HMC and gradient flow with machine-learned classically perfect fixed-point actions

K. Holland,[a] **A. Ipp,**[b] **D. I. Müller**[b] **and U. Wenger**[c,*]

[c] *Institute for Theoretical Physics, Albert Einstein Center for Fundamental Physics, University of Bern, CH-3012 Bern, Switzerland*

[b] *Institute for Theoretical Physics, TU Wien, Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria*

[a] *University of the Pacific, 3601 Pacific Ave., Stockton, CA 95211, USA*

E-mail: wenger@itp.unibe.ch

Fixed-point (FP) lattice actions are classically perfect, i.e., they have continuum classical properties unaffected by discretization effects and are expected to have suppressed lattice artifacts at weak coupling. Therefore they provide a possible way to extract continuum physics with coarser lattices, allowing to circumvent problems with critical slowing down and topological freezing towards the continuum limit. We use machine-learning methods to parameterize a FP action for four-dimensional SU(3) gauge theory using lattice gauge-covariant convolutional neural networks. The large operator space allows us to find superior parameterizations compared to previous studies and we show how such actions can be efficiently simulated with the Hybrid Monte Carlo algorithm. Furthermore, we argue that FP lattice actions can be used to define a classically perfect gradient flow without any lattice artifacts at tree level. We present initial results for scaling of the gradient flow with the FP action.

The 41st International Symposium on Lattice Field Theory (LATTICE2024)
*28 July - 3 August 2024*
Liverpool, UK

---

[*]Speaker

1. Introduction

One of the most difficult challenges in lattice QCD simulations is the critical slowing down towards the continuum limit where the lattice spacing a is taken to zero while keeping some physical scales fixed, cf. Fig. 1 for an illustration. The critical behaviour manifests itself in long autocorrelation times in the measurement of physical observables [1], for example gradient-flow quantities or the topological charge. The slow decorrelation in the latter quantity is usually referred to as topological freezing. While the problem is absent at sufficiently large lattice spacings, simulations at coarse resolution in general lead to large lattice artifacts and correspondingly large systematic errors due to the long extrapolation to $a \rightarrow 0$, unless one uses highly improved lattice discretizations of the underlying theory.

One way to construct highly improved lattice actions is through renormalization group transformations (RGT). The aim of an RGT is to describe the physics of a fine lattice on a coarse lattice by means of adjusting the effective couplings of the discretized theory, cf. Fig. 1, such that the resulting effective lattice action maintains the low-energy physics exactly.
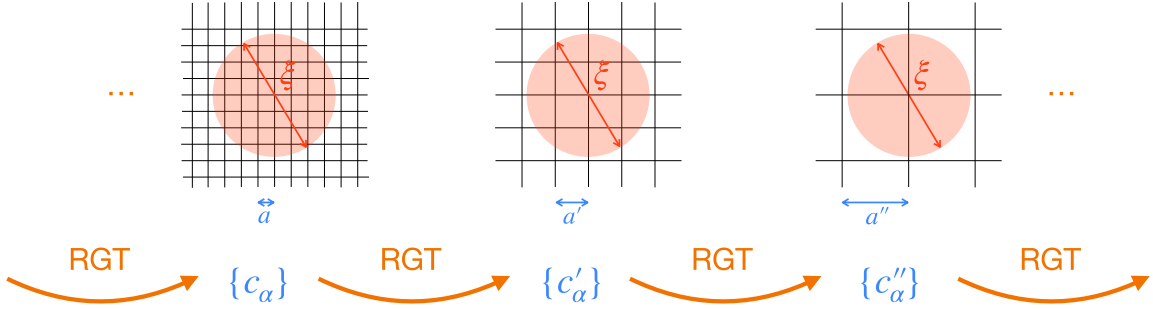
Figure 1: Illustration of the continuum limit towards the left with lattice spacings $a/\xi < a'/\xi < a''/\xi$ at fixed physical scale ξ (*upper row*), and renormalization group transformation (RGT) steps yielding effective couplings $\{c_\alpha\} \rightarrow \{c'_\alpha\} \rightarrow \{c''_\alpha\}$ (*lower row*).

