Letter

Novel approach for computing gradients of physical observables

Simone Bacchio
Computation-based Science and Technology Research Center, The Cyprus Institute, Nicosia, Cyprus

(Received 17 May 2023; accepted 15 November 2023; published 30 November 2023)

We show that an infinitesimal step of gradient flow can be used for defining a novel approach for computing gradients of physical observables with respect to action parameters. Compared to the commonly used perturbative expansion, this approach does not require calculating any disconnected contribution or vacuum expectation value and can provide results up to 3 orders of magnitudes more precise. On the other hand, it requires a nontrivial condition to be satisfied by the flow action, the calculation of its force and its Laplacian, and the force of the observable, whose gradient needs to be measured. As a proof of concept, we measure gradients in β of Wilson loops in a four-dimensional SU(3) Yang-Mills theory simulated on a 16^4 lattice using the Wilson action.

DOI: 10.1103/PhysRevD.108.L091508

Introduction. In a lattice gauge theory with an action $S_\theta \equiv S(U,\theta)$, the expectation value of an observable $\mathcal{O}(U)$ is given by the path integral

$$\langle \mathcal{O} \rangle_\theta = \frac{1}{\mathcal{Z}_\theta} \int \mathrm{D}[U] \mathcal{O}(U) \exp(-S(U,\theta)), \tag{1}$$

where $\mathcal{Z}_\theta \equiv \int \mathrm{D}[U] \exp(-S(U,\theta))$ is the partition function and θ is a parameter of the action. The focus of this work is the calculation of the gradient of the expectation value with respect to the parameter θ, which might result in a challenging task since the so-called sea effects have to be measured. Indeed, it is standard to perturbatively expand the expectation value with respect to an infinitesimal change of the parameter, $d\theta$, obtaining

$$\langle \mathcal{O}_{\theta+d\theta} \rangle_{\theta+d\theta} = \frac{\langle d\theta \frac{\partial \mathcal{O}_\theta}{\partial \theta} + \mathcal{O}_\theta \exp(-d\theta \frac{\partial S_\theta}{\partial \theta}) \rangle_\theta}{\langle \exp(-d\theta \frac{\partial S_\theta}{\partial \theta}) \rangle_\theta} + O(d\theta^2),$$

and, thus, the gradient with respect to θ is given by

$$\begin{aligned} \frac{d\langle \mathcal{O}_\theta \rangle_\theta}{d\theta} &= \lim_{d\theta \to 0} \frac{\langle \mathcal{O}_{\theta+d\theta} \rangle_{\theta+d\theta} - \langle \mathcal{O}_\theta \rangle_\theta}{d\theta} \\ &= \left\langle \frac{\partial \mathcal{O}_\theta}{\partial \theta} - \mathcal{O}_\theta \frac{\partial S_\theta}{\partial \theta} \right\rangle_\theta + \langle \mathcal{O}_\theta \rangle_\theta \left\langle \frac{\partial S_\theta}{\partial \theta} \right\rangle_\theta, \end{aligned} \tag{2}$$

where, in the above equations, for generality's sake, we have introduced an optional explicit dependence on θ also in the observable $\mathcal{O}_\theta \equiv \mathcal{O}(U,\theta)$. Since on both sides of Eq. (2) all terms depend on θ, we rewrite the equation with the shorthand notation

$$\frac{d\langle \mathcal{O} \rangle}{d\theta} = \left\langle \frac{\partial \mathcal{O}}{\partial \theta} - \mathcal{O} \frac{\partial S}{\partial \theta} \right\rangle + \langle \mathcal{O} \rangle \left\langle \frac{\partial S}{\partial \theta} \right\rangle, \tag{3}$$

and we refer to

$$\begin{aligned} \left\langle \frac{\partial \mathcal{O}}{\partial \theta} \right\rangle &: \text{ as connected contribution,} \\ \left\langle \mathcal{O} \frac{\partial S}{\partial \theta} \right\rangle &: \text{ as disconnected contribution, and} \\ \langle \mathcal{O} \rangle \left\langle \frac{\partial S}{\partial \theta} \right\rangle &: \text{ as vacuum expectation value.} \end{aligned}$$

Disconnected contributions are well known to be noisy. Furthermore, when the vacuum expectation value is nonzero, the latter and the disconnected contributions are usually large values, whose difference needs to be taken accurately for reliably measuring the gradient.

