Machine-Learned Fixed-Point Actions in Four-Dimensional SU(3) Gauge Theory

Andreas Ipp

Institute for Theoretical Physics, TU Wien, Austria

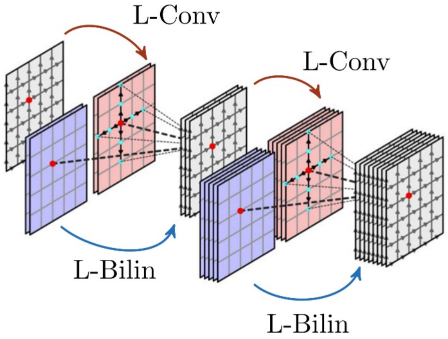
EuCAIFCon 2025
Cagliari, Sardinia, June 18, 2025

INSTITUTE FOR THEORETICAL PHYSICS

Der Wissenschaftsfonds.

Overview

Lattice gauge equivariant convolutional neural networks (L-CNNs)



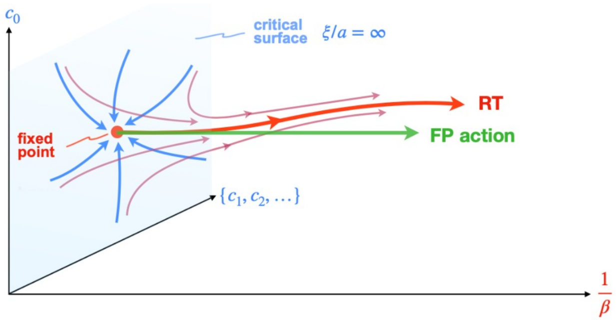
Favoni, AI, Müller, Schuh,
Phys.Rev.Lett. 128 (2022) 032003

Andreas Ipp

David Müller

Open source: https://gitlab.com/openpixi/lge-cnn

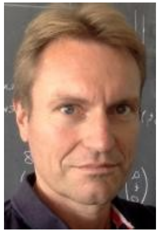
Learning fixed-point actions



Holland, AI, Müller, Wenger,
Phys.Rev.D 110 (2024) 7, 074502

Holland, AI, Müller, Wenger,
arXiv:2504.15870

Kieran Holland
(U. Pacific)



Urs Wenger
(U. Bern)

QCD on the lattice

Lattice Gauge Theory allows for precision determination of non-perturbative quantities in quantum chromodynamics (QCD).

Calculate hadron masses, decay constants, thermodynamic quantities at finite temperature (QCD phase transition), topological properties, or glueball spectra.

Solve Feynman path integral numerically using Monte Carlo methods

Gauge transformation:
Rotation in internal color space
at each space-time point

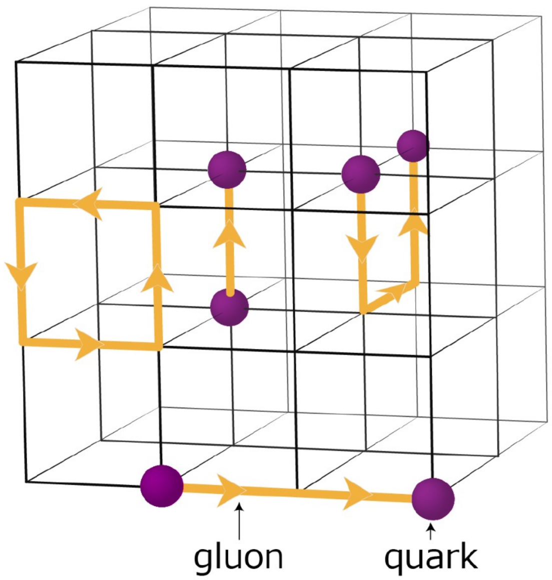


Image from Bi et al. EPJ Web Conf. 245 (2020) 09008

Machine learning approaches:

- Normalizing flow
  Kanwar, Albergo, Boyda, Cranmer, Hackett, Racanière, Rezende, Shanahan (2020), ...
- Continuous normalizing flow
  Gerdes, de Haan, Bondesan, Cheng (2024)
- Diffusion models
  Zhu, Aarts, Wang, Zhou, Wang (2024)
- Fixed point action
  Holland, AI, Müller, Wenger (2024)

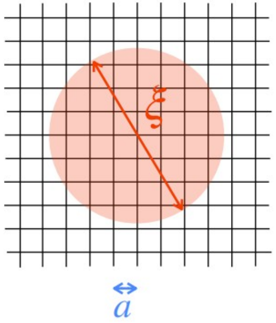
Wilson action for gluons:

$$S_W[U] = \frac{2}{g^2} \sum_{x\in\Lambda} \sum_{\mu<\nu} \mathrm{Tr}\left[\mathbb{1} - U_{x,\mu\nu}\right]$$

Gauge transformation of link variables:

$$U_\mu(n) \;\to\; U'_\mu(n) = \Omega(n)\, U_\mu(n)\, \Omega(n+\hat{\mu})^\dagger$$

Quantum field theory on a lattice



Lattice spacing *a*

Partition function for SU(N_c) gauge theory

$$Z(\beta) = \int \mathscr{D}U \exp\{-\beta A[U]\}$$

Action

with gauge coupling $\beta = \dfrac{2N_c}{g^2}$

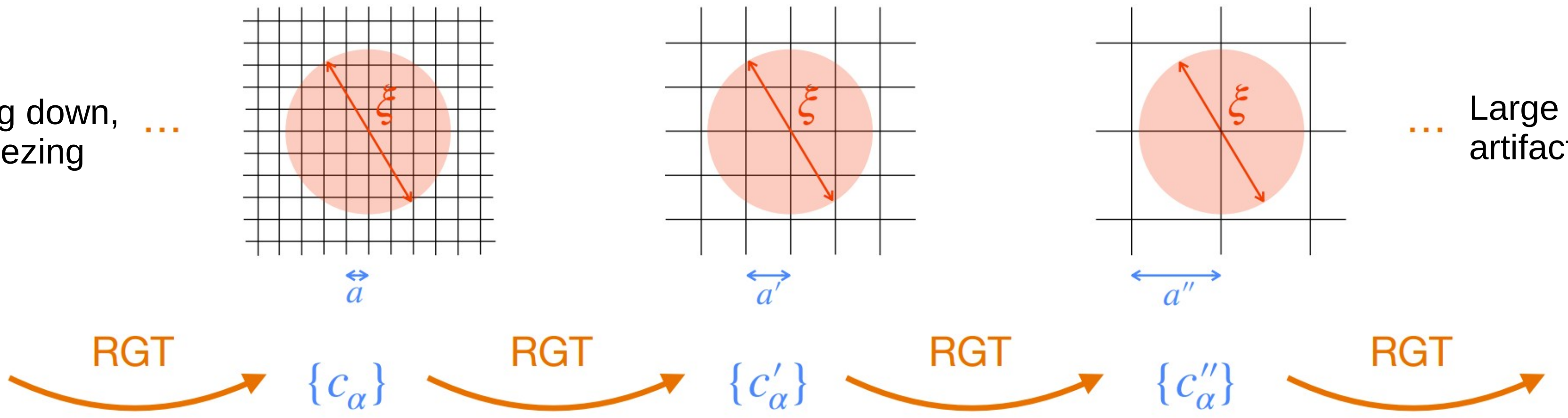
Expectation values of observables can be calculated as

$$\langle \mathcal{O}_\xi(\beta) \rangle = \frac{1}{Z} \int \mathscr{D}U \exp\{-\beta A[U]\}\, \mathcal{O}_\xi[U]$$

for a characteristic length scale ξ

Renormalization group transformation

Critical slowing down, topological freezing …



… Large lattice artifacts

Introduce a (real space) renormalization group transformation (RGT)

$$\exp\left\{-\beta' A'[V]\right\} = \int \mathscr{D}U \exp\left\{-\beta\left(A[U] + T[U,V]\right)\right\}$$