In these proceedings we describe the status of our project to construct an RG-improved lattice action which is classically perfect, i.e., which has no lattice discretization artifacts at tree level in the gauge coupling, and to employ it in Hybrid Monte Carlo (HMC) simulations. Previous reports of this ongoing project were given in [2, 3].

2. Renormalization group transformations and fixed-point actions

For gauge theories, an RGT for a lattice action $\mathcal{A}$ from a fine to a coarse lattice can be defined through

$$\exp(-\beta' \mathcal{A}'[V]) = \int \mathcal{D}U \exp(-\beta\{\mathcal{A}[U] + T[U,V]\}), \qquad (1)$$

where the integration on the RHS is over gauge-field configurations U on the fine lattice restricted by the RG blocking kernel $T[U,V]$. In this process, the gauge coupling β and the action $\mathcal{A}$ parameterized by the couplings $\{c_\alpha\}$ on the fine lattice are mapped to β' and $\mathcal{A}'$, i.e., $\{c'_\alpha\}$, on the coarse lattice. The RG blocking is defined as follows. First, a smeared link S_μ^{smeared} is constructed from the fine gauge links by adding up the original link $U_\mu(x)$ with weight s_0, the planar staples
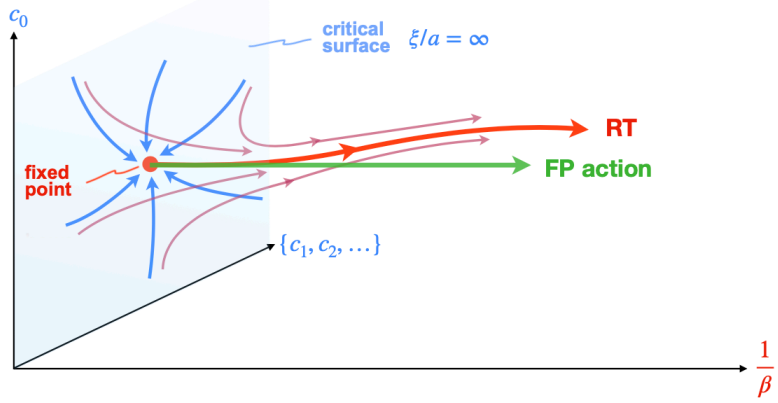
2
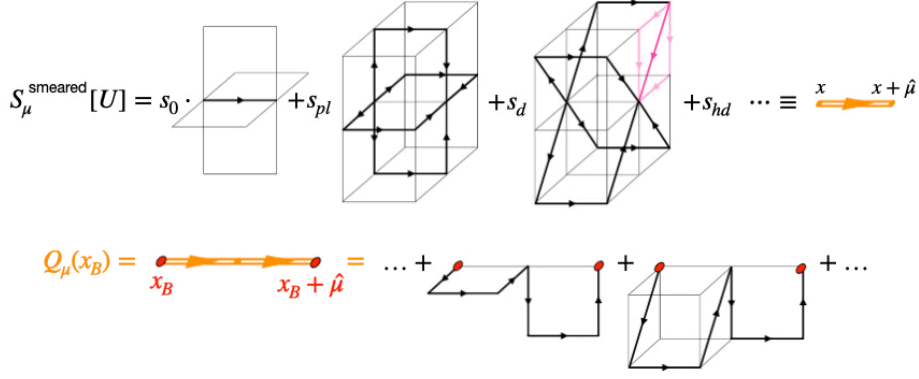
Figure 2: Illustration of the RGT blocking employed here.