In this work, we present a novel and alternative approach, based on gradient-flow techniques [1], which is free from any of the aforementioned problems. Namely, if a flow action that flows along the parameter θ is found, then gradients of observables can be computed directly using the ordinary differential equation (ODE) of the operator. Our numerical results show that sea effects can be computed up to 3 orders of magnitude more precisely than the perturbative expansion in Eq. (3), when an exact flow action is found. Therefore, this approach can possibly be used to improve the accuracy of gradients in, for example, among others:

(i) applications of the Feynman-Hellmann theorem [2–7],
(ii) leading isospin breaking corrections [8–12],
(iii) leading QED corrections [12–14],
(iv) fine-tuning of simulation parameters [15],
(v) leading contribution of the QCD Θ term to the neutron electric dipole moment [16–18].

In the following, we present the novel approach in the form of a theorem, followed by its proof and a numerical case study.

Main result.

Notation: In the following,

(i) $\tilde{S}$ is referred to as *flow action* and is to be determined;
(ii) $\mathcal{L}_0 = -\sum_{x,\mu,a} \partial^a_{x,\mu}\partial^a_{x,\mu}$ denotes the Laplacian;
(iii) $(A,B) = \sum_{x,\mu,a} A^a_\mu(x) B^a_\mu(x)$ denotes the scalar product over algebra-valued fields;
(iv) ∂ is the force of a scalar function, defined as

$$\partial^a_{x,\mu} f(U) = \frac{d}{d\tau} f(U_\tau)\Big|_{\tau=0} \tag{4}$$

with

$$U_\tau(y,\nu) = \begin{cases} e^{\tau T^a} U(x,\mu), & \text{for } (x,\mu) = (y,\nu), \\ U(y,\nu) & \text{for } (x,\mu) \neq (y,\nu). \end{cases} \tag{5}$$

Theorem, gradient-flow approach: The gradient of an expectation value with respect to a parameter θ of the action is given by

$$\frac{d\langle\mathcal{O}\rangle}{d\theta} = \left\langle \frac{\partial\mathcal{O}}{\partial\theta} + (\partial\mathcal{O}, \partial\tilde{S}) - \mathcal{O}\mathcal{C} \right\rangle + \langle\mathcal{O}\rangle\langle\mathcal{C}\rangle, \tag{6}$$

where

$$\mathcal{C} = \mathcal{L}_0\tilde{S} + (\partial S, \partial\tilde{S}) + \frac{\partial S}{\partial\theta}. \tag{7}$$

Corollary, ideal gradient-flow approach: If the flow action $\tilde{S}$ is such that $\mathcal{C} =$ constant, then Eq. (6) simplifies to

$$\frac{d\langle\mathcal{O}\rangle}{d\theta} = \left\langle \frac{\partial\mathcal{O}}{\partial\theta} + (\partial\mathcal{O}, \partial\tilde{S}) \right\rangle. \tag{8}$$

We, therefore, refer to *ideal* flow action as an $\tilde{S}$ such that

$$\mathcal{L}_0\tilde{S} + (\partial S, \partial\tilde{S}) + \frac{\partial S}{\partial\theta} = \text{constant}. \tag{9}$$

Corollary, standard perturbative approach: If $\tilde{S} \equiv$ constant, then Eq. (6) simplifies to Eq. (3), since $\mathcal{L}_0\tilde{S} = \partial\tilde{S} = 0$ and $\mathcal{C} = \partial S_\theta/\partial\theta$.

Summary: The corollaries are two special applications of the theorem, and our numerical results show that gradients computed using the ideal gradient-flow approach are significantly more precise than those computed using the standard perturbative approach. If Eq. (9) is not satisfied exactly, then the theorem can be used to obtain correct and possibly improved results.

Proof. The proof is based on gradient-flow techniques that have been introduced in Ref. [1], which we refer to for further details. Under a change of variables $U = \mathcal{F}(V)$, the expectation value of an observable transforms as

$$\begin{aligned} \langle\mathcal{O}\rangle &= \frac{1}{\mathcal{Z}} \int \mathrm{D}[U]\mathcal{O}(U)\exp(-S(U)) \\ &= \frac{1}{\mathcal{Z}_\mathcal{F}} \int \mathrm{D}[V]\mathcal{O}(\mathcal{F}(V))\exp(-S_\mathcal{F}(V)), \end{aligned} \tag{10}$$

where $\mathcal{Z}_\mathcal{F} \equiv \int \mathrm{D}[V]\exp(-S_\mathcal{F}(V))$ and

$$S_\mathcal{F}(V) = S(\mathcal{F}(V)) - \ln\det\mathcal{F}_*(V) \tag{11}$$

with $\mathcal{F}_*$ being the Jacobian of the transformation. In the following, we consider an infinitesimal transformation of the field employing an Euler integration step and the force of an action $\tilde{S}$:

$$\begin{aligned} \mathcal{F}(V) &= \exp(-\epsilon\partial\tilde{S}(V))V \\ &= V - \epsilon\partial\tilde{S}(V)V + O(\epsilon^2) \quad \text{with} \quad 0 < \epsilon \ll 1. \end{aligned} \tag{12}$$

