Machine-learned RG-improved gauge actions and classically perfect gradient flows

Kieran Holland,
1, * Andreas Ipp,
2, † David I. Müller,
2, ‡ and Urs Wenger
3, §

¹University of the Pacific, 3601 Pacific Avenue, Stockton, California 95211, USA

²Institute for Theoretical Physics, TU Wien, Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria

Albert Einstein Center for Fundamental Physics, Institute for Theoretical Physics,

University of Bern, Sidlerstraße 5, CH-3012 Bern, Switzerland

Extracting continuum properties of quantum field theories from discretized spacetime is challenging due to lattice artifacts. Renormalization-group (RG)-improved lattice actions can preserve continuum properties, but are in general difficult to parameterize. Machine learning (ML) with gauge-equivariant convolutional neural networks provides a way to efficiently describe such actions. We test a machine-learned RG-improved lattice gauge action, the classically perfect fixed-point (FP) action, for four-dimensional SU(3) gauge theory through Monte Carlo simulations. We establish that the gradient flow of the FP action is free of tree-level discretization effects to all orders in the lattice spacing, making it classically perfect. This allows us to test the quality of improvement of the FP action, without introducing additional artifacts. We find that discretization effects in gradient-flow observables are highly suppressed and less than 1% up to lattice spacings of 0.14 fm, allowing continuum physics to be extracted from coarse lattices. The quality of improvement achieved motivates the use of the FP action in future gauge theory studies. The advantages of ML-based parameterizations also highlight the possibility of realizing quantum perfect actions in lattice gauge theory.

Introduction — The Standard Model of particle physics gives excellent agreement between experimental measurements and theoretical predictions to very high precision. A central element is Quantum Chromodynamics (QCD), the quantum field theory of quark-gluon interactions, where Monte Carlo (MC) simulations play a crucial role in testing nonperturbative dynamics [1]. A necessary step in lattice simulations is spacetime discretization, which ultimately must be removed to obtain physical continuum predictions. However, the extrapolation to the continuum limit is often the dominant source of uncertainty in lattice studies and remains a serious challenge: reducing the lattice spacing $a \to 0$, simulations become less efficient [2], requiring longer Markov chains for high statistical accuracy (*critical slowing down*) and to fully explore the phase space (topological freezing). These systematic effects are a barrier to increasing the precision of lattice predictions for the Standard Model. To overcome this hurdle, one can use improved lattice actions [3–5], designed to remove the leading lattice artifacts, often through a perturbative expansion in a. The efficacy of the improvement must be empirically tested via MC simulations. The problem has also stimulated the study of various alternatives, such as normalizing flows, which use machine learning (ML) to transform random samples to target distributions [6–14]. However, normalizing flows face significant challenges, such as achieving high efficiency in four-dimensional quantum field theories—measured by effective sample size—and ensuring that models trained at one lattice spacing can generalize effectively to generate ensembles at other lattice spacings. Combined with the persistent difficulties of extrapolating to the continuum limit, these challenges underscore the urgent need for new methods that can achieve accurate results with small lattice artifacts already on coarse lattices, thus evading critical slowing down and topological freezing.

In this Letter, we employ an alternative approach. based on fixed-point (FP) lattice actions [15] which are implicitly and nonperturbatively defined using the properties of Wilson's renormalization group (RG) for asymptotically free theories [16–18]. It is used to extract continuum quantities from coarse lattices with significantly reduced lattice artifacts. We demonstrate the effectiveness of the FP action in MC simulations of four-dimensional SU(3) gauge-field theory, parameterized using a convolutional neural network with exact gauge invariance [19]. This result is made possible by our new theoretical insight, presented here, that FP actions yield a classically perfect gradient flow, i.e., without tree-level lattice artifacts to all orders in the lattice spacing. This result also builds on the success of our recent work [20–22], where we demonstrated that ML-based parameterizations can accurately capture the FP action across a wide range of gauge-field fluctuations—from very fine to very coarse lattice spacings. The resulting parameterization achieves an average relative error in the action value of less than 0.2%, representing an improvement by over an order of magnitude compared to previous approaches developed earlier [23–25]. The ML framework also enables accurate parameterization of the derivatives of the FP action w.r.t. the gauge fields, which are crucial for both efficient network training and the practical implementations of MC algorithms. The FP action can be generally used in any lattice-QCD study and provides a new way to reduce discretization artifacts on coarse lattices, making it broadly useful.

Renormalization group and fixed-point actions — For SU(N) gauge theories, an RG transformation of a lattice action \mathcal{A} from a fine to a coarse lattice can be defined

through

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

where $\beta = 2N/g^2$ and the fine lattice configuration U is connected to the coarse configuration V through the RG blocking kernel

$$\mathcal{T}[U,V] = -\kappa \sum_{n_B,\mu} \left\{ \operatorname{ReTr}(V_{n_B,\mu} Q_{n_B,\mu}^{\dagger}[U]) - \mathcal{N}_{\mu}^{\infty} \right\}.$$
(2)

The variables $Q_{n_B,\mu}$ are blocked links constructed from the fine links $U_{n,\mu}$ and the normalization constant $\mathcal{N}^{\infty}_{\mu} = \max_{W \in \mathrm{SU}(N)} \left[\operatorname{ReTr}(WQ^{\dagger}_{n_B,\mu}) \right]$ guarantees the invariance of the partition function after an RG blocking step. The blocking details and parameter κ can be freely chosen to maximize the compactness of the FP action [24] and in this work we follow the RG blocking therein.

In the limit $1/g^2 \to \infty$ of asymptotically free quantum field theories, the FP of the RG is scale-invariant and defines the continuum limit of the gauge theory. Moreover, the path integral on the RHS of Eq. (1) reduces to a saddle point problem yielding the classical FP equation [15]

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

The resulting FP action \mathcal{A}^{FP} is *classically perfect*, i.e., it has no lattice artifacts at tree level and exactly reproduces continuum classical properties at finite lattice spacing, e.g., dispersion relation [15], topology [23, 26], or index theorem [27], chirality [28, 29] when fermions are included. A key theoretical result of this work is the derivation that the gradient flow of the FP action is also classically perfect.

Classically perfect gradient flow — To rigorously test the quality of improvement in any lattice action, it is essential to use precisely measurable observables that avoid introducing significant cutoff effects. Observables based on the gradient flow (GF) of gauge fields [30] are a natural choice, given their small systematics and high statistical accuracy. In addition, due to the renormalization properties of the GF [31], these observables are widely used for the setting of physical scales and the measurement of the β -function of the gauge coupling. We show here that the GF based on the FP action has the additional desirable property of being classically perfect: it has no tree-level discretization effects to any order in the lattice spacing a.