Blocking kernel

The effective action $\beta' A'[V]$ is described

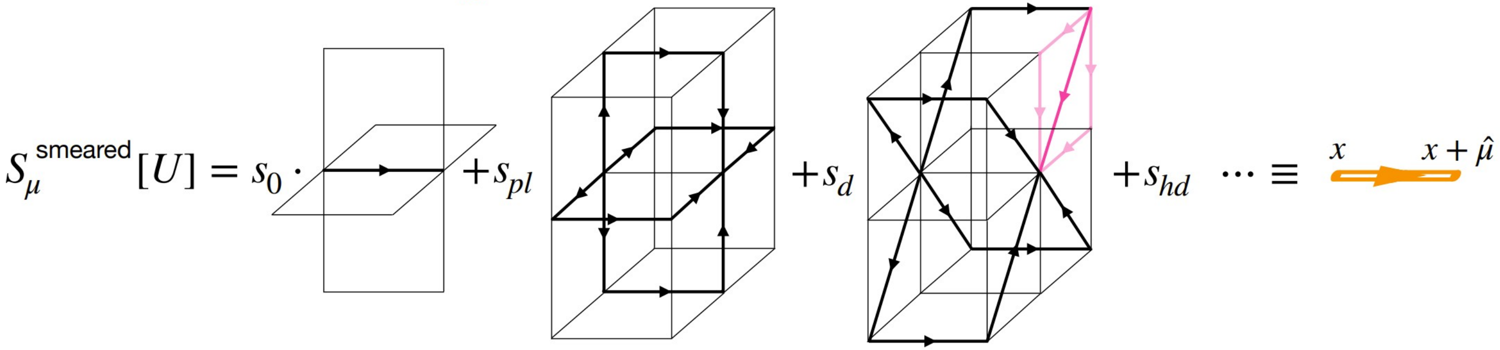
by infinitely many couplings $\{c'_\alpha\}$

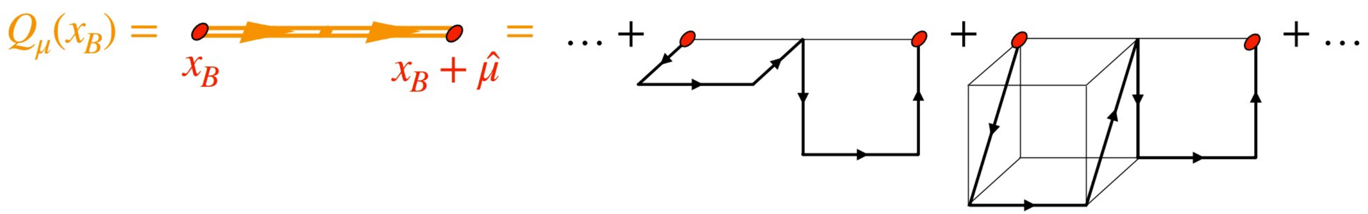
The fixed point is the saddle point in the classical limit $\beta \to \infty$, which can be found by a minimization condition.

P. Hasenfratz, F. Niedermayer, Nucl.Phys.B 414 (1994) 785

Blocking kernel

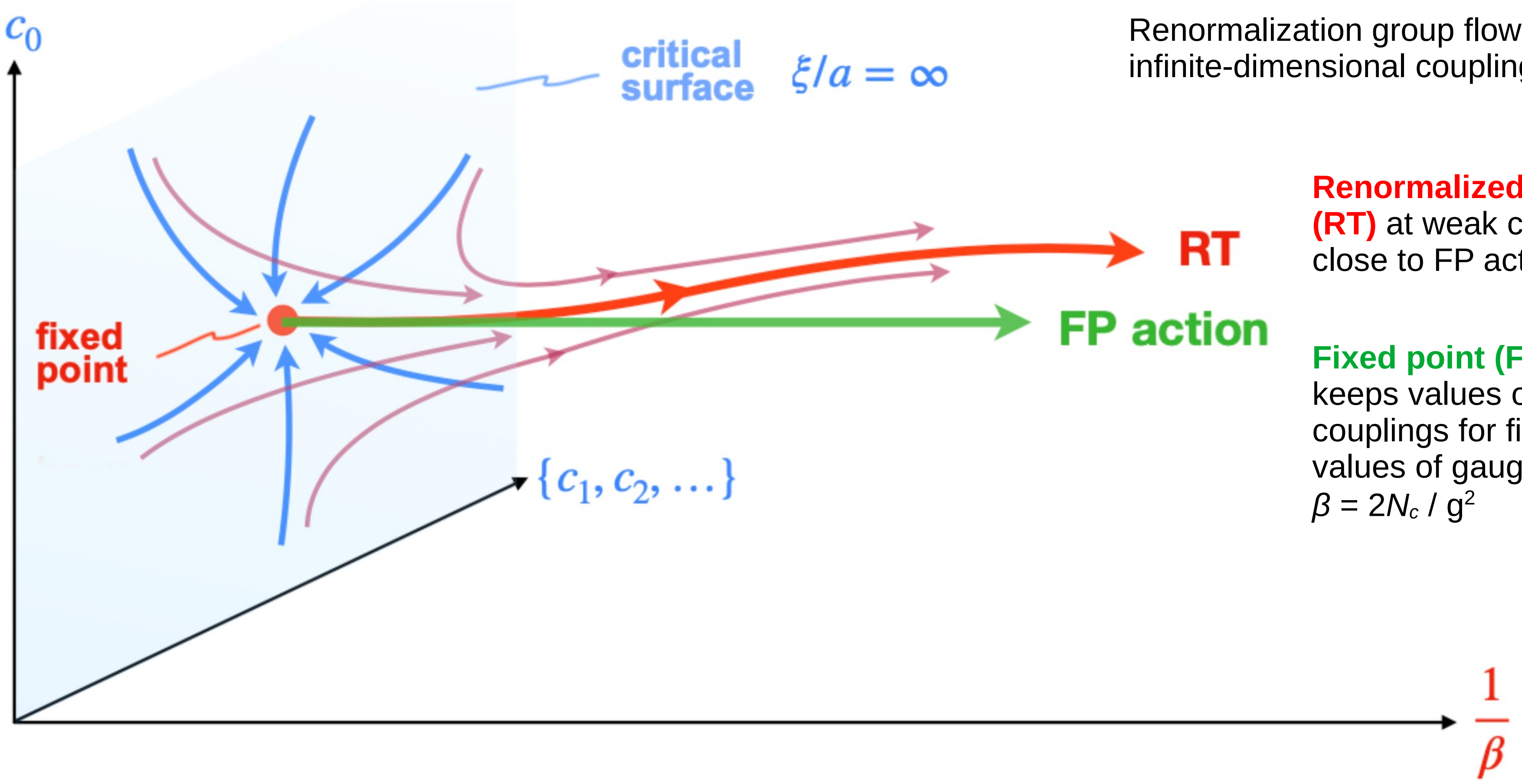
$$T[U,V] = -\frac{\kappa}{N_c} \sum_{x_B,\mu} \left\{ \mathrm{ReTr}\left(V_\mu(x_B) \cdot Q_\mu^\dagger(x_B) \right) - \mathcal{N}_\mu^\beta \right\}$$





Choice of blocking kernel determines how couplings are modified across scales.