Two properties of such flow are [1]

$$\ln\det\mathcal{F}_*(V) = \epsilon\mathcal{L}_0\tilde{S}(V) + O(\epsilon^2) \tag{13}$$

and

$$\mathcal{O}(\mathcal{F}(V)) = \mathcal{O}(V) - \epsilon(\partial\mathcal{O}(V), \partial\tilde{S}(V)) + O(\epsilon^2). \tag{14}$$

By using Eqs. (13) and (14) in Eq. (11), we obtain

$$S_\mathcal{F} = S - \epsilon[(\partial S, \partial\tilde{S}) + \mathcal{L}_0\tilde{S}] + O(\epsilon^2), \tag{15}$$

where all terms now depend on the field V. The intent is then to interpret the difference $S_\mathcal{F} - S$ as an infinitesimal change of action along the parameter θ, requiring

$$\frac{\partial S}{\partial\theta} = \lim_{\epsilon\to 0}\frac{S_\mathcal{F} - S}{\epsilon} = -(\partial S, \partial\tilde{S}) - \mathcal{L}_0\tilde{S}. \tag{16}$$

We then define $\mathcal{C}$, in Eq. (7), as the difference between the left- and right-hand sides. By using $\mathcal{C}$ in Eq. (15), we can remove any dependence on $\tilde{S}$, obtaining

$$S_{\mathcal{F}} = S + \epsilon\left(\frac{\partial S}{\partial\theta} - \mathcal{C}\right) + O(\epsilon^2). \tag{17}$$

In the following, we track explicitly the dependence on the parameter θ, and from Eq. (17), we obtain

$$S_{\mathcal{F}} = S_{\theta+\epsilon} - \epsilon\mathcal{C}_\theta + O(\epsilon^2). \tag{18}$$

By using the latter in the definition of $\mathcal{Z}_{\mathcal{F}}$, we obtain

$$\mathcal{Z}_{\mathcal{F}} = \mathcal{Z}_{\theta+\epsilon} - \epsilon\mathcal{Z}_\theta\langle\mathcal{C}_\theta\rangle_\theta + O(\epsilon^2). \tag{19}$$

Finally, by using the above in Eq. (10), we obtain

$$\begin{aligned}\langle\mathcal{O}_\theta\rangle_\theta = \langle\mathcal{O}_\theta\rangle_{\theta+\epsilon} + \epsilon\langle\mathcal{O}_\theta\mathcal{C}_\theta\rangle_\theta - \epsilon\langle\mathcal{O}_\theta\rangle_\theta\langle\mathcal{C}_\theta\rangle_\theta \\ - \epsilon\langle(\partial\mathcal{O}_\theta, \partial\tilde{S}_\theta)\rangle_\theta + O(\epsilon^2),\end{aligned} \tag{20}$$

and the theorem is proved since

$$\begin{aligned}\frac{d\langle\mathcal{O}_\theta\rangle_\theta}{d\theta} &= \lim_{\epsilon\to 0}\frac{\langle\mathcal{O}_{\theta+\epsilon}\rangle_{\theta+\epsilon} - \langle\mathcal{O}_\theta\rangle_\theta}{\epsilon} \\ &= \left\langle\frac{\partial\mathcal{O}_\theta}{\partial\theta} + (\partial\mathcal{O}_\theta, \partial\tilde{S}_\theta) - \mathcal{O}_\theta\mathcal{C}_\theta\right\rangle_\theta + \langle\mathcal{O}_\theta\rangle_\theta\langle\mathcal{C}_\theta\rangle_\theta.\end{aligned} \tag{21}$$

Case study and numerical results.