with weight s_{pl}, diagonal staples with weight s_d, and hyper-diagonal staples with weight s_{hd}, as illustrated in the top row of Fig. 2. Two smeared links are then multiplied together to construct the blocked link $Q_\mu(x_B)$ generating a combination of many gauge-link paths within 16 hypercubes of the fine lattice, cf. bottom row of Fig. 2. Finally, the gauge invariant blocking kernel $T[U,V]$ is defined by

$$T[U,V] = -\kappa \sum_{n_B,\mu} \left\{ \mathrm{ReTr}\left(V_{n_B,\mu} Q^\dagger_{n_B,\mu}[U] \right) - \mathcal{N}^\beta_\mu[U] \right\}, \tag{2}$$

where the term $\mathcal{N}^\beta_\mu[U]$ guarantees the correct normalization of the partition function under the RGT. The parameters $s_0, s_{pl}, s_d, s_{hd}, \kappa$ of the blocking kernel can be chosen to optimize the locality of the resulting effective gauge action. For asymptotically free theories, there is one marginally relevant operator parameterized by β and hence the critical surface, where $\xi/a = \infty$, is located at $\beta = \infty$. Under repeated RGT steps, any lattice action defined on the critical surface is driven to the Fixed Point (FP), cf. Fig. 3. Lattice actions slightly off the critical surface flow to-


Figure 3: Illustration of the RG flow for asymptotically free gauge theories.

wards the Renormalized Trajectory (RT) which emerges from the FP along the relevant direction. Actions on the RT are effective actions with no lattice artifacts at all.

The FP at $\beta \to \infty$ is defined by the FP equation [4]

$$\mathcal{A}^{\mathrm{FP}}[V] = \min_U \left(\mathcal{A}^{\mathrm{FP}}[U] + T[U,V] \right), \tag{3}$$

and using $\mathcal{A}^{\mathrm{FP}}$ with fixed couplings $\{c^{\mathrm{FP}}_\alpha\}$ at finite β defines the straight line in Fig. 3. It is expected to stay close to the RT at weak coupling and hence to have suppressed lattice artifacts. While the
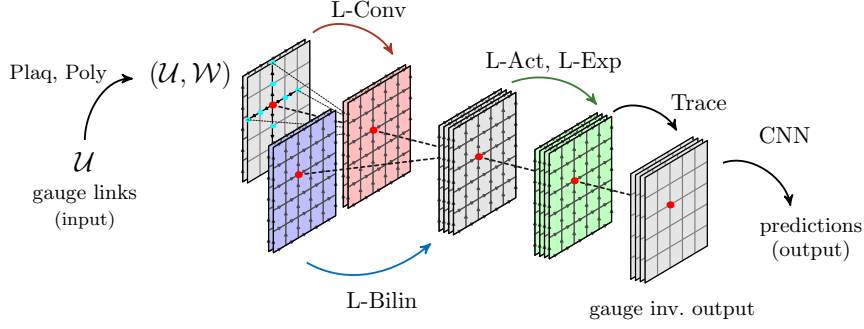
Figure 4: Example of an L-CNN that processes lattice gauge data while preserving gauge symmetry through specialized convolutional (L-Conv), bilinear (L-Bilin), and activation layers [5].

FP Eq. (3) implicitly provides data for each coarse configuration V, finding the explicit form of the FP action, i.e., the couplings $\{c_\alpha^{\text{FP}}\}$, is a difficult parameterization problem. It can be addressed efficiently using convolutional neural networks and machine-learning techniques.

3. Neural network parameterization and HMC simulation

Since the FP action in general involves infinitely many gauge invariant operators with their corresponding couplings $\{c_\alpha^{\text{FP}}\}$, any useful parameterization requires high flexibility while maintaining exact gauge invariance. The lattice gauge-covariant convolutional neural network (L-CNN) architecture described in Ref. [5] is well suited for this task. Taking as input a gauge field $\mathcal{U}$ and gauge covariant objects $\mathcal{W}$, e.g., untraced Wilson plaquettes and Polyakov loops, the network produces arbitrarily complicated loops through convolutional layers with parallel transports, while additional bilinear layers combine products of covariant objects locally, retaining the gauge symmetry at each step. Finally, a trace layer produces gauge invariant output, cf. Fig. 4 for an illustration. We employ machine-learning techniques to train the network to reproduce the FP action values given by the RHS of Eq. (3) by parameterizing the FP action on the LHS. In addition, the derivatives of the FP action with respect to the gauge links are also determined by the FP equation, and they provide $8 \times 4 \times (\text{volume})$ data per gauge configuration. The L-CNN generates these derivatives exactly through backpropagation, hence these large datasets are essential in training the network.