In continuum language, for an action \mathcal{A} the GF evolves the non-Abelian gauge field $A_{\mu}(t)$ along a fictitious flow time t according to [30–33]

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

where the starting value $A_{\mu}(t=0)$ is the original gauge field and the action is $\mathcal{A} = (1/g^2) \int d^4x (-\frac{1}{2} \text{Tr} F_{\mu\nu} F_{\mu\nu})$, with field strength $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} + [A_{\mu}, A_{\nu}]$ and gauge coupling g. The action density $E(t) = -\frac{1}{2}\text{Tr} F_{\mu\nu}F_{\mu\nu}$ at flow time t is a renormalized quantity, and for SU(N) gauge theory it can be connected perturbatively to a renormalized coupling g in a given scheme, e.g., the $\overline{\text{MS}}$ scheme, as [31]

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

where $\langle \cdot \rangle$ denotes the gauge-field average.

On the lattice, the perturbative expansion is in the bare coupling g_0 with [34]

$$t^{2}\langle E(t)\rangle = \frac{3(N^{2}-1)g_{0}^{2}}{128\pi^{2}} \left(C(a^{2}/t) + \mathcal{O}(g_{0}^{2})\right), \quad (6)$$

where

$$C(a^2/t) = \frac{64\pi^2 t^2}{3} \times \int_{-\pi/a}^{\pi/a} \frac{d^4 p}{(2\pi)^4} \operatorname{Tr} \left[e^{-t(\mathcal{S}^f + \mathcal{G})} (\mathcal{S}^g + \mathcal{G})^{-1} e^{-t(\mathcal{S}^f + \mathcal{G})} \mathcal{S}^e \right]$$
(7)

contains the tree-level lattice artifacts as deviations from 1. Here, the action density in momentum space to quadratic order in the fields is $A_{\mu}(p)S_{\mu\nu}(p)A_{\nu}(-p)$, and S^e, S^f, S^g correspond to the (possibly different) choices of the discretized action for the action-density observable E, for the GF, and for the MC simulation, respectively, while \mathcal{G} is a gauge fixing term. For any choice of action discretizations, the value of $C(a^2/t)$ can be calculated and used to define a tree-level improved coupling at finite lattice spacing [34] (see also [35]). Using the same lattice action for all three gives

$$C(a^2/t) = \frac{64\pi^2 t^2}{3} \int_{-\pi/a}^{\pi/a} \frac{d^4 p}{(2\pi)^4} \operatorname{Tr}\left[e^{-2t(\mathcal{S}+\mathcal{G})}\right], \quad (8)$$

since the various factors under the trace commute with each other.

For the exact FP action \mathcal{A}^{FP} , the gluon propagator has poles $1/(p + 2\pi l)^2$ for momenta $-\pi/a \leq p_{\mu} \leq \pi/a$ and all possible integers l [23], since the iterative RG connection to finer and finer lattices extends the momentum integration range to the full interval $\pm \infty$ even at finite lattice spacing, producing precisely the continuum dispersion relation. Hence, for \mathcal{S}^{FP} one obtains

$$C(a^2/t) = \frac{64\pi^2 t^2}{3} \cdot 3\left(\int_{-\infty}^{+\infty} \frac{dp}{2\pi} e^{-2tp^2}\right)^4 = \frac{64\pi^2 t^2}{(\sqrt{8\pi t})^4} = 1,$$

i.e., the exact FP action has no tree-level cutoff dependence in $t^2 \langle E(t) \rangle$; lattice artifacts appear in loop corrections, i.e., as quantum effects, at $\mathcal{O}(a^2 g_0^2)$. This is a new example of FP actions keeping continuum classical properties intact at finite lattice spacing.

Machine learning architecture — The practical limitations in using the FP action in MC simulations arise from two factors: the size of quantum lattice artifacts, and the inherent imperfections in the fixed-point parameterization for four-dimensional SU(3) gauge theory. The latter is addressed here using a gauge-equivariant convolutional neural network [19] and ML techniques. The ML architecture designed for the FP action is described in detail in [21]. In brief, the main building block of the neural network is a bilinear convolution which computes various products of Wilson loops, starting from 1×1 plaquettes and the links $U_{n,\mu}$ of a gauge-field configuration as input. Gauge covariance is exactly maintained throughout by accounting for parallel transport in the convolution. As the network depth increases, larger and more complex closed loops are generated. The final layer is a trace giving the action value as output. The network respects translational symmetry and is (ultra-)local, which allows it to be used on any lattice size. Additionally, our architecture is set up to reproduce a perturbatively defined FP action for smooth fields [24] in order to guarantee the correct continuum limit.

The goal of training a model is to approximate the action values according to the FP equation in Eq.(3). A crucial element is that the exact derivatives of \mathcal{A}^{FP} with respect to gauge links $U_{n,\mu}$ can be easily obtained from the FP equation, with each gauge configuration giving $8 \times 4 \times (\text{volume})$ data for learning. In the network, the derivatives can be calculated exactly via backpropagation. The loss function for training hence contains both action values and derivatives. The best architecture has been obtained from a hyperparameter scan and consists of three gauge-covariant convolutional layers. It has been trained on equilibrated configurations for a variety of lattice sizes and a large range of coupling parameters from very smooth to coarse gauge-field fluctuations, and has further been improved by finetuning on instanton configurations. Our best model reaches an average relative error in the action value of less than 0.2% with the exact derivatives more accurately reproduced than in previous parameterizations [25] across all lattice spacings. The fidelity of the model for very smooth configurations is particularly important for the GF because the gauge-field fluctuations become very smooth at large flow time.

Fixed-point action simulations — The availability of accurate gauge-field derivatives from our ML parameterization model makes the Hybrid Monte Carlo (HMC) algorithm a suitable choice for simulations, where artificial conjugate momenta are introduced to update the gauge links through molecular dynamics [36, 37]. The simulations can be accelerated with a coarser choice of time step when solving the equations of motion [38] using in particular the 4MN4FP variant of Omelyan integration scheme [39]. The GF for gauge links $dU_{\mu}/dt =$ $-i(\delta \mathcal{A}^{\rm FP}/\delta U_{\mu})U_{\mu}$ is solved using the third-order Runge-Kutta integrator [30]. The ML network is memory intensive; the largest lattice volumes we simulate are 18⁴, for which we can generate a new HMC trajectory in ~ 4 minutes on one NVidia A100 SXM6 64GB GPU. With the FP action, a smaller lattice volume is compensated by coarser lattice spacing to yield near-continuum-physics

results in large enough physical volume. As a consistency check of continuum predictions with the FP action, we also simulate four-dimensional SU(3) gauge theory using either Wilson or tree-level Symanzik improved gauge actions, measuring with Wilson GF, using the publicly available openqcd package [40].