Renormalization group transformation and Fixed point action



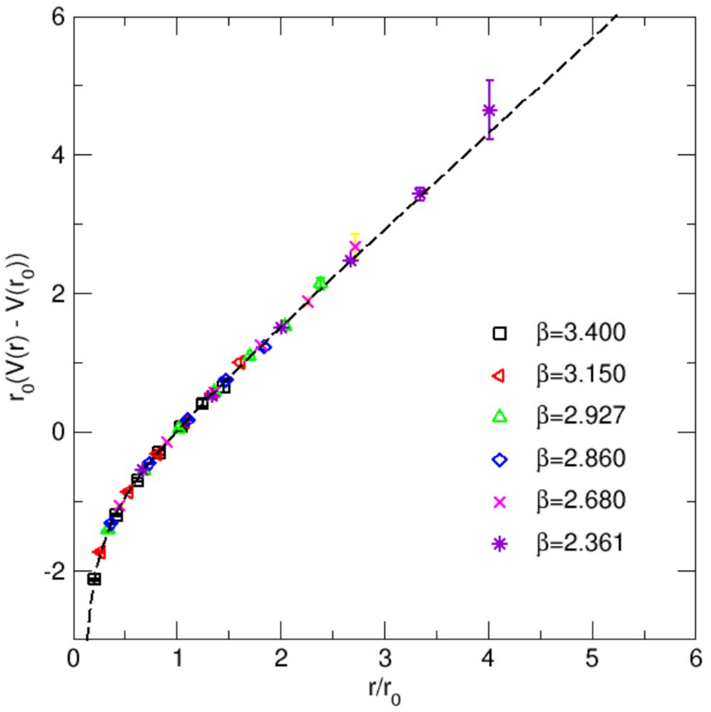
Renormalization group flow in infinite-dimensional coupling space.

Renormalized trajectory (RT) at weak coupling close to FP action.

Fixed point (FP) action keeps values of critical couplings for finite values of gauge coupling $\beta = 2N_c / g^2$

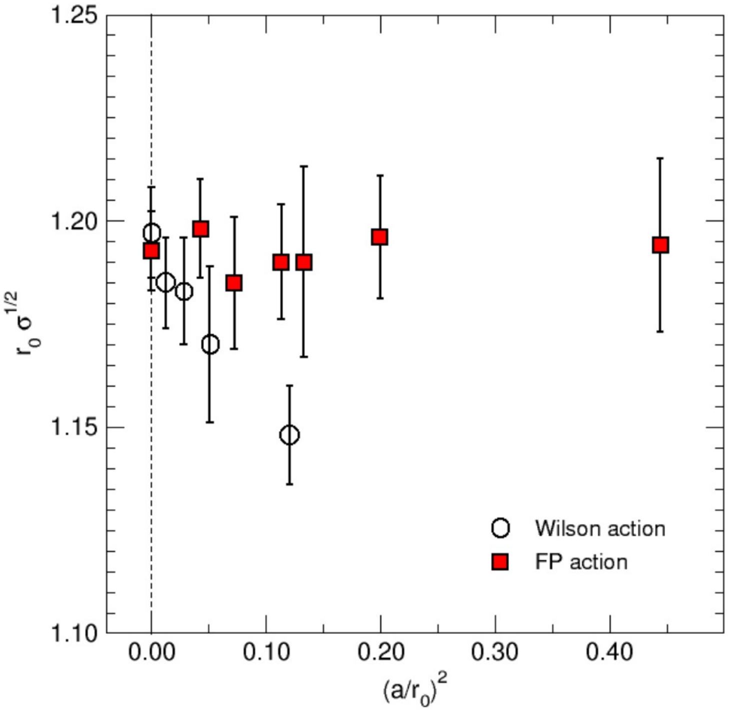
Fixed point action using older parametrizations

Static quark-antiquark potential



fit to $V(r) = V_0 - \frac{\alpha}{r} + \sigma r$

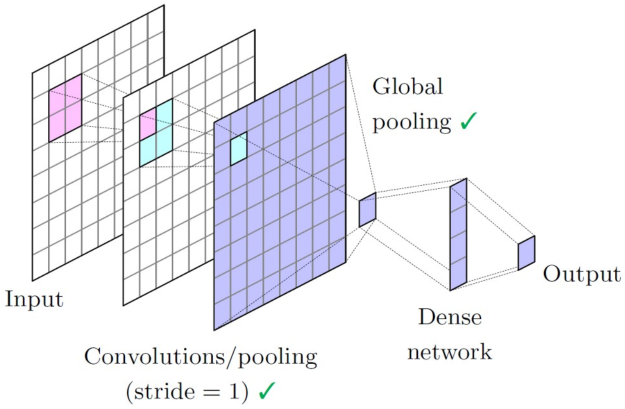
Extraction of string tension σ



Niedermayer, Rüfenacht, Wenger, Nucl.Phys.B 597 (2001) 413, hep-lat/0007007

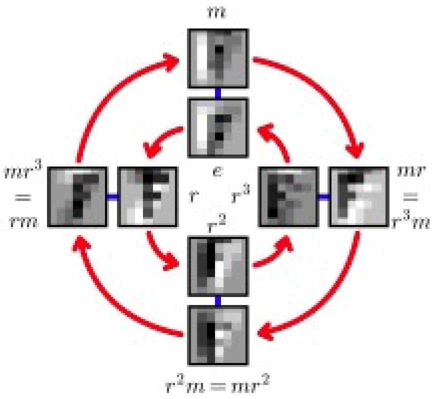
Symmetries on the lattice

Translational symmetry
→ Convolutional neural networks (CNNs)



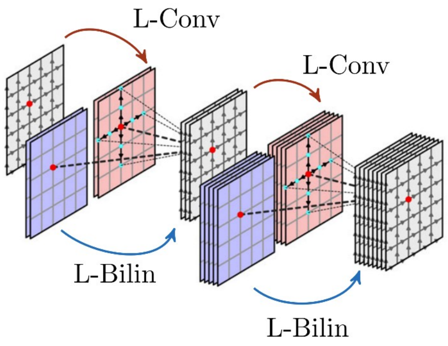
Bulusu, Favoni, AI, Müller, Schuh,
Phys. Rev. D 104 (2021) 074504

Rotation, mirror symmetry
→ Group equivariant CNNs (G-CNNs)



Cohen, Welling, ICML 2016

Lattice gauge symmetry
→ Lattice gauge equivariant CNNs (L-CNNs)



Favoni, AI, Müller, Schuh,
Phys.Rev.Lett. 128 (2022) 032003

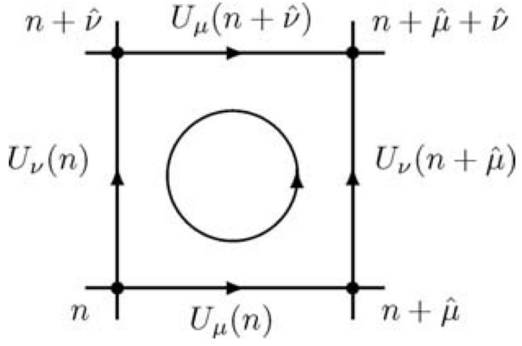
L-CNN data

Combine lattice links *U*
and locally transforming objects *W*

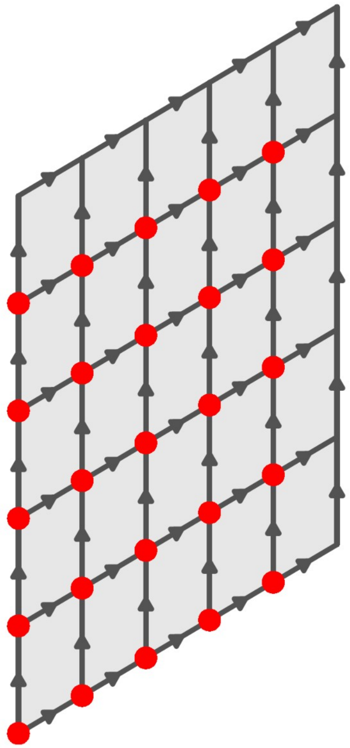
$$\text{tuple}\ \ (\mathcal{U}, \mathcal{W})$$