As our best parameterization, we use an L-CNN consisting of three layers with 12, 24, 24 channels and kernel sizes 2, 2, 1. In Fig. 5 we show the comparison with previous parameterizations based on traced loops up to length six and up to four powers of the traces (IIIc-4) [6], and parameterizations based on powers of traced plaquettes constructed from hypercubically APE-smeared fat links (APE444, APE431) [7]. To test the quality of different parameterizations, we take as input ensembles of coarse configurations generated with the Wilson gauge action at a range of β_{wil} values to cover a broad span of lattice spacings, where numerical minimization on the RHS of Eq. (3) provides the underlying true action values. The left panel shows the relative error of the parameterized FP action values on these ensembles at each β_{wil} indicated on the x-axis, while the right panel shows the size of the parameterization errors of the derivatives. We find that the L-CNN consistently performs better than the previous parameterizations all the way from coarse to fine lattice spacing.
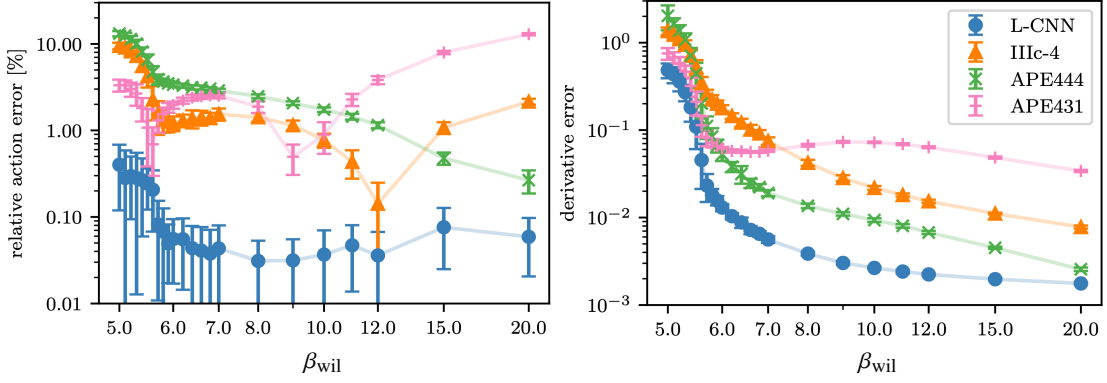
Figure 5: Comparison of the new parameterization of the FP action (L-CNN) and previous parameterizations using either powers of traced loops up to length six (IIIc-4) or powers of traced plaquettes containing APE-smeared fat links (APE 431, APE 444), from Ref. [2].

While an accurate parameterization is a necessary first step in the FP approach, there is no guarantee that the lattice artifacts will in fact be small when this action is used for nonperturbative calculations. This can only be tested with an actual scaling study involving MC simulations
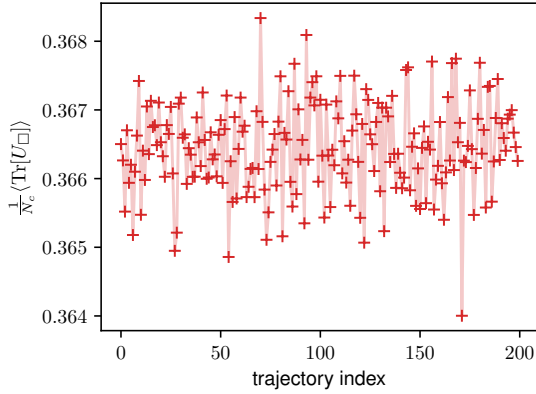
Figure 6: HMC evolution of the plaquette value for the FP action at $\beta_{FP} = 2.95$ corresponding to $a \simeq 0.11$ fm.