Scaling tests — The GF provides several physical quantities that can be extracted. For example, a reference scale t_c/a^2 can be set through the condition $t_c^2 \langle E(t_c) \rangle = c$, where a common choice is c = 0.3 [30]. An alternative scale w_d/a is set through the derivative condition

$$t \cdot \frac{d}{dt} \left[t^2 \langle E(t) \rangle \right]_{t=t_d} = d, \qquad w_d = \sqrt{t_d}, \tag{9}$$

where d = 0.3 is also frequently used [41]. At finite lattice spacing, the values of t_c/a^2 and w_d/a are determined by the choice of bare coupling $6/g_0^2$ at which MC simulations are performed, as well as the action discretization used in the GF and the action density. Ratios of scales, e.g., t_c/w_d^2 or t_d/t_c , are physical properties, which have a welldefined value in the continuum limit independent of the discretizations employed. The necessary consistency in the continuum limit is a consequence of universality and provides a stringent test of the FP simulation results.

In addition, the renormalizability of the GF action density E(t) allows one to define a scheme for a renormalized coupling. For SU(3) gauge theory the GF scheme is defined by

$$t^2 \langle E(t) \rangle \doteq \frac{3g_{\rm GF}^2(t)}{16\pi^2},\tag{10}$$

where the GF renormalized coupling $g_{\rm GF}^2(t)$ runs with the RG scale $\mu = 1/\sqrt{8t}$. The β -function of the gauge coupling $\mu^2 \cdot dg_{\rm GF}^2/d(\mu^2) = -t \cdot dg_{\rm GF}^2/dt$ can be measured directly from the GF at the desired value of the coupling [42, 43]. For example, the choice c = 0.3 corresponds to strong coupling $g_{\rm GF}^2 \approx 15.79$. In the perturbative regime, it can be compared with the known 3-loop calculation [44].

Repeating the procedure for a set of gauge ensembles at different bare couplings, the continuum limit is taken as $a^2/t_c \rightarrow 0$. We present in Figures 1 and 2 results for the ratios $t_{0.3}/w_{0.3}^2$ and $t_{0.5}/t_{0.3}$, and the β -function $t \cdot dg_{\rm GF}^2/dt$ at fixed $g_{\rm GF}^2 \approx 15.79$, using FP, Wilson and Symanzik action simulations. Further examples can be found in the Supplemental Material. The leading lattice artifacts are expected to be $\mathcal{O}(a^2)$ for the Wilson data, $\mathcal{O}(a^4)$ for the Symanzik data, and $\mathcal{O}(g^2a^2)$ for the FP



FIG. 1. Continuum-limit extrapolations for the ratios $t_{0.3}/w_{0.3}^2$ and $t_{0.5}/t_{0.3}$. Results from Wilson and Symanzik MC simulations are shown using plaquette and clover discretizations of the action density.

data, and we show various possible fits (polynomial in $a^2/t_{0.3}$) describing the continuum-limit extrapolations.¹ The fit curves are shaded according to their Akaike Information Criterion (AIC) weight [45], cf. Supplemental Material.

The FP data for $t_{0.3}/w_{0.3}^2$ and the β -function show mild variation with lattice spacing, with artifacts smaller than 1% at $a \simeq 0.14$ fm, using $\sqrt{t_{0.3}} = 0.1679$ fm [46] to convert from lattice to physical scales. The ratio $t_{0.5}/t_{0.3}$ has essentially no lattice-spacing dependence below $a \lesssim 0.14$ fm using the FP action. By comparison, discretization errors with the Wilson or Symanzik action



FIG. 2. The β -function of the GF renormalized coupling at $g_{\rm GF}^2 = 15.79$ with highly suppressed lattice artifacts in the results for the FP compared to the Wilson and Symanzik gauge actions.

and either plaquette or clover version of the action density E are much larger, hence finer lattice spacings and larger lattice volumes are necessary to reach a scaling window of $\mathcal{O}(a^2)$ behavior while simultaneously keeping finite-volume effects small.

To estimate systematic effects in the continuum extrapolation, such as choosing the lattice-spacing range over which to fit or the order of the fit polynomial in $a^2/t_{0.3}$, we use a procedure based on the AIC, cf. Supplemental Material.

The advantageous features of the FP action are visible in other choices of c and d, cf. Supplemental Material for further examples. In each case, FP lattice artifacts are small, and in absolute value the FP continuum predictions are accurate to per mille level. We summarize a comparison of our continuum results using the FP, Wilson or Symanzik action in Figure 3, for values of c and d employed by other groups [47–54]. There is clear consistency between FP, Wilson and Symanzik results. Simulating on smaller coarser lattices, one can measure precisely the continuum properties of the gauge theory with the FP approach.

Conclusions and outlook — The explicit form of the exact RG-improved FP action is unknown and must be parameterized as a solution of the nonperturbative FP equation. Our previous study [21] demonstrated that ML networks can describe the FP action for four-dimensional SU(3) gauge theory to much higher accuracy than previously possible [24, 25, 55]. In this work, we show that the FP action defines a classically perfect gradient flow, i.e., without tree-level lattice artifacts to all orders in the lattice spacing. As a consequence, the FP action successfully reduces lattice artifacts in gradient-flow ob-

¹ We note that the leading $\mathcal{O}(a^4)$ behaviour of the Symanzik data is masked by the $\mathcal{O}(a^2)$ discretization effects from the flow action (Wilson) and the action density (plaquette and clover).



FIG. 3. The comparison of continuum predictions for fourdimensional SU(3) gauge theory from MC simulations using either the FP, Wilson or tree-level Symanzik improved lattice action shows very good consistency. The β -function results are rescaled by a factor of 50 for visibility.

servables, paving the way for various future applications. As an RG-improved lattice action, it can be used in any lattice-QCD study. One application is a new measurement of the Λ -parameter of SU(3) pure gauge theory, as a consistency check of recent results [51, 52, 56] which differ from the older literature average [1], and as quantitative input for the determination of the strong coupling $\alpha_S(M_Z)$ through the decoupling of quarks [57]. Another possibility is using ML to parameterize the FP Dirac operator [28, 58, 59] and reduce cutoff effects in lattice-QCD simulations with dynamical quarks. A third is using ML to construct quantum perfect actions [15], where all lattice artifacts are removed. As a general lesson, our study shows a new way in which ML can enhance particle physics simulations.