$$\mathcal{U} = \{U_{x,\mu}\}\ \ \mathrm{SU}(N_c)\ \text{matrices}$$

$$\mathcal{W} = \{W_{x,i}\}\ \text{with}\ W_{x,i} \in \mathbb{C}^{N_c \times N_c}$$



from: Gattringer, Lang (2010)



$$\mathcal{U} = \{U_{\mathbf{x},\mu}\}$$

$$\mathcal{W} = \{W_{\mathbf{x},\mu\nu}\}$$

Gauge transformation

$$T_\Omega U_{x,\mu} = \Omega_x U_{x,\mu} \Omega^\dagger_{x+\mu}$$

$$T_\Omega W_{x,i} = \Omega_x W_{x,i} \Omega^\dagger_x$$

Gauge equivariant (gauge covariant) function

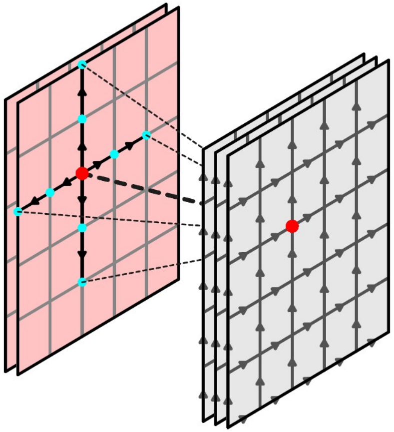
$$f(T_\Omega \mathcal{U}, T_\Omega \mathcal{W}) = T'_\Omega f(\mathcal{U}, \mathcal{W})$$

Gauge invariant function

$$f(T_\Omega \mathcal{U}, T_\Omega \mathcal{W}) = f(\mathcal{U}, \mathcal{W})$$

Lattice gauge equivariant layers

Convolution (L-Conv)

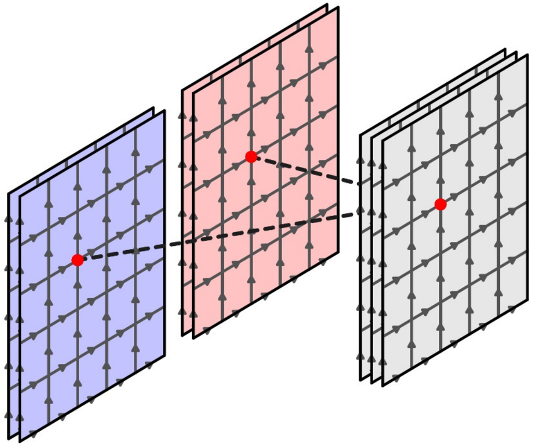


Convolution wish shared weights and proper parallel transport along coordinate axes

$$(\mathcal{U}, \mathcal{W}) \to (\mathcal{U}, \mathcal{W}')$$

$$W'_{\mathbf{x},i} = \sum_{j,\mu,k} \omega_{i,j,\mu,k} U_{\mathbf{x},k\cdot\mu} W_{\mathbf{x}+k\cdot\mu,j} U^{\dagger}_{\mathbf{x},k\cdot\mu}$$

Bilinear layer (L-Bilin)

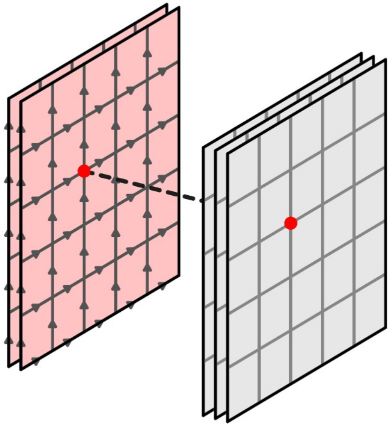


Multiply *W* at each lattice point

$$(\mathcal{U}, \mathcal{W}) \times (\mathcal{U}, \mathcal{W}') \to (\mathcal{U}, \mathcal{W}'')$$

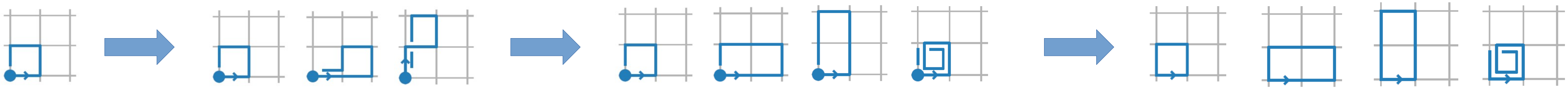
$$W''_{\mathbf{x},i} = \sum_{j,k} \alpha_{ijk} W_{\mathbf{x},j} W'_{\mathbf{x},k}$$

Trace layer

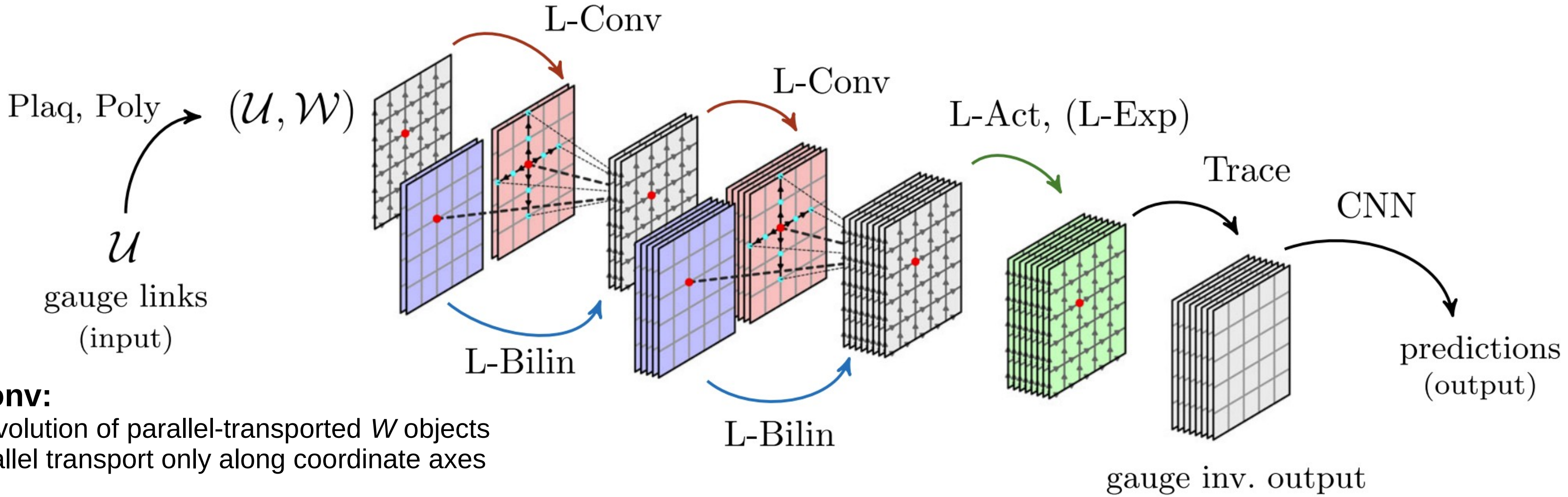


Generate gauge invariant output

$$w_{\mathbf{x},i} = \operatorname{Tr} W_{\mathbf{x},i} \in \mathbb{C}$$



Generic L-CNN



L-Conv:
* convolution of parallel-transported *W* objects
* parallel transport only along coordinate axes

L-Bilin:
* bilinear layer, product of locally transforming objects

L-Act:
* activation functions multiply *W* objects by scalar, gauge-invariant functions