at a range of β values, to make quantitative comparison with other lattice actions and the continuum limit of appropriately chosen quantities. Given that the derivatives of the FP action with respect to gauge links are automatically provided by the L-CNN through backpropagation, it is natural to choose the Hybrid Monte Carlo (HMC) algorithm [8, 9] for simulations with the FP action, adding fictitious conjugate momenta to evolve the gauge links according to the equations of motion. Standard techniques for accelerating the HMC algorithm can be used for the FP action, for example we use the 4MN4FP variant [10] of the 4th order Omelyan integrator [11] to allow for a coarser time step when solving the equations of motion. Among the tests of our implementation, we check that the Hamiltonian constraint $\langle \exp(-\Delta H) \rangle = 1$ is satisfied by ensembles generated with the HMC algorithm. We show in Fig. 6 an example of the HMC evolution of the plaquette values for a stream of 200 trajectories obtained by simulating the FP action at $\beta_{FP} = 2.95$, which corresponds to a lattice spacing of $a \simeq 0.11$ fm.

4. Classically perfect gradient flow

A precise observable is necessary to test the quality of improvement of the FP action, without introducing additional lattice artifacts. This is naturally given by the gradient flow (GF) of the gauge fields [12, 13]. In continuum language, given the gauge action $\mathcal{A} = (1/g^2) \int d^4x (-\frac{1}{2} \text{Tr} F_{\mu\nu} F_{\mu\nu})$,

with the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ and gauge coupling g, the non-Abelian gauge fields $A_\mu(t)$ evolve with a fictitious time t as

$$\frac{dA_\mu(t)}{dt} = -\frac{\delta\mathcal{A}}{\delta A_\mu(t)}, \tag{4}$$

where the flow starts with the original gauge field $A_\mu(t = 0)$ and the flow time t has dimension $(length)^2$. The resulting action density of flowed gauge fields $E(t) = -\frac{1}{2}\text{Tr}F_{\mu\nu}F_{\mu\nu}$ is renormalized, which allows a number of connections to physical quantities. For example, perturbatively the flowed action density for SU(N) gauge theory is

$$t^2\langle E(t)\rangle = \frac{3(N^2-1)g^2}{128\pi^2}(1 + O(g^2)), \tag{5}$$

where g is the renormalized gauge coupling in the $\overline{\text{MS}}$ scheme with RG scale $\mu = 1/\sqrt{8t}$. Hence a specific value of $g^2(t)$ implicitly defines a value for the flow time t in physical units, which allows the scale of lattice simulations to be determined. Another use is the gauge coupling β-function defined through the GF scheme, e.g., for SU(3) gauge theory

$$t^2\langle E(t)\rangle \doteq \frac{3g_{\text{GF}}^2(t)}{16\pi^2}, \tag{6}$$

where the β-function $\mu^2 \cdot dg_{\text{GF}}^2/d(\mu^2) = -t \cdot dg_{\text{GF}}^2/dt$ can be measured directly from the GF on gauge configurations generated at chosen values of the coupling [14, 15]. In general, GF quantities can be measured with high accuracy, making them good probes of discretization effects in lattice simulations.

On the lattice, the flow is of the gauge links, with $dU_\mu/dt = -i(\delta\mathcal{A}^f/\delta U_\mu)U_\mu$ for some choice $\mathcal{A}^f$ of lattice flow action. In addition, one has a separate choice of lattice action $\mathcal{A}^e$ for the discretized action density observable $E(t)$. The gauge configurations are generated via MC simulation, where there is a third choice of discretized action $\mathcal{A}^g$. The collective discretization artifacts of these choices can be calculated perturbatively [16], with

$$\begin{aligned}
t^2\langle E(t)\rangle &= \frac{3(N^2-1)g_0^2}{128\pi^2}\left[C(a^2/t) + O(g_0^2)\right], \\
C(a^2/t) &= \frac{64\pi^2t^2}{3}\int_{-\pi/a}^{\pi/a}\frac{d^4p}{(2\pi)^4}\text{Tr}\left[e^{-t(\mathcal{A}^f+\mathcal{G})}(\mathcal{A}^g+\mathcal{G})^{-1}e^{-t(\mathcal{A}^f+\mathcal{G})}\mathcal{A}^e\right],
\end{aligned} \tag{7}$$