ACKNOWLEDGEMENTS

K.H. wishes to thank the AEC and ITP at the University of Bern for their support. U.W. acknowledges funding from the Swiss National Science Foundation (SNSF) project No. 200020_208222. A.I. and D.M. acknowledge funding from the Austrian Science Fund (FWF) projects P 32446, P 34455, P 34764 and PAT3667424. K.H. acknowledges funding from the US National Science Foundation under grants 2014150 and 2412320. The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC) and LEONARDO at CINECA, Italy, via an AURELEO (Austrian Users at LEONARDO supercomputer) project, and UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

- * kholland@pacific.edu
- [†] ipp@hep.itp.tuwien.ac.at
- [‡] dmueller@hep.itp.tuwien.ac.at
- [§] wenger@itp.unibe.ch
- Y. Aoki *et al.* (Flavour Lattice Averaging Group (FLAG)), FLAG Review 2021, Eur. Phys. J. C 82, 869 (2022), arXiv:2111.09849 [hep-lat].
- [2] S. Schaefer, R. Sommer, and F. Virotta (ALPHA), Critical slowing down and error analysis in lattice QCD simulations, Nucl. Phys. B 845, 93 (2011), arXiv:1009.5228 [hep-lat].
- [3] K. Symanzik, Continuum Limit and Improved Action in Lattice Theories. 1. Principles and φ^4 Theory, Nucl. Phys. B **226**, 187 (1983).
- [4] K. Symanzik, Continuum Limit and Improved Action in Lattice Theories. 2. O(N) Nonlinear Sigma Model in Perturbation Theory, Nucl. Phys. B 226, 205 (1983).
- [5] M. Lüscher and P. Weisz, On-shell improved lattice gauge theories, Commun. Math. Phys. 98, 433 (1985), [Erratum: Commun.Math.Phys. 98, 433 (1985)].
- [6] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière, D. J. Rezende, and P. E. Shanahan, Equivariant flow-based sampling for lattice gauge theory, Phys. Rev. Lett. **125**, 121601 (2020), arXiv:2003.06413 [hep-lat].
- [7] R. Abbott *et al.*, Gauge-equivariant flow models for sampling in lattice field theories with pseudofermions, Phys. Rev. D **106**, 074506 (2022), arXiv:2207.08945 [hep-lat].
- [8] R. Abbott, A. Botev, D. Boyda, D. C. Hackett, G. Kanwar, S. Racanière, D. J. Rezende, F. Romero-López, P. E. Shanahan, and J. M. Urban, Applications of flow models to the generation of correlated lattice QCD ensembles, Phys. Rev. D 109, 094514 (2024), arXiv:2401.10874 [heplat].
- [9] L. Del Debbio, J. M. Rossney, and M. Wilson, Efficient modeling of trivializing maps for lattice \u03c64 theory using normalizing flows: A first look at scalability, Phys. Rev. D 104, 094507 (2021), arXiv:2105.12481 [hep-lat].
- [10] D. Albandea, L. Del Debbio, P. Hernández, R. Kenway, J. Marsh Rossney, and A. Ramos, Learning trivializing flows, Eur. Phys. J. C 83, 676 (2023), arXiv:2302.08408 [hep-lat].
- [11] S. Bacchio, P. Kessel, S. Schaefer, and L. Vaitl, Learning trivializing gradient flows for lattice gauge theories, Phys. Rev. D 107, L051504 (2023), arXiv:2212.08469 [hep-lat].
- [12] M. Gerdes, P. de Haan, C. Rainone, R. Bondesan, and M. C. N. Cheng, Learning lattice quantum field theories with equivariant continuous flows, SciPost Phys. 15, 238 (2023), arXiv:2207.00283 [hep-lat].
- [13] M. Gerdes, P. de Haan, R. Bondesan, and M. C. N. Cheng, Continuous normalizing flows for lattice gauge theories, (2024), arXiv:2410.13161 [hep-lat].
- [14] K. A. Nicoli, C. J. Anders, T. Hartung, K. Jansen, P. Kessel, and S. Nakajima, Detecting and mitigating mode-collapse for flow-based sampling of lattice field theories, Phys. Rev. D 108, 114501 (2023), arXiv:2302.14082 [hep-lat].
- [15] P. Hasenfratz and F. Niedermayer, Perfect lattice action for asymptotically free theories, Nucl. Phys. B 414, 785 (1994), arXiv:hep-lat/9308004.
- [16] K. G. Wilson and J. B. Kogut, The Renormalization

group and the epsilon expansion, Phys. Rept. **12**, 75 (1974).

- [17] K. G. Wilson, The Renormalization Group: Critical Phenomena and the Kondo Problem, Rev. Mod. Phys. 47, 773 (1975).
- [18] K. G. Wilson, The renormalization group and critical phenomena, Rev. Mod. Phys. 55, 583 (1983).
- [19] M. Favoni, A. Ipp, D. I. Müller, and D. Schuh, Lattice Gauge Equivariant Convolutional Neural Networks, Phys. Rev. Lett. **128**, 032003 (2022), arXiv:2012.12901 [hep-lat].
- [20] U. Wenger, K. Holland, A. Ipp, and D. I. Müller, HMC and gradient flow with machine-learned classically perfect fixed point actions, PoS LATTICE2024, 466 (2025), arXiv:2502.03315 [hep-lat].
- [21] 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, 074502 (2024), arXiv:2401.06481 [hep-lat].
- [22] K. Holland, A. Ipp, D. I. Müller, and U. Wenger, Fixed point actions from convolutional neural networks, PoS LATTICE2023, 038 (2024), arXiv:2311.17816 [hep-lat].
- [23] 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, 587 (1995), arXiv:hep-lat/9506030.
- [24] M. Blatter and F. Niedermayer, New fixed point action for SU(3) lattice gauge theory, Nucl. Phys. B 482, 286 (1996), arXiv:hep-lat/9605017.
- [25] F. Niedermayer, P. Rufenacht, and U. Wenger, Fixed point gauge actions with fat links: Scaling and glueballs, Nucl. Phys. B 597, 413 (2001), arXiv:hep-lat/0007007.
- [26] M. Blatter, R. Burkhalter, P. Hasenfratz, and F. Niedermayer, Instantons and the fixed point topological charge in the two-dimensional O(3) sigma model, Phys. Rev. D 53, 923 (1996), arXiv:hep-lat/9508028.
- [27] P. Hasenfratz, V. Laliena, and F. Niedermayer, The Index theorem in QCD with a finite cutoff, Phys. Lett. B 427, 125 (1998), arXiv:hep-lat/9801021.
- [28] P. Hasenfratz, Lattice QCD without tuning, mixing and current renormalization, Nucl. Phys. B 525, 401 (1998), arXiv:hep-lat/9802007.
- [29] F. Niedermayer, Exact chiral symmetry, topological charge and related topics, Nucl. Phys. B Proc. Suppl. 73, 105 (1999), arXiv:hep-lat/9810026.
- [30] M. Lüscher, Properties and uses of the Wilson flow in lattice QCD, JHEP 08, 071, [Erratum: JHEP 03, 092 (2014)], arXiv:1006.4518 [hep-lat].
- [31] M. Lüscher and P. Weisz, Perturbative analysis of the gradient flow in non-abelian gauge theories, JHEP 02, 051, arXiv:1101.0963 [hep-th].
- [32] R. Narayanan and H. Neuberger, Infinite N phase transitions in continuum Wilson loop operators, JHEP 03, 064, arXiv:hep-th/0601210.
- [33] R. Lohmayer and H. Neuberger, Continuous smearing of Wilson Loops, PoS LATTICE2011, 249 (2011), arXiv:1110.3522 [hep-lat].
- [34] 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, 018, arXiv:1406.0827 [hep-lat].
- [35] A. Ramos and S. Sint, Symanzik improvement of the gradient flow in lattice gauge theories, Eur. Phys. J. C 76, 15 (2016), arXiv:1508.05552 [hep-lat].