L-Exp:
* update link variables using exponential map

Trace:
* calculate gauge invariant trace

Plaq:
* generate all possible plaquettes

Poly:
* generate all possible Polyakov loops

Favoni, AI, Müller, Schuh,
Phys.Rev.Lett. 128 (2022) 032003

Fixed point action using L-CNNs

Parametrize **action model** in particular way:

$$\mathcal{A}^{\text{L-CNN}}[V] = \sum_x \mathcal{A}_x^{\text{pre}}[V] \sum_{n=0}^{\infty} b^{(n)} (N_x[V] - N_x[\mathbb{1}])^n$$

Prefactor controls continuum behavior

L-CNN

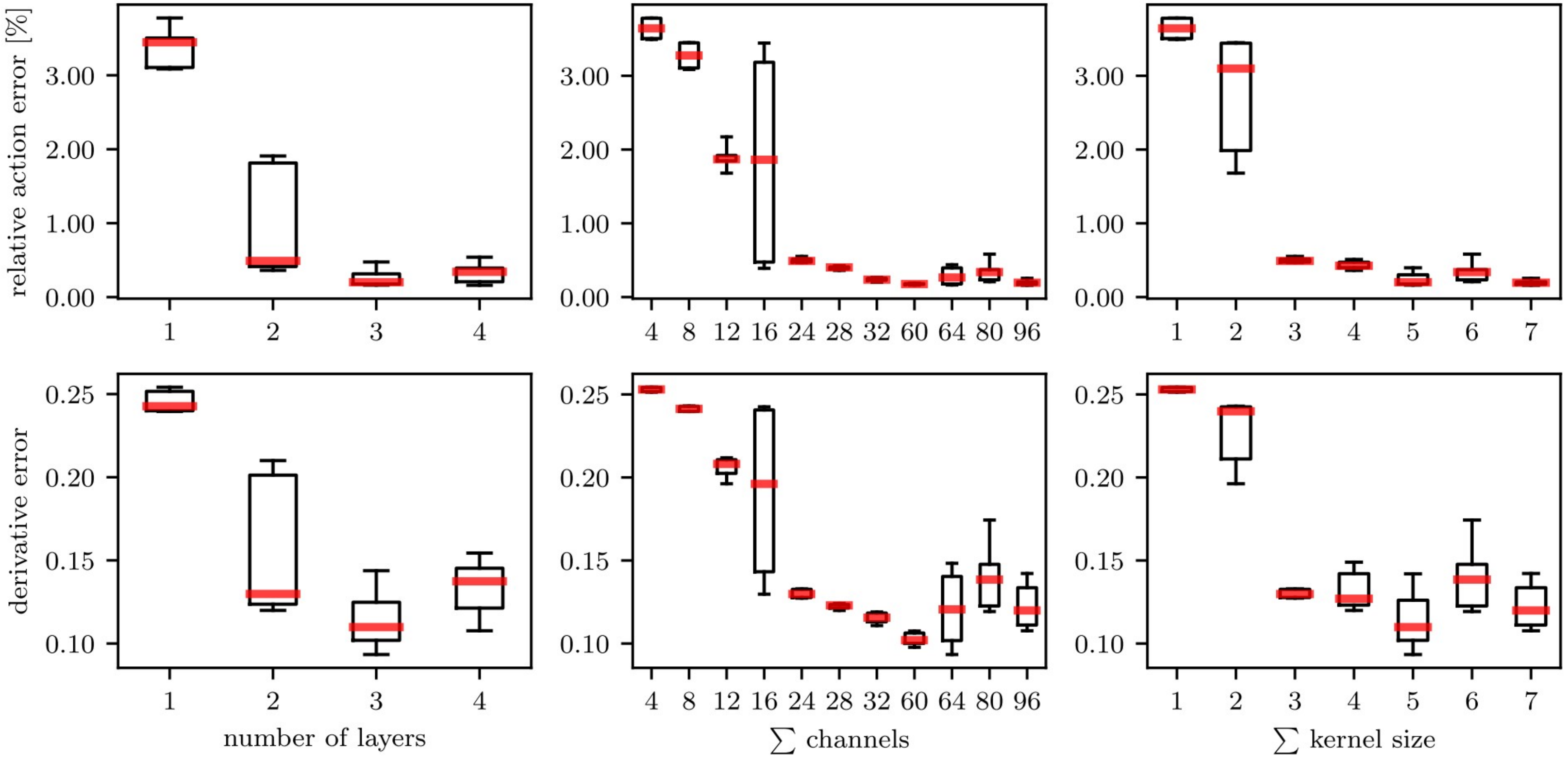
Loss function combines action values and its derivatives $\mathcal{L} = w_1 \mathcal{L}_1 + w_2 \mathcal{L}_2$

$$\mathcal{L}_1 = \frac{1}{L^4 N_{\text{cfg}}} \sum_{i=1}^{N_{\text{cfg}}} |\mathcal{A}^{\text{FP}}[V_i] - \mathcal{A}^{\text{L-CNN}}[V_i]|,$$

$$\mathcal{L}_2 = \frac{1}{32 L^4 N_{\text{cfg}}} \sum_{i=1}^{N_{\text{cfg}}} \sum_{x,\mu} \text{Tr}\left[(D_{x,\mu}^{\text{FP}}[V_i] - D_{x,\mu}^{\text{L-CNN}}[V_i])^2 \right]$$

Technical remark: derivatives of L-CNNs are obtained through backpropagation

Scan through various architectures



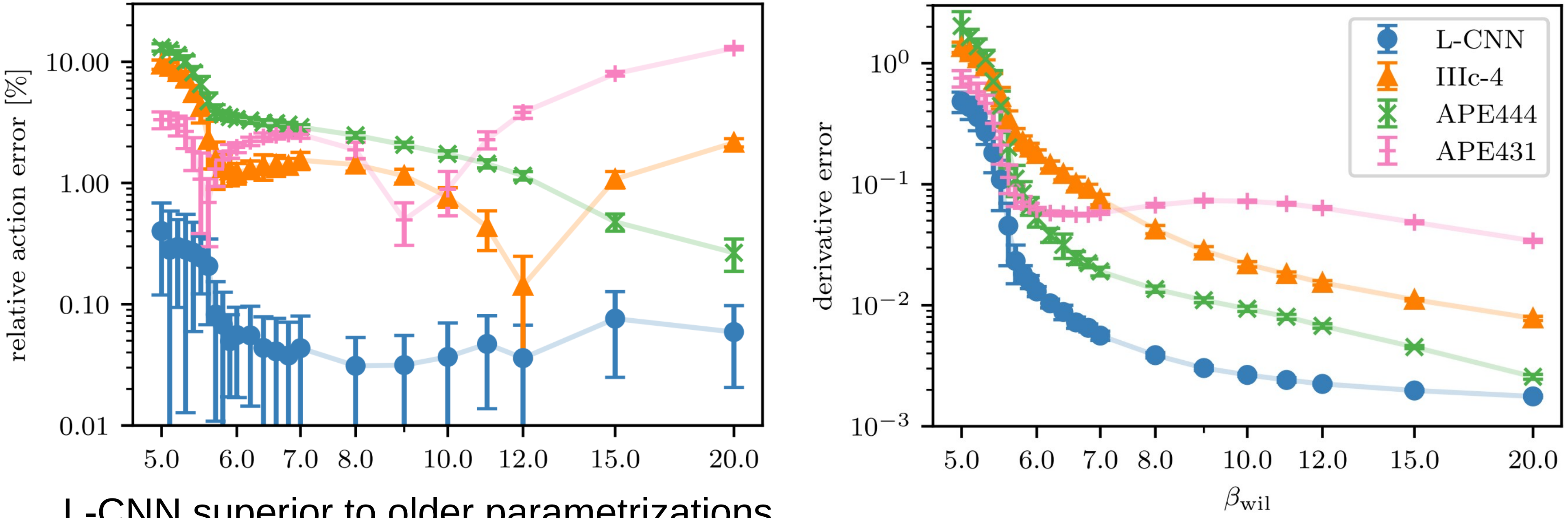
Supervised learning from coarse configurations and corresponding minimized action values on fine configurations.

