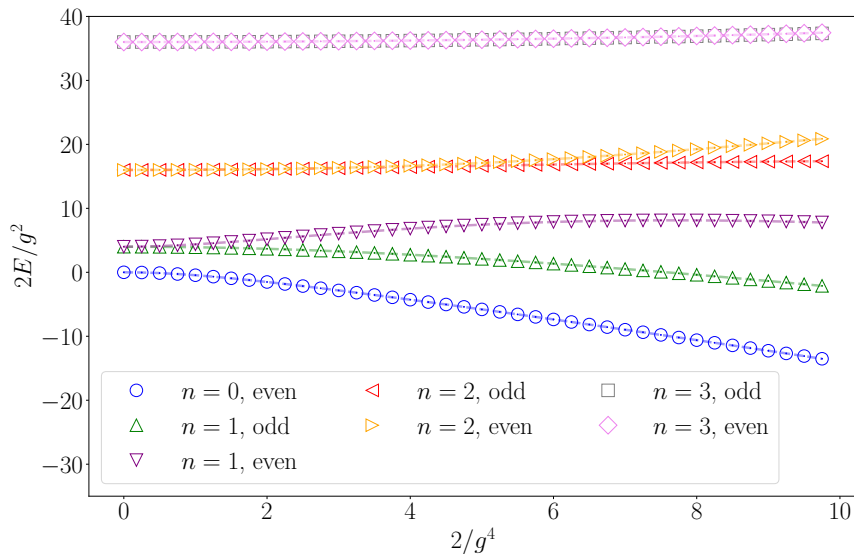
$$H = -\frac{g^2}{2} \sum_k \frac{\partial^2}{\partial \phi_k^2} - \frac{2}{g^2} \sum_{\square} \cos \phi_{\square}.$$

Schrödinger equation after gauge fixing

$$\left(-2g^2 \frac{d^2}{d\omega^2} - \frac{2}{g^2} \cos \omega \right) \psi(\omega) = E \psi(\omega).$$



U(1) spectrum: PINNs VS exact results



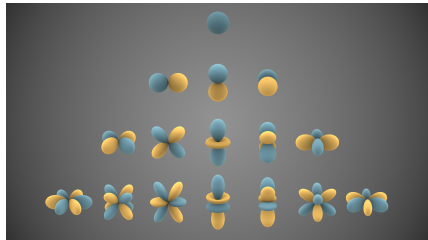
SU(2): single plaquette

Gauge link

$$\mathcal{U}(\omega, \theta, \phi) = \begin{pmatrix} \cos \frac{\omega}{2} - i \sin \frac{\omega}{2} \cos \theta & -i \sin \frac{\omega}{2} \sin \theta e^{-i\phi} \\ -i \sin \frac{\omega}{2} \sin \theta e^{i\phi} & \cos \frac{\omega}{2} + i \sin \frac{\omega}{2} \cos \theta \end{pmatrix}$$

Hamiltonian

$$H = 2g^2 \left[\frac{\mathbb{L}^2}{4 \sin^2 \frac{\omega}{2}} - \frac{\partial^2}{\partial \omega^2} - \cot \frac{\omega}{2} \frac{\partial}{\partial \omega} \right] - \frac{4}{g^2} \cos \frac{\omega}{2}.$$