- [36] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Hybrid Monte Carlo, Phys. Lett. B 195, 216 (1987).
- [37] 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, 2531 (1987).
- [38] 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, 272 (2003).
- [39] T. Takaishi and P. de Forcrand, Testing and tuning new symplectic integrators for hybrid Monte Carlo algorithm in lattice QCD, Phys. Rev. E 73, 036706 (2006), arXiv:hep-lat/0505020.
- [40] M. Lüscher, //luscher.web.cern.ch/luscher/openQCD/.
- [41] S. Borsányi, S. Dürr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, and C. McNeile (BMW), High-precision scale setting in lattice QCD, JHEP 09, 010, arXiv:1203.4469 [hep-lat].
- [42] 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, 08027 (2018), arXiv:1711.04833 [hep-lat].
- [43] A. Hasenfratz and O. Witzel, Continuous renormalization group β function from lattice simulations, Phys. Rev. D 101, 034514 (2020), arXiv:1910.06408 [hep-lat].
- [44] R. V. Harlander and T. Neumann, The perturbative QCD gradient flow to three loops, JHEP 06, 161, arXiv:1606.03756 [hep-ph].
- [45] H. Akaike, A new look at the statistical model identification, IEEE Trans. Automatic Control 19, 716 (1974).
- [46] L. Giusti and M. Lüscher, Topological susceptibility at $T > T_c$ from master-field simulations of the SU(3) gauge theory, Eur. Phys. J. C **79**, 207 (2019), arXiv:1812.02062 [hep-lat].
- [47] M. Asakawa, T. Hatsuda, T. Iritani, E. Itou, M. Kitazawa, and H. Suzuki, Determination of Reference Scales for Wilson Gauge Action from Yang–Mills Gradient Flow, (2015), arXiv:1503.06516 [hep-lat].
- [48] R. Sommer, Scale setting in lattice QCD, PoS LAT-TICE2013, 015 (2014), arXiv:1401.3270 [hep-lat].
- [49] F. Knechtli, T. Korzec, B. Leder, and G. Moir (ALPHA), Power corrections from decoupling of the charm quark, Phys. Lett. B 774, 649 (2017), arXiv:1706.04982 [heplat].
- [50] A. Ramos Martinez and G. Catumba, Testing universality of gauge theories, PoS LATTICE2022, 383 (2023).
- [51] C. H. Wong, S. Borsanyi, Z. Fodor, K. Holland, and J. Kuti, Toward a novel determination of the strong QCD coupling at the Z-pole, PoS LATTICE2022, 043 (2023), arXiv:2301.06611 [hep-lat].
- [52] A. Hasenfratz, C. T. Peterson, J. van Sickle, and O. Witzel, Λ parameter of the SU(3) Yang-Mills theory from the continuous β function, Phys. Rev. D **108**, 014502 (2023), arXiv:2303.00704 [hep-lat].
- [53] N. Brambilla, V. Leino, J. Mayer-Steudte, and A. Vairo, Static force from generalized Wilson loops on the lattice using the gradient flow, Phys. Rev. D 109, 114517 (2024), arXiv:2312.17231 [hep-lat].
- [54] S. Spiegel and C. Lehner, A high-precision continuum limit study of the HVP short-distance window, (2024), arXiv:2410.17053 [hep-lat].

- [55] T. A. DeGrand, A. Hasenfratz, P. Hasenfratz, and F. Niedermayer, Nonperturbative tests of the fixed point action for SU(3) gauge theory, Nucl. Phys. B 454, 615 (1995), arXiv:hep-lat/9506031.
- [56] M. Dalla Brida and A. Ramos, The gradient flow coupling at high-energy and the scale of SU(3) Yang–Mills theory, Eur. Phys. J. C 79, 720 (2019), arXiv:1905.05147 [heplat].
- [57] M. Dalla Brida, R. Höllwieser, F. Knechtli, T. Korzec, A. Nada, A. Ramos, S. Sint, and R. Sommer (AL-PHA), Determination of $\alpha_s(m_Z)$ by the non-perturbative decoupling method, Eur. Phys. J. C 82, 1092 (2022), arXiv:2209.14204 [hep-lat].
- [58] P. Hasenfratz, S. Hauswirth, K. Holland, T. Jörg, F. Niedermayer, and U. Wenger, The construction of generalized Dirac operators on the lattice, Int. J. Mod. Phys. C 12, 691 (2001), arXiv:hep-lat/0003013.
- [59] P. Hasenfratz, S. Hauswirth, T. Jorg, F. Niedermayer, and K. Holland, Testing the fixed point QCD action and the construction of chiral currents, Nucl. Phys. B 643, 280 (2002), arXiv:hep-lat/0205010.

SUPPLEMENTAL MATERIAL

Simulation details and gradient-flow measurements

Here we provide information on the ensembles generated using MC simulations of the FP action. The HMC integrator evolves up to trajectory length $\tau = 3.0$ using $N_{\tau} = 32$ steps for all ensembles. We use the 4MN4FP variant of Omelyan integration scheme as described in [39], generating N_{traj} trajectories.

In Table I we provide for each ensemble the linear lattice size $N_{\rm latt}$, the FP action β -value, and the number of thermalized trajectories $N_{\rm traj}$. We use symmetric lattice volumes $N_{\rm latt}^4$ throughout. In addition, we list the measured gradient-flow observables $t_{0.3}$, $t_{0.4}$, $t_{0.5}$ and $w_{0.3}$ in lattice units. For the gradient flow we use the third-order Runge-Kutta scheme [30] with timestep $\Delta t_f = 0.01$ on all ensembles. We also list in Table I the lattice volume L and lattice spacing a in physical units, computed using $\sqrt{t_{0.3}} = 0.1679 \,{\rm fm}$ [46].

In Table II we compile the FP measurements of flowtime ratios $t_c/t_{0.3}$ and t_c/w_c^2 for c = 0.3, 0.4, 0.5 and the β -function at renormalized coupling $g_{\rm GF}^2 = 15.79$ which corresponds to a flow time of $t_{0.3}$.