Also use derivatives of fixed point action for learning.

Holland, Al, Müller, Wenger, Phys.Rev.D 110 (2024) 7, 074502

Train 130 models of various sizes for 4^4 lattice, SU(3) gauge group, and $\beta_{\mathrm{wil}} \in [5,10]$.

Learning the fixed point action with L-CNNs



L-CNN superior to older parametrizations of FP action.

Best model: L-CNN with
3 layers with 12, 24, 24 channels
and kernel size 2, 2, 1.

L-CNN: Holland, AI, Müller, Wenger, Phys.Rev.D 110 (2024) 7, 074502

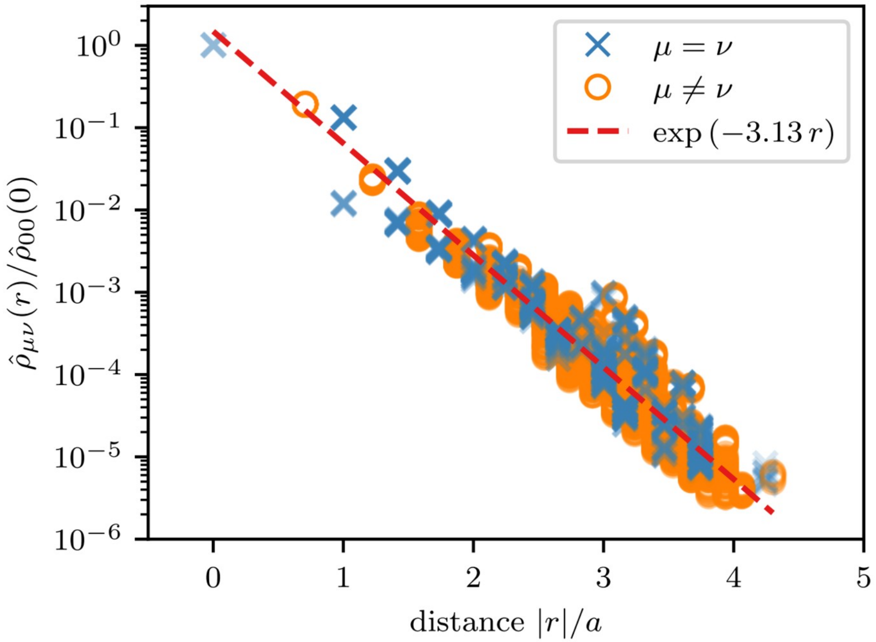
APE444, APE431: Niedermayer, Rüfenacht, Wenger (2000)

Type IIIa, IIIb, IIIc: Blatter, Niedermayer (1996)

Quadratic, Type I, II: DeGrand, Hasenfratz, Blatter, Niedermayer (1995)

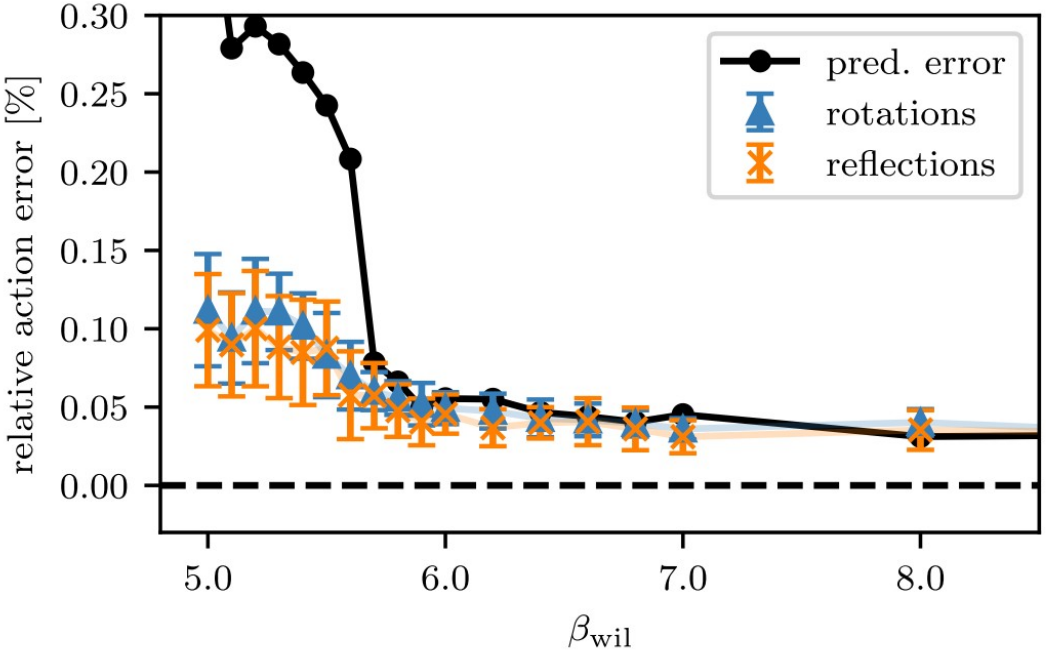
Properties of the learned FP action

Locality measurement



Couplings of learned fixed point actions decrease exponentially with separation.

Rotation and reflection error



Learned symmetry errors are smaller than or comparable to prediction errors.