where $C(a^2/t)$ contains the tree-level lattice artifacts as deviations from 1, the action density in momentum space to quadratic order in the fields is $A_\mu(p)\mathcal{A}_{\mu\nu}(p)A_\nu(-p)$, g_0 is the bare gauge coupling and $\mathcal{G}$ is a gauge fixing term. One can exploit the freedom to choose different gauge actions for $\mathcal{A}^e, \mathcal{A}^f$ and $\mathcal{A}^g$ to minimize the artifacts in a tree-level improved renormalized gauge coupling [16]. On the other hand, using the same lattice action for all three gives

$$C(a^2/t) = \frac{64\pi^2t^2}{3}\int_{-\pi/a}^{\pi/a}\frac{d^4p}{(2\pi)^4}\text{Tr}[e^{-2t(\mathcal{A}+\mathcal{G})}]. \tag{8}$$

As a pedagogical guide, examples of the tree-level coefficient $C(a^2/t)$ for two different regulators are shown in Fig. 7. One is a hard-momentum cutoff $\pm\pi/a$ with the continuum dispersion
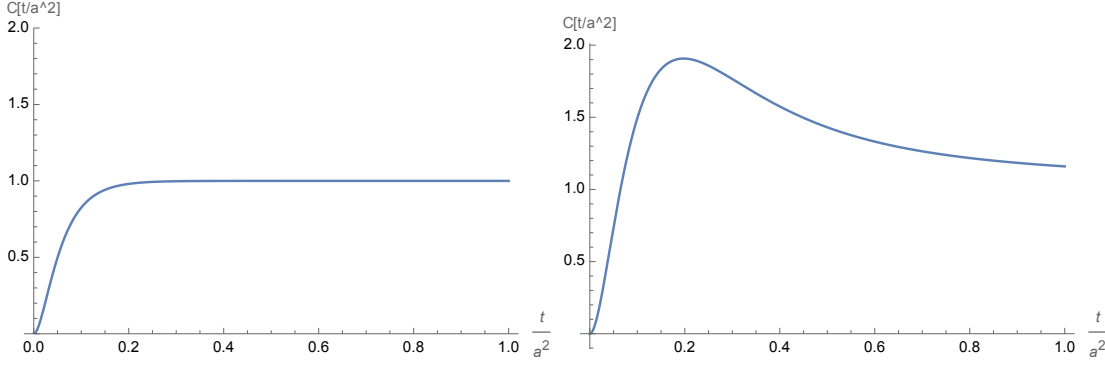
Figure 7: The tree-level coefficient $C(a^2/t)$ in the gauge field gradient flow for different regulators, with a hard momentum cutoff (*left*) or Wilson lattice action (*right*). The exact FP propagator extends the momentum range to $\pm\infty$ even at finite lattice spacing, resulting in $C^{\mathrm{FP}}(a^2/t) = 1$ with no artifacts.

relation, the second is the Wilson lattice gauge action for which $C(a^2/t)$ is analytically calculable in terms of the Bessel function I_0. For both, the coefficient approaches 1 in the limit $t/a^2 \to \infty$.