Wilson action: To demonstrate the effectiveness of Eq. (6) compared to Eq. (3), we consider a four-dimensional SU(3) Yang-Mills theory defined on a lattice using the standard Wilson action

$$S_W(\beta, U) = -\frac{\beta}{6}\mathcal{W}_0(U), \tag{22}$$

where $\mathcal{W}_0$ denotes the sum of plaquettes. We aim at computing the slope in β of various observables, and therefore the approach outlined here requires a flow action $\tilde{S}$ that satisfies Eq. (9) in the following way:

$$\mathcal{L}_0\tilde{S} - \frac{\beta}{6}(\partial\mathcal{W}_0, \partial\tilde{S}) - \frac{1}{6}\mathcal{W}_0 = \text{constant}. \tag{23}$$

Inspired by Lüscher's t expansion [1], it is easy to show that an analytical solution to Eq. (23) is

$$\begin{aligned}\tilde{S} &= \frac{1}{6}\sum_{k=0}^{\infty}\left(\frac{\beta}{6}\right)^k\tilde{S}^{(k)} \quad \text{with} \\ \tilde{S}^{(0)} &= \mathcal{L}_0^{-1}\mathcal{W}_0 = \frac{3}{16}\mathcal{W}_0 \quad \text{and} \\ \tilde{S}^{(k)} &= \mathcal{L}_0^{-1}(\partial\mathcal{W}_0, \partial\tilde{S}^{(k-1)}) \quad \text{for } k > 0,\end{aligned} \tag{24}$$

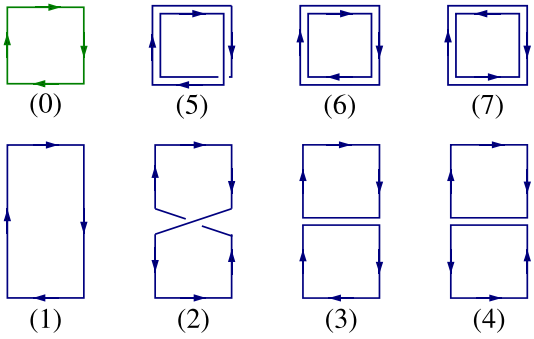

FIG. 1. Representation of the loops entering the β expansion up to the NLO. The plaquette, shown by the green loop, enters the LO expansion, while the loops shown in blue enter the NLO. The enumeration of the loops follows Ref. [1].

which we refer to as β expansion, and it is equivalent to Lüscher's t expansion evaluated at $t = \beta$ divided by β. We refer to $\tilde{S}^{(0)}$ as the leading order (LO) and to $\tilde{S}^{(1)}$ as the next-to-leading order (NLO). The calculation of the NLO can be found in Ref. [1], where the term $(\partial\mathcal{W}_0, \partial\mathcal{W}_0)$ is computed. It results in a linear combination of the Wilson loops depicted in Fig. 1, which are defined as

$$\begin{aligned}\mathcal{W}_i &= \sum_{C\in\Gamma_i}\text{tr}\{U(C)\} \quad \text{for } i = 0, 1, 2, 5 \\ \mathcal{W}_i &= \sum_{C,C'\in\Gamma_i}\text{tr}\{U(C)\}\text{tr}\{U(C')\} \quad \text{for } i = 3, 4, 6, 7,\end{aligned} \tag{25}$$

where Γ_i are all unique loops for a given shape, including loops in the perpendicular direction—i.e., chair shaped—for $i = 1,2,3,4$.

Gradients of Wilson loops: Since the convergence of the β expansion in Eq. (24) deteriorates as β grows, we show results for simulations performed at various values of β from $\beta = 0$ up to $\beta = 6$. At $\beta = 0$, the LO of the β expansion solves Eq. (23) exactly, while, on the other hand, the region of physical interest starts from $\beta \gtrsim 5.8$, where a lattice spacing of about 0.14 fm is measured [19]. We compute the slope in β for a set of Wilson loops, $\mathcal{W}_i$, using either the perturbative expansion (PE) or the new approach using the LO or the NLO of Eq. (24) computed at each β. For all experiments, we use a lattice of size 16^4 and measure expectation values over 10,000 configurations simulated at each value of β using an Hamiltonian Monte Carlo (HMC) algorithm. To compare gradients for various Wilson loops, we study the deviation of the value measured using an approach "T" from the weighted average of all approaches, namely,

$$\Delta_{T,i} = \frac{1}{\sigma_{\text{PE,i}}}\left(\left.\frac{d\langle W_i\rangle}{d\beta}\right|_T - \left.\frac{d\langle W_i\rangle}{d\beta}\right|_{\text{PE}}\right), \tag{26}$$