Holland, AI, Müller, Wenger,
Phys.Rev.D 110 (2024) 7, 074502

Hybrid Monte Carlo

$$H(p,U) = \frac{p^2}{2} + \mathcal{A}(U)$$

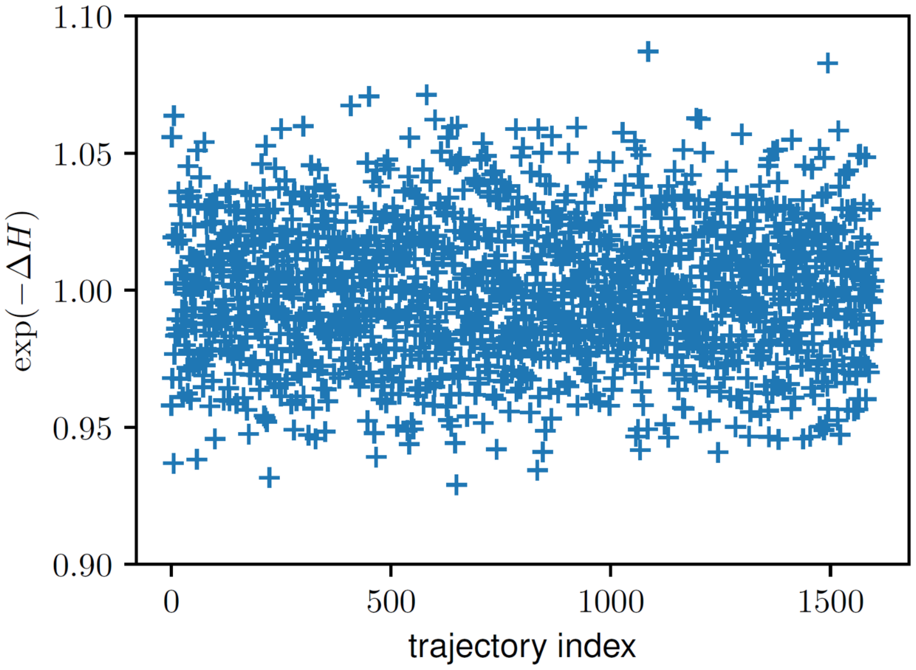
add momentum p

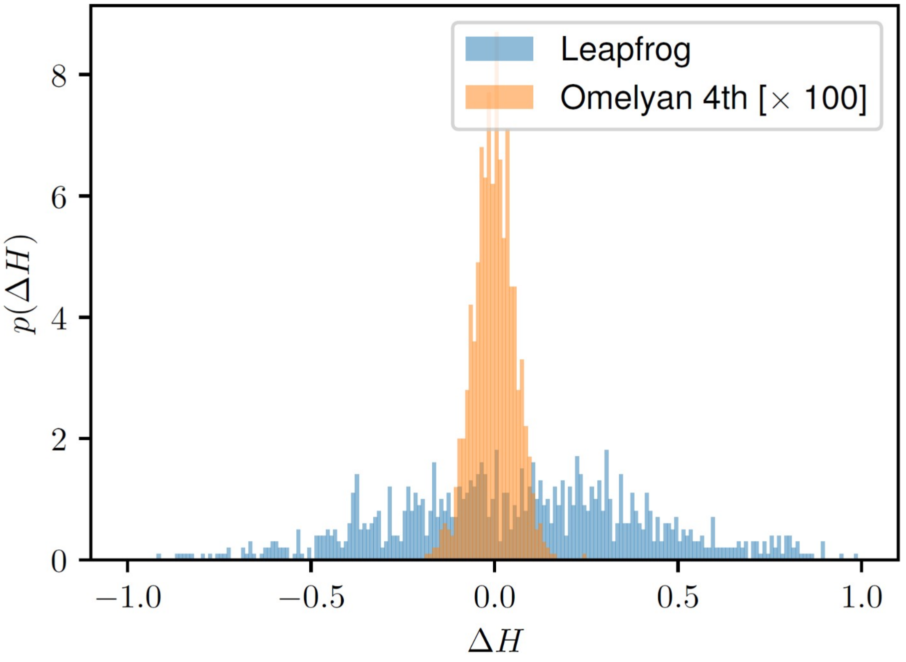
$$\frac{dU}{dt} = \frac{\partial H}{\partial p} = p$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial U} = -\frac{\partial \mathcal{A}}{\partial U}$$

- sample momenta p
- integrate eqs. of motion (leapfrog, Omelyan)
- correct for H non-conservation

$$\langle \exp(-\Delta H) \rangle = 1$$





Gradient flow

Gradient flow (GF) in the continuum:

$$\frac{dA_\mu}{dt} = D_\nu G_{\nu\mu} \qquad \text{with } A_\mu(t)$$

$$G_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

Observable:

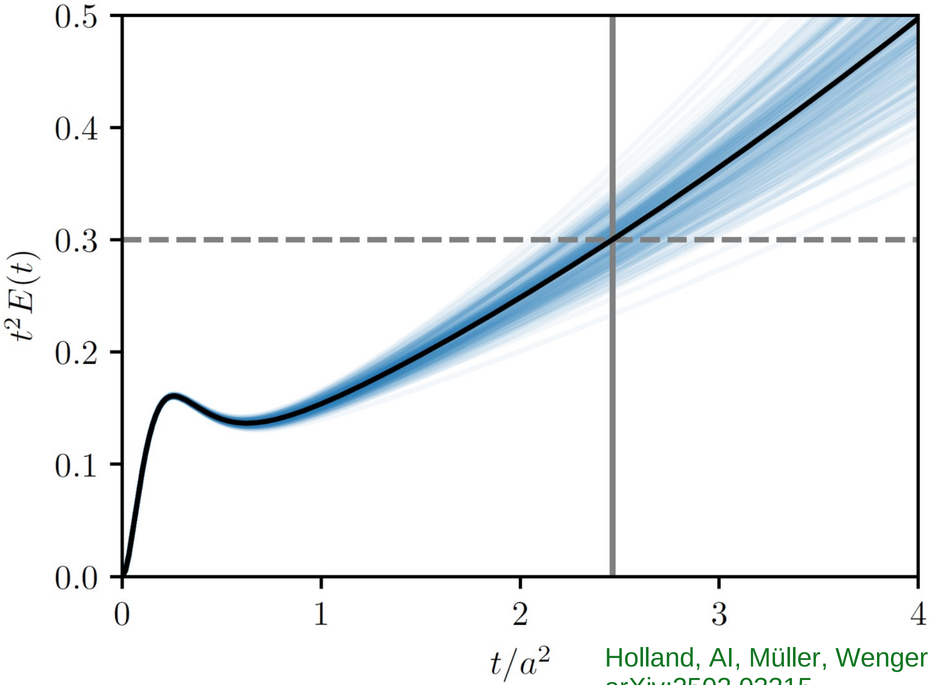
$$E = \frac{1}{4} G^a_{\mu\nu} G^a_{\mu\nu}$$

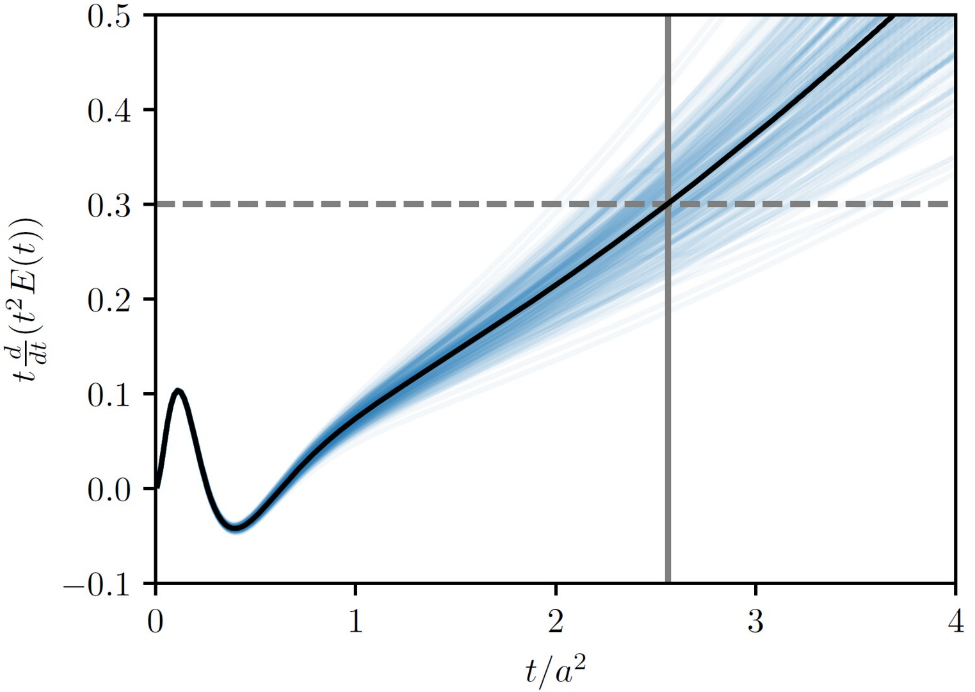
Lüscher JHEP 08 (2010) 071

Scale definitions:

$$t^2 \langle E \rangle|_{t=t_{0.3}} = 0.3$$

$$t^2 \frac{d}{dt}\left(t^2 \langle E \rangle\right)|_{t=w^2_{0.3}} = 0.3$$





Holland, Al, Müller, Wenger
arXiv:2502.03315

HMC + Gradient Flow with learned fixed point action

Hybrid Monte Carlo (HMC) with learned fixed point (FP) action.