SU(2) after gauge fixing

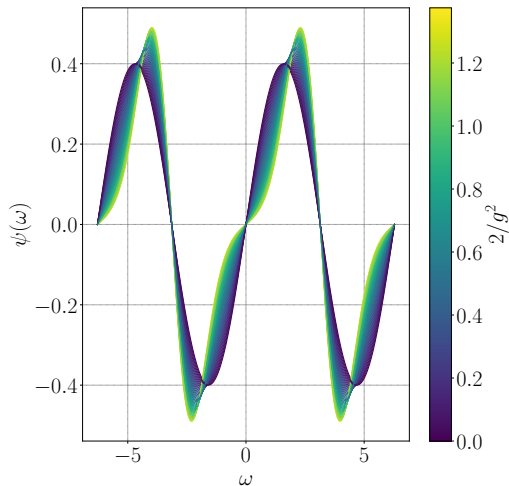
Gauss's law

$$G_a^2 = \mathbb{L}^2$$

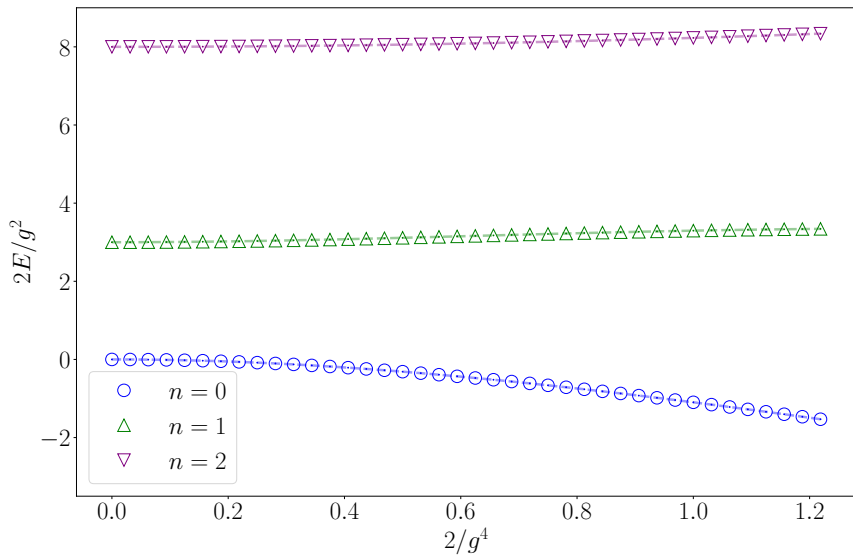
Gauge-invariant sector

$$\psi = Y_0^0(\theta, \phi) \frac{u(\omega)}{\sin(\omega/2)}$$

$$\left[-\frac{d^2}{d\omega^2} - \frac{1}{4} - \frac{2}{g^4} \cos \frac{\omega}{2} \right] u = \frac{E}{2g^2} u$$



SU(2) spectrum: PINNs VS exact results



Conclusion

Achievements

- ▶ PINNs solve lattice gauge Schrödinger eq. **without Hilbert-space truncation**
- ▶ **L-CNN** $\implies$ gauge invariance by construction, no Gauss penalty
- ▶ Adiabatic training follows the RG flow $g : \infty \rightarrow 0$
- ▶ U(1) and SU(2) single-plaquette: exact agreement

Outlook

SU(3), larger volumes, fermions, real-time dynamics

Thank you for your attention!



Bibliography I

- [1] Simone Romiti. “SU(N) lattice gauge theories with physics-informed neural networks”. In: *Phys. Rev. D* 113 (2026), p. 054511. DOI: [10.1103/mb67-9rkf](https://doi.org/10.1103/mb67-9rkf) (cit. on p. 2).
- [2] Kieran Holland, Andreas Ipp, David I. Müller, and Urs Wenger. “Machine learning a fixed point action for SU(3) gauge theory with a gauge equivariant convolutional neural network”. In: *Phys. Rev. D* 110.7 (2024), p. 074502. DOI: [10.1103/PhysRevD.110.074502](https://doi.org/10.1103/PhysRevD.110.074502). arXiv: [2401.06481 \[hep-lat\]](https://arxiv.org/abs/2401.06481) (cit. on p. 15).
- [3] Matteo Favoni, Andreas Ipp, David I. Müller, and Daniel Schuh. “Lattice Gauge Equivariant Convolutional Neural Networks”. In: (Dec. 2020). arXiv: [2012.12901 \[hep-lat\]](https://arxiv.org/abs/2012.12901) (cit. on p. 16).