Autocorrelation times τ_c of the action density for c = 0.3, 0.4, 0.5 are estimated through single exponential fits of the autocorrelation function of $t_c^2 E(t_c)$, combining N_{chains} independent runs of unequal length, the shortest of which has length N_{s}^* , cf. Table III. Gradient-flow measurements are made every $N_{\text{skip}} = 32$ trajectories, except for the ensemble marked \dagger with more frequent measurements at $N_{\text{skip}} = 4$. The measurements in ensembles at $\beta = 2.80, 2.75, 2.65$ show no significant autocorrelations, which is why the corresponding entries are missing from the Table. Finally, the autocorrelation times, as

tabulated in Table III, are reexpressed in units of HMC trajectories.

The autocorrelation times are in general short, growing somewhat towards finer lattice spacing, but are significantly less than 32 trajectories. In the statistical analysis, gradient-flow measurements are further blocked with $N_{\rm block} = 8$ to suppress any remnant autocorrelation effects in the data, and we use $N_{\rm resample} = 2000$ in the bootstrap determination of statistical errors.

				P	8	•••••••••••••••••	[
$N_{\rm latt}$	β	$N_{\rm traj}$	$t_{0.3}/a^2$	$w_{0.3}/a$	$t_{0.4}/a^2$	$t_{0.5}/a^2$	$L/{ m fm}$	$a/{ m fm}$
10	2.65	51232	1.2321(27)	1.1455(15)	1.6526(37)	2.0342(49)	1.5126(16)	0.15126(16)
10	2.75	76128	1.5709(35)	1.2848(19)	2.1000(50)	2.5843(68)	1.3396(15)	0.13396(15)
16	2.75	68256	1.5702(13)	1.28442(67)	2.0989(18)	2.5822(24)	2.1438(9)	0.13399(6)
14	2.80	79392	1.7655(19)	1.3606(10)	2.3594(30)	2.9034(39)	1.7690(10)	0.12636(7)
14	2.85	74400	1.9798(26)	1.4394(13)	2.6448(39)	3.2553(51)	1.6706(11)	0.11933(8)
14	2.90	82368	2.2028(31)	1.5171(14)	2.9417(46)	3.6211(62)	1.5837(11)	0.11312(8)
16	2.90	81600	2.2085(25)	1.5192(12)	2.9493(37)	3.6299(48)	1.8077(10)	0.11298(6)
14	2.95	81312	2.4687(44)	1.6067(19)	3.2980(65)	4.0604(87)	1.4960(14)	0.10686(10)
16	2.95	70336	2.4667(33)	1.6062(15)	3.2954(51)	4.0579(65)	1.7104(11)	0.10690(7)
16	3.00	69696	2.7433(43)	1.6923(19)	3.6631(67)	4.5082(87)	1.6219(13)	0.10137(8)
16	3.05	69600	3.0379(48)	1.7809(19)	4.0571(76)	4.995(10)	1.5413(13)	0.09633(8)
18	3.05	76896	3.0435(38)	1.7828(16)	4.0651(59)	5.0050(81)	1.7323(11)	0.09624(6)
18	3.10	79808	3.3840(52)	1.8799(20)	4.5197(80)	5.565(11)	1.6429(13)	0.09127(7)
18	3.15	66688	3.7456(76)	1.9777(28)	5.003(11)	6.162(16)	1.5615(16)	0.08675(9)
18	3.20	68608	4.1508(83)	2.0811(30)	5.542(13)	6.823(18)	1.4834(14)	0.08241(8)

TABLE I. Overview of parameters for the FP ensembles used in this work together with a selection of gradient-flow scales in lattice units. The linear lattice extent L and the lattice spacing a are obtained using $\sqrt{t_{0.3}} = 0.1679 \text{ fm } [46]$.

TABLE II. FP results for gradient-flow observables. The β -function in the last column is given at the renormalized coupling $g_{\text{GF}}^2 = 15.79$ which corresponds to a flow time of $t_{0.3}$.