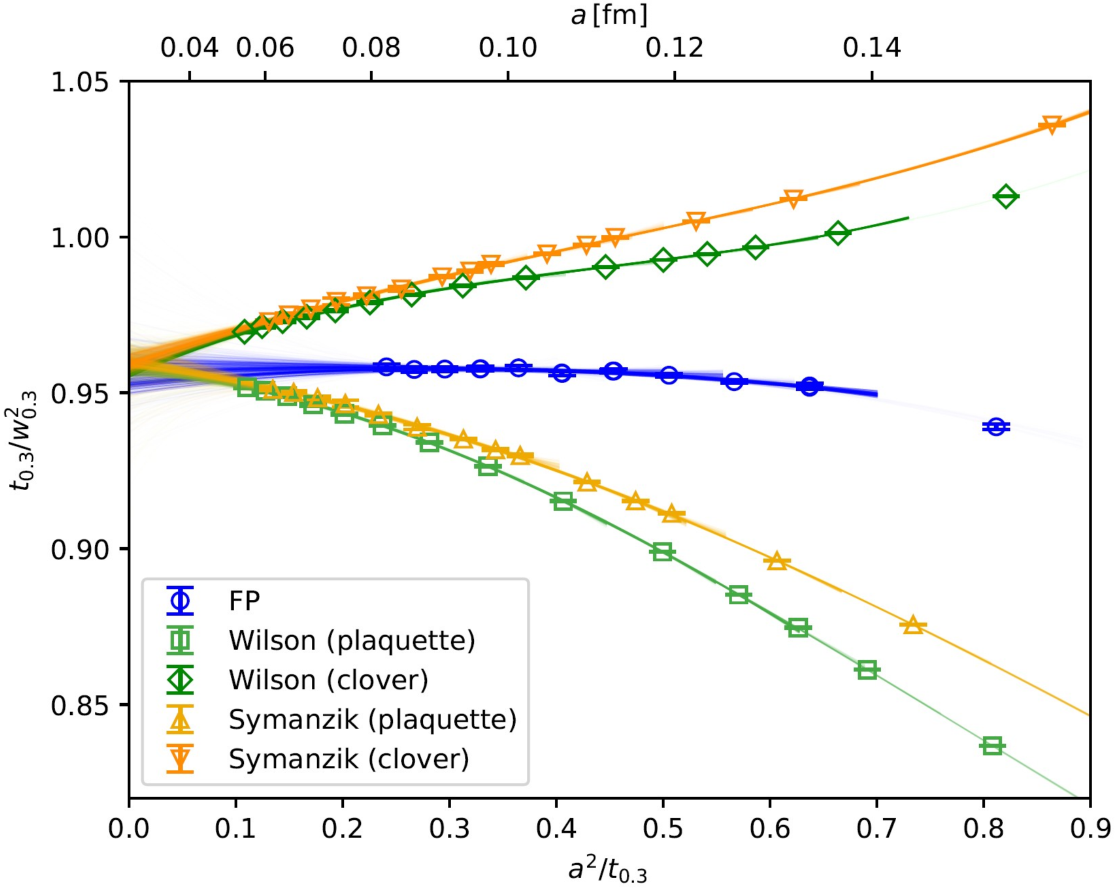
Gradient flow (GF) with learned FP action.
(GF with FP action is classically perfect!)

GF scale definitions:

$$t^2 \langle E \rangle |_{t=t_{0.3}} = 0.3$$

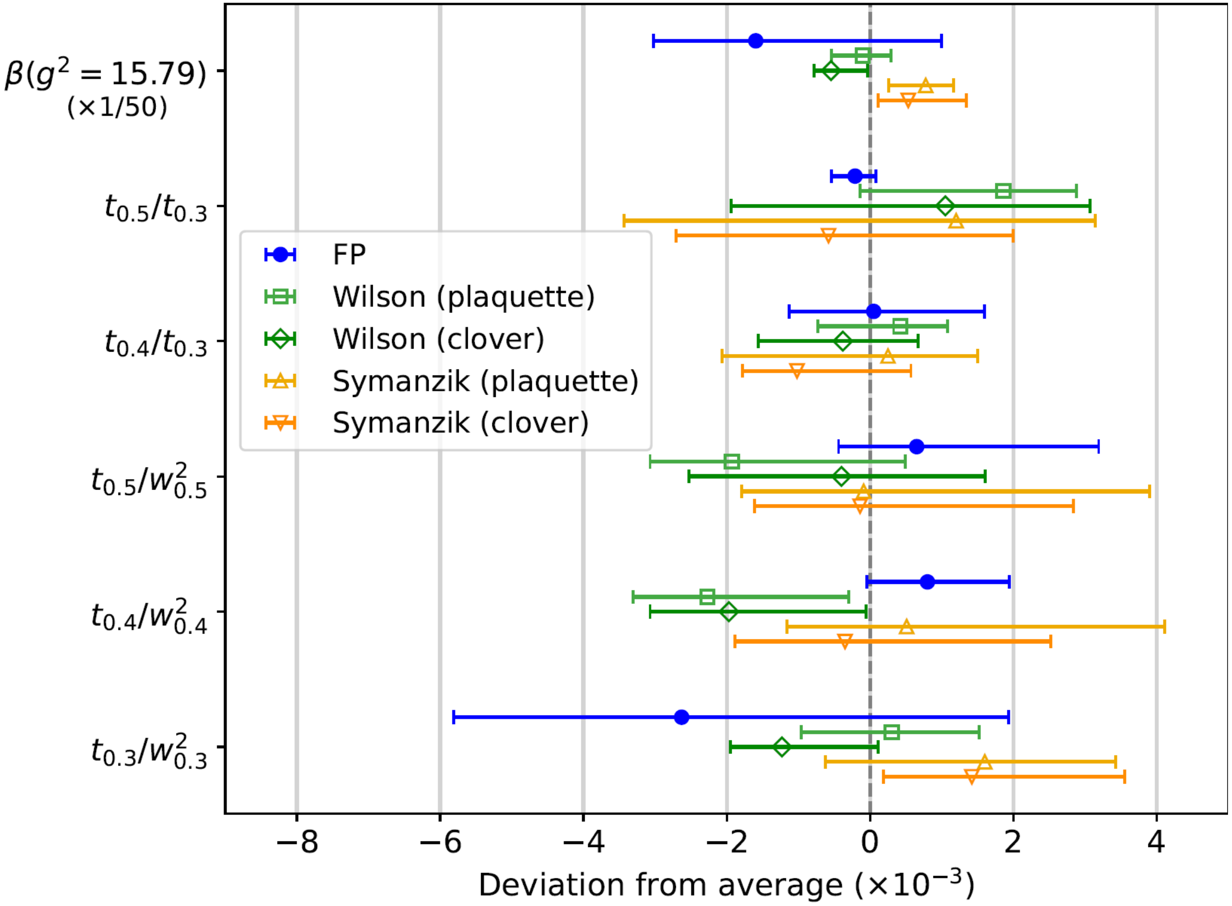
$$t^2 \frac{d}{dt} \left(t^2 \langle E \rangle \right) |_{t=w_{0.3}^2} = 0.3$$

Holland, AI, Müller, Wenger,
Phys.Rev.D 110 (2024) 7, 074502



Holland, AI, Müller, Wenger,
arXiv:2504.15870

Comparison of various continuum predictions



Fixed point (FP) action shows very good consistency with best results from Wilson or Symanzik improved lattice actions for 4D SU(3).

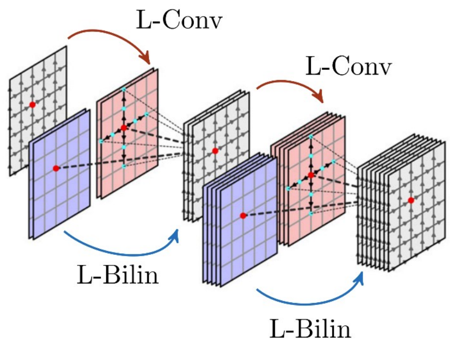
Holland, Al, Müller, Wenger, arXiv:2504.15870

Summary

- L-CNNs achieve higher accuracy than previous hand-crafted parametrizations
- Fixed-point (FP) actions define a classically perfect gradient flow
- FP actions reduces lattice artifacts in gradient-flow observables

Outlook:

- Measure various quantities (critical temperature, glueballs) on the lattice using the FP action
- Construct quantum perfect actions



Open source: https://gitlab.com/openpixi/lge-cnn