For the exact FP action, the FP propagator is substituted into Eq. (8). At any finite lattice spacing, the gauge configuration can be connected back to an arbitrarily fine initial configuration through multiple RG transformations, where the propagator is also RG blocked. Although on the lattice the momenta are restricted to $-\pi/a \le p_\mu \le \pi/a$, the iterated RG transformations result in the FP propagator having poles at $(p + 2\pi l)^2$ for all integers l. Hence the momentum range is extended to $\pm\infty$ and the exact FP propagator has the continuum dispersion relation. This results in $C^{\mathrm{FP}}(a^2/t) = 1$, i.e., the exact FP action has no tree-level artifacts in the gradient flow. This new finding is in keeping with FP actions preserving classical properties exactly on the lattice, e.g., the continuum dispersion relation, exact classical (instanton) solutions and invariant topological charge under RG transformations, and chirality of FP fermions as a solution of the Ginsparg-Wilson relation [17–19].
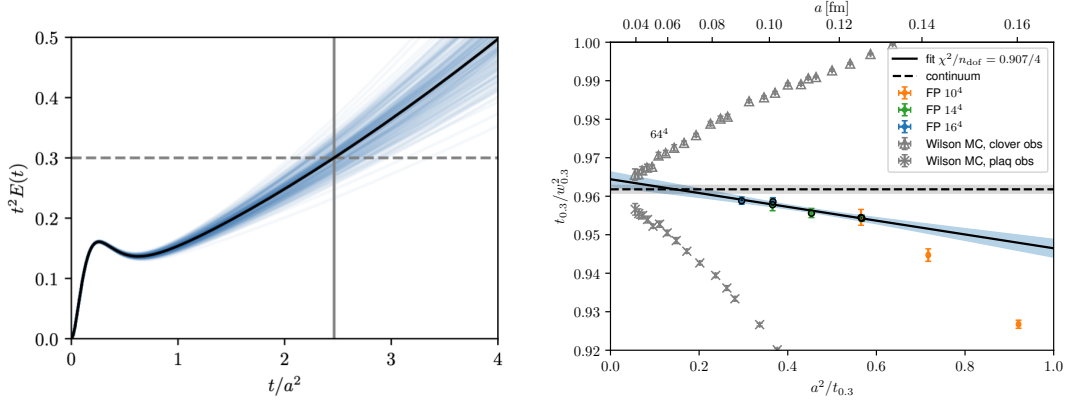
5. Scaling of gradient-flow scales

Having generated MC ensembles with the FP action, we integrate the flow equation on each configuration with the same FP action using a third-order Runge-Kutta scheme [12]. In the left plot of Fig. 8 we show the resulting flow of the action density for one FP gauge ensemble as an example. Physical reference scales are defined through the conditions [12, 20]

$$t^2 \langle E \rangle |_{t=t_{0.3}} = 0.3, \qquad t \frac{d}{dt} \left(t^2 \langle E \rangle \right) \bigg|_{t=w_{0.3}^2} = 0.3 \,, \tag{9}$$

and we use the scaling of a dimensionless ratio of these scales to probe the FP action for any remaining lattice artifacts. With the classically perfect GF at hand, any such lattice artifacts are due either to an imperfect parameterization or to quantum effects. In the right plot of Fig. 8 we show the scaling of the ratio $t_{0.3}/w_{0.3}^2$ towards the continuum limit $a^2/t_{0.3} \to 0$ for the parameterized FP action and compare it with that of the Wilson gauge action using either the plaquette or clover discretization of the observable $E(t)$. We use $w_{0.3} = 0.1755$ fm [20] for a visual guide of the

Figure 8: (*Left*) Example of gradient flow on one FP gauge ensemble, the light blue curves are from individual configurations, the black curve is the ensemble average. (*Right*) Scaling test of FP simulation results for the combination of scales $t_{0.3}/w_{0.3}^2$.

lattice spacing. Discretization effects in the parameterized FP action are very mild and an $O(a^2)$ extrapolation of the data with $a \lesssim 0.13$ fm is sufficient for a reliable continuum limit. This is in contrast to the Wilson gauge action which requires simulations at substantially smaller lattice spacings for an equally reliable extrapolation.

6. Conclusions and outlook

In this work, we have demonstrated how a fixed-point lattice gauge action can be parameterized using modern machine-learning tools, argued that the exact FP gradient flow is free of tree-level lattice artifacts, and presented preliminary results for a scaling test of the parameterized FP action through MC simulations and measurements of the gradient flow in four-dimensional SU(3) gauge theory. Lattice gauge-covariant convolutional neural networks (L-CNNs) are an essential element to achieve a superior parameterization of the FP action with exact gauge invariance. The initial results show that lattice artifacts in the parameterized FP action are much reduced, making it broadly useful for any gauge theory study, allowing continuum properties to be extracted from simulations on coarser lattices.