$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	-							
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$N_{\rm latt}$	β	$t_{0.3}/w_{0.3}^2$	$t_{0.4}/w_{0.4}^2$	$t_{0.5}/w_{0.5}^2$	$t_{0.4}/t_{0.3}$	$t_{0.5}/t_{0.3}$	$\beta-{\rm fn}$
10 2.75 $0.9518(11)$ $1.0319(12)$ $1.0769(12)$ $1.33682(59)$ $1.6451(12)$ $14.740(21)$ 16 2.75 $0.95177(35)$ $1.03223(38)$ $1.07802(43)$ $1.33672(19)$ $1.64455(43)$ $14.7399(69)$ 14 2.80 $0.95372(56)$ $1.03181(61)$ $1.07686(66)$ $1.33635(30)$ $1.64453(66)$ $14.798(11)$ 14 2.85 $0.95556(66)$ $1.03181(61)$ $1.07686(66)$ $1.33592(34)$ $1.64428(75)$ $14.847(13)$ 14 2.90 $0.95706(70)$ $1.03175(75)$ $1.07602(76)$ $1.33550(38)$ $1.64385(79)$ $14.885(14)$ 16 2.90 $0.95696(54)$ $1.03216(59)$ $1.07645(60)$ $1.33542(30)$ $1.64354(64)$ $14.882(11)$ 14 2.95 $0.95640(79)$ $1.03112(84)$ $1.07544(89)$ $1.33586(41)$ $1.64354(64)$ $14.875(15)$ 16 2.95 $0.95615(65)$ $1.03066(69)$ $1.07436(74)$ $1.33601(34)$ $1.64522(77)$ $14.870(13)$ 16 3.00 $0.95784(76)$ $1.03229(83)$ $1.07644(82)$ $1.33550(43)$ $1.64425(94)$ $14.911(16)$ 18 3.05 $0.95749(67)$ $1.03108(69)$ $1.07495(73)$ $1.33550(43)$ $1.64445(88)$ $14.904(15)$ 18 3.10 $0.95747(75)$ $1.03109(80)$ $1.07496(81)$ $1.33559(40)$ $1.64451(11)$ $14.909(18)$ 18 3.15 $0.95750(92)$ $1.033(10)$ $1.0759(12)$ $1.33519(53)$ $1.6437(12)$ $14.927(19)$	10	2.65	0.93909(81)	1.02812(89)	1.07663(97)	1.34137(46)	1.6511(10)	14.401(15)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	10	2.75	0.9518(11)	1.0319(12)	1.0769(12)	1.33682(59)	1.6451(12)	14.740(21)
14 2.80 $0.95372(56)$ $1.03181(61)$ $1.07686(66)$ $1.33635(30)$ $1.64453(66)$ $14.798(11)$ 14 2.85 $0.95556(66)$ $1.03158(68)$ $1.07618(72)$ $1.33592(34)$ $1.64428(75)$ $14.847(13)$ 14 2.90 $0.95706(70)$ $1.03175(75)$ $1.07602(76)$ $1.33550(38)$ $1.64385(79)$ $14.885(14)$ 16 2.90 $0.95696(54)$ $1.03216(59)$ $1.07645(60)$ $1.33542(30)$ $1.64354(64)$ $14.882(11)$ 14 2.95 $0.95640(79)$ $1.03112(84)$ $1.07544(89)$ $1.33586(41)$ $1.64522(77)$ $14.875(15)$ 16 2.95 $0.95615(65)$ $1.03066(69)$ $1.07436(74)$ $1.33601(34)$ $1.64522(77)$ $14.870(13)$ 16 3.00 $0.95784(76)$ $1.03229(83)$ $1.07644(82)$ $1.33520(41)$ $1.64330(88)$ $14.907(15)$ 16 3.05 $0.95780(82)$ $1.03129(86)$ $1.07502(95)$ $1.33550(43)$ $1.64425(94)$ $14.911(16)$ 18 3.10 $0.95749(67)$ $1.03109(80)$ $1.07495(73)$ $1.33550(43)$ $1.64446(75)$ $14.904(15)$ 18 3.15 $0.95750(92)$ $1.0303(10)$ $1.0739(10)$ $1.3351(49)$ $1.6437(12)$ $14.907(15)$ 18 3.20 $0.95841(99)$ $1.0316(11)$ $1.0755(12)$ $1.33519(53)$ $1.6437(12)$ $14.927(19)$	16	2.75	0.95177(35)	1.03223(38)	1.07802(43)	1.33672(19)	1.64455(43)	14.7399(69)
14 2.85 $0.95556(66)$ $1.03158(68)$ $1.07618(72)$ $1.33592(34)$ $1.64428(75)$ $14.847(13)$ 14 2.90 $0.95706(70)$ $1.03175(75)$ $1.07602(76)$ $1.33550(38)$ $1.64385(79)$ $14.885(14)$ 16 2.90 $0.95696(54)$ $1.03216(59)$ $1.07645(60)$ $1.33542(30)$ $1.64354(64)$ $14.882(11)$ 14 2.95 $0.95640(79)$ $1.03112(84)$ $1.07544(89)$ $1.33586(41)$ $1.64458(91)$ $14.875(15)$ 16 2.95 $0.95615(65)$ $1.03066(69)$ $1.07436(74)$ $1.33601(34)$ $1.64522(77)$ $14.870(13)$ 16 3.00 $0.95784(76)$ $1.03229(83)$ $1.07644(82)$ $1.33520(41)$ $1.64425(94)$ $14.907(15)$ 16 3.05 $0.95780(82)$ $1.03129(86)$ $1.07502(95)$ $1.33550(43)$ $1.64425(94)$ $14.907(15)$ 18 3.10 $0.95747(75)$ $1.03108(69)$ $1.07495(73)$ $1.33559(40)$ $1.64445(88)$ $14.904(15)$ 18 3.15 $0.95750(92)$ $1.0303(10)$ $1.0739(10)$ $1.33581(49)$ $1.6437(12)$ $14.927(19)$ 18 3.20 $0.95841(99)$ $1.0316(11)$ $1.0755(12)$ $1.33519(53)$ $1.6437(12)$ $14.927(19)$	14	2.80	0.95372(56)	1.03181(61)	1.07686(66)	1.33635(30)	1.64453(66)	14.798(11)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	14	2.85	0.95556(66)	1.03158(68)	1.07618(72)	1.33592(34)	1.64428(75)	14.847(13)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	14	2.90	0.95706(70)	1.03175(75)	1.07602(76)	1.33550(38)	1.64385(79)	14.885(14)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	16	2.90	0.95696(54)	1.03216(59)	1.07645(60)	1.33542(30)	1.64354(64)	14.882(11)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	14	2.95	0.95640(79)	1.03112(84)	1.07544(89)	1.33586(41)	1.64458(91)	14.875(15)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	16	2.95	0.95615(65)	1.03066(69)	1.07436(74)	1.33601(34)	1.64522(77)	14.870(13)
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	16	3.00	0.95784(76)	1.03229(83)	1.07644(82)	1.33520(41)	1.64330(88)	14.907(15)
183.050.95749(67)1.03108(69)1.07495(73)1.33562(34)1.64446(75)14.904(13)183.100.95747(75)1.03109(80)1.07496(81)1.33559(40)1.64445(88)14.904(15)183.150.95750(92)1.0303(10)1.0739(10)1.33581(49)1.6451(11)14.909(18)183.200.95841(99)1.0316(11)1.0755(12)1.33519(53)1.6437(12)14.927(19)	16	3.05	0.95780(82)	1.03129(86)	1.07502(95)	1.33550(43)	1.64425(94)	14.911(16)
183.100.95747(75)1.03109(80)1.07496(81)1.33559(40)1.64445(88)14.904(15)183.150.95750(92)1.0303(10)1.0739(10)1.33581(49)1.6451(11)14.909(18)183.200.95841(99)1.0316(11)1.0755(12)1.33519(53)1.6437(12)14.927(19)	18	3.05	0.95749(67)	1.03108(69)	1.07495(73)	1.33562(34)	1.64446(75)	14.904(13)
183.150.95750(92)1.0303(10)1.0739(10)1.33581(49)1.6451(11)14.909(18)183.200.95841(99)1.0316(11)1.0755(12)1.33519(53)1.6437(12)14.927(19)	18	3.10	0.95747(75)	1.03109(80)	1.07496(81)	1.33559(40)	1.64445(88)	14.904(15)
$18 \qquad 3.20 \qquad 0.95841(99) \qquad 1.0316(11) \qquad 1.0755(12) \qquad 1.33519(53) \qquad 1.6437(12) \qquad 14.927(19)$	18	3.15	0.95750(92)	1.0303(10)	1.0739(10)	1.33581(49)	1.6451(11)	14.909(18)
	18	3.20	0.95841(99)	1.0316(11)	1.0755(12)	1.33519(53)	1.6437(12)	14.927(19)

TABLE III. Statistics and autocorrelation times τ_c of $t_c^2 E(t_c)$ in units of HMC trajectories for the FP ensembles with a < 0.12 fm. The ensembles are assembled from $N_{\rm chains}$ independent chains of unequal length, the shortest of which has length $N_{\rm s}^*$. Gradient-flow measurements are done every $N_{\rm skip} = 32$ trajectories, except for the ensemble marked † with $N_{\rm skip} = 4$. Autocorrelation times for ensembles with $\beta \leq 2.800$ are too short to be measured accurately.