7. Acknowledgments

References

[1] ALPHA collaboration, S. Schaefer, R. Sommer and F. Virotta, *Critical slowing down and error analysis in lattice QCD simulations*, *Nucl. Phys. B* **845** (2011) 93–119, [1009.5228].

[2] K. Holland, A. Ipp, D. I. Müller and U. Wenger, *Machine learning a fixed point action for SU(3) gauge theory with a gauge equivariant convolutional neural network*, *Phys. Rev. D* **110** (2024) 074502, [2401.06481].

[3] K. Holland, A. Ipp, D. I. Müller and U. Wenger, *Fixed point actions from convolutional neural networks*, *PoS* **LATTICE2023** (2024) 038, [2311.17816].

[4] P. Hasenfratz and F. Niedermayer, *Perfect lattice action for asymptotically free theories*, *Nucl. Phys. B* **414** (1994) 785–814, [hep-lat/9308004].

[5] M. Favoni, A. Ipp, D. I. Müller and D. Schuh, *Lattice Gauge Equivariant Convolutional Neural Networks*, *Phys. Rev. Lett.* **128** (2022) 032003, [2012.12901].

[6] M. Blatter and F. Niedermayer, *New fixed point action for SU(3) lattice gauge theory*, *Nucl. Phys. B* **482** (1996) 286–304, [hep-lat/9605017].

[7] F. Niedermayer, P. Rufenacht and U. Wenger, *Fixed point gauge actions with fat links: Scaling and glueballs*, *Nucl. Phys. B* **597** (2001) 413–450, [hep-lat/0007007].

[8] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, *Hybrid Monte Carlo*, *Phys. Lett. B* **195** (1987) 216–222.

[9] S. A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken and R. L. Sugar, *Hybrid Molecular Dynamics Algorithms for the Numerical Simulation of Quantum Chromodynamics*, *Phys. Rev. D* **35** (1987) 2531–2542.

[10] T. Takaishi and P. de Forcrand, *Testing and tuning new symplectic integrators for hybrid Monte Carlo algorithm in lattice QCD*, *Phys. Rev. E* **73** (2006) 036706, [hep-lat/0505020].

[11] I. P. Omelyan, I. M. Mryglod and R. Folk, *Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations*, *Comput. Phys. Commun.* **151** (2003) 272–314.

[12] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, *JHEP* **08** (2010) 071, [1006.4518].

[13] M. Luscher and P. Weisz, *Perturbative analysis of the gradient flow in non-abelian gauge theories*, *JHEP* **02** (2011) 051, [1101.0963].

[14] Z. Fodor, K. Holland, J. Kuti, D. Nogradi and C. H. Wong, *A new method for the beta function in the chiral symmetry broken phase*, *EPJ Web Conf.* **175** (2018) 08027, [1711.04833].

[15] A. Hasenfratz and O. Witzel, *Continuous renormalization group β function from lattice simulations*, *Phys. Rev. D* **101** (2020) 034514, [`1910.06408`].

[16] Z. Fodor, K. Holland, J. Kuti, S. Mondal, D. Nogradi and C. H. Wong, *The lattice gradient flow at tree-level and its improvement*, *JHEP* **09** (2014) 018, [`1406.0827`].

[17] T. A. DeGrand, A. Hasenfratz, P. Hasenfratz and F. Niedermayer, *The Classically perfect fixed point action for SU(3) gauge theory*, *Nucl. Phys. B* **454** (1995) 587–614, [`hep-lat/9506030`].

[18] P. Hasenfratz, *Lattice QCD without tuning, mixing and current renormalization*, *Nucl. Phys. B* **525** (1998) 401–409, [`hep-lat/9802007`].

[19] P. Hasenfratz, V. Laliena and F. Niedermayer, *The Index theorem in QCD with a finite cutoff*, *Phys. Lett. B* **427** (1998) 125–131, [`hep-lat/9801021`].

[20] BMW collaboration, S. Borsányi, S. Dürr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg et al., *High-precision scale setting in lattice QCD*, *JHEP* **09** (2012) 010, [`1203.4469`].