β	$N_{\rm latt}$	$N_{\rm chains}$	$N_{ m s}^*$	$ au_{0.3}$	$ au_{0.4}$	$ au_{0.5}$
2.650	10	2	25536	-	-	-
2.750	10	3	25280	-	-	-
2.750	16	24	2656	-	-	-
2.800	14	8	9408	-	-	-
2.850^{+}	14	1	10844	7.1	7.6	7.8
2.850	14	8	2496	6.9	6.4	6.2
2.900	14	8	10176	9.3	9.7	9.9
2.900	16	16	3776	8.8	9.3	9.4
2.950	14	8	9888	12.8	13.3	13.4
2.950	16	16	2496	9.4	9.9	10.2
3.000	16	16	3520	14.7	15.2	15.5
3.050	16	16	3168	11.3	12.0	12.5
3.050	18	28	2688	14.2	15.1	15.6
3.100	18	29	2720	16.4	17.0	17.6
3.150	18	32	1632	19.8	20.7	21.3
3.200	18	32	2112	17.3	19.2	20.7

Continuum limits and AIC averaging procedure

We perform the continuum limit for each dimensionless quantity z by fitting the results from the simulations with polynomials of order n in $a^2/t_{0.3}$. For the FP results we use n = 0, 1, and 2, while for the results from the Wilson and tree-level Symanzik improved action we use n = 1, 2, and 3. The results obtained with the plaquette and clover discretization of the action density E are treated as independent and we do not perform combined continuum limits.

The continuum limits are illustrated in Figures 1, 2, and 4 where we show the results of the simulations together with a selection of (bootstrapped) polynomial fits indicating the included data points by the extent of the lines and shaded according to their Akaike Information Criterion (AIC) weight [45] calculated as follows. For each possible fit *i* given n_{tot} data of which *n* are fitted using *k* parameters, the weight ω_i for the continuum limit z_i is given by

$$AIC_{i} = \chi_{i}^{2} + 2k + 2(n_{tot} - n), \qquad (11)$$

$$\omega_i = \exp(-\text{AIC}_i/2). \tag{12}$$

The AIC weights are also used to estimate systematic effects in the continuum extrapolation as follows. The probability distribution function (PDF) of z is built from a weighted combination of each distribution of z_i parameterized as a normal distribution $\mathcal{N}(z_i, \sigma_i)$. The median of the resulting PDF is taken as the final result, and the 16th and 84th percentiles of the cumulative distribution function as (possibly asymmetric) uncertainty ranges. This procedure combines both statistical and systematic errors in the extrapolation. In Figures 5 and 6 we show the PDFs for each continuum limit and quantity. The multipeak structures seen in the PDFs reflect the fact that at the given high statistical accuracy the continuum extrapolations in general depend on the order of the fitted polynomial. As a consequence, the total error is dominated by the systematics of the continuum extrapolation as pointed out in the introduction. The final results from the continuum extrapolations are collected in Table IV.

In comparing the results, we calculate the weighted average PDF_{avg} from the individual PDFs using

$$PDF_{avg} = \frac{\sum_{k} \omega_k PDF_k}{\sum_{k} \omega_k},$$
(13)

where the index k runs over the five cases: FP, Wilson (plaquette), Wilson (clover), Symanzik (plaquette), and Symanzik (clover). PDF_k denotes the respective individual PDF. Since the data from the plaquette and the clover definition of the action density for either the Wilson or the Symanzik action are highly correlated, we follow a conservative approach and suppress these contributions by a factor of $\omega_k = 1/2$, while $\omega_{\rm FP} = 1$. Finally, Figure 3 shows the deviation $z_k - z_{\rm avg}$, where $z_{\rm avg}$ is the median of



FIG. 4. Continuum-limit extrapolations of FP, Wilson and Symanzik action measurements, for the ratios $t_{0.4}/t_{0.3}$ (top), $t_{0.4}/w_{0.4}^2$ (middle), and $t_{0.5}/w_{0.5}^2$ (bottom). Lattice artifacts are small with the FP action and gradient flow. The Wilson and Symanzik actions have much larger cutoff effects for both the clover and plaquette discretizations of E. A variety of polynomial fits with different ranges and highest power $(a^2/t_0)^n$ are shown, shaded according to their respective AIC weight.



FIG. 5. Comparison of various PDFs for the quantity $\beta(g^2 = 15.79 \ (top), \ t_{0.5}/t_{0.3} \ (middle), \ t_{0.4}/t_{0.3} \ (bottom)$. Below the plot, the median, $16^{\rm th}$ and $84^{\rm th}$ percentiles are marked.

 PDF_{avg} , for all five cases, together with their respective



FIG. 6. Comparison of various PDFs for the quantity $t_{0.5}/w_{0.5}^2$ (top), $t_{0.4}/w_{0.4}^2$ (middle), $t_{0.3}/w_{0.3}^2$ (bottom). Below the plot, the median, 16^{th} and 84^{th} percentiles are marked.

errors.

TABLE IV. Continuum values of gradient-flow observables using AIC model averaging with fits up to a^6 -terms (Wilson and Symanzik) and up to a^4 -terms (FP).

Observable	FP	Wilson clover	Wilson plaquette	Symanzik clover	Symanzik plaquette	Weighted average
$t_{0.3}/w_{0.3}^2$	$0.9553(^{+46}_{-32})$	$0.9567(^{+13}_{-7})$	$0.9582(^{+12}_{-13})$	$0.9594(^{+21}_{-12})$	$0.9595(^{+18}_{-22})$	$0.9579(^{+24}_{-35})$
$t_{0.4}/w_{0.4}^2$	$1.0305(^{+11}_{-9})$	$1.0277(^{+19}_{-11})$	$1.0274(^{+20}_{-10})$	$1.0293(^{+29}_{-15})$	$1.0302(^{+36}_{-17})$	$1.0297(^{+19}_{-25})$
$t_{0.5}/w_{0.5}^2$	$1.0726(^{+25}_{-11})$	$1.0715(^{+20}_{-21})$	$1.0700(^{+24}_{-11})$	$1.0718(^{+30}_{-15})$	$1.0718(^{+40}_{-17})$	$1.0719(^{+28}_{-21})$
$t_{0.4}/t_{0.3}$	$1.3356(^{+15}_{-12})$	$1.3352(^{+10}_{-12})$	$1.3360(^{+7}_{-11})$	$1.3345(^{+16}_{-8})$	$1.3358(^{+12}_{-23})$	$1.3356(^{+12}_{-15})$
$t_{0.5}/t_{0.3}$	$1.6443(^{+3}_{-3})$	$1.6456(^{+20}_{-30})$	$1.6464(^{+10}_{-20})$	$1.6440(^{+26}_{-21})$	$1.6457(^{+19}_{-46})$	$1.6445(^{+25}_{-18})$
$\beta(g^2 = 15.79)$	$14.853(^{+130}_{-71})$	$14.906(^{+26}_{-12})$	$14.928(^{+20}_{-22})$	$14.960(^{+40}_{-21})$	$14.972(^{+20}_{-26})$	$14.933(^{+48}_{-92})$