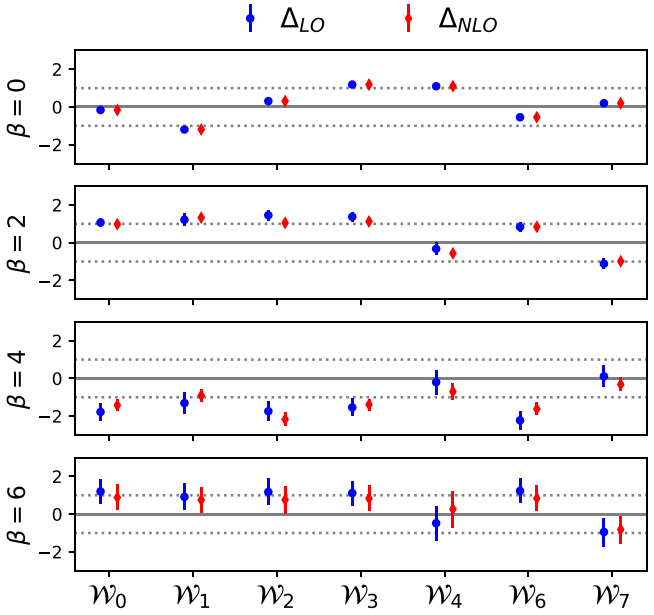

FIG. 2. Accuracy on the gradients in β for the Wilson loops, $\mathcal{W}_i$ depicted in Fig. 1, measured using the perturbative expansion in Eq. (3) or alternatively using the new approach in Eq. (6) with the LO or the NLO terms for the flow action $\tilde{S}$ in Eq. (24). We depict the deviation Δ_T from the values measured using PE normalized such that the errors of PE are unitary; see Eq. (26). Thus, the continuous lines are the central values of PE, while the dashed lines are plus or minus one unit of the error of PE. Results at various β are computed over 10,000 configurations with lattice size 16^4, simulated with HMC and separated by 4 MDUs.

where $\frac{d\langle W_i\rangle}{d\beta}|_T$ is the value from the approach T and we normalize the results using the error from the perturbative expansion PE, such that the error of $\Delta_{\mathrm{PE},i}$ is always unity. Results are depicted in Fig. 2 and summarized in Table I. As can be seen, we observe excellent statistical agreement between the results measured using the perturbative expansion and the novel approach discussed in this work, supporting the correctness of the derivation. We also observe an impressive improvement in the error achieved at small β. This is highlighted in Table I, in which we list the minimum and maximum values of the ratio between the errors obtained using the various approaches. At $\beta = 0$, results for LO and NLO are the same, since LO solves exactly Eq. (23). Here, we achieve the largest improvement for the signal-to-noise ratio, which is of more than 3 orders of magnitude in calculating the slope in β of the plaquette. Additionally, looking at Table I, we observe a very strong connection between the improvement in the errors and the ratio of the standard deviations of C and $\partial S/\partial\beta$. The latter, therefore, can be used to have an estimate of the gain one would achieve in a generic situation.

TABLE I. Smallest and largest value for the ratio of the error obtained using the new approach defined by Eq. (6) and the error of the perturbative expansion given by Eq. (3) over the measured Wilson loops. Results are given for the cases when the LO or the NLO is used in the flow action. We report also the ratio between the standard deviation of C and the one of $S' = \partial S/\partial\beta$, noting a strong connection between this ratio and the one between the errors.

		$\beta=0$	$\beta=1$	$\beta=2$	$\beta=3$	$\beta=4$	$\beta=5$	$\beta=6$
LO	min	0.08%	12.0%	23.1%	37.6%	45.4%	67.6%	65.3%
	max	0.45%	16.7%	33.5%	48.2%	66.2%	81.1%	92.5%
	$\sigma_C/\sigma_{S'}$	0.00%	15.0%	29.2%	47.1%	65.6%	75.6%	102.0%
NLO	min	0.08%	1.8%	6.9%	17.0%	29.6%	55.3%	68.9%
	max	0.45%	2.5%	10.4%	22.4%	43.9%	66.3%	99.1%
	$\sigma_C/\sigma_{S'}$	0.00%	2.4%	9.7%	22.8%	42.3%	62.2%	100.7%

Convergence of the β expansion: On a side note, our experiments let us comment on the convergence of the β expansion in Eq. (24). This provides some interesting insights into the ability to flow in β using gradient flows [20]. On one side, if all $\tilde{S}^{(k)}$ have similar magnitude, the series would be converging only for $\beta < 6$. On the other hand, $\tilde{S}^{(k)}$ might decrease in magnitude at increasing k, therefore converging for all β. In Ref. [1], an analytical study on the convergence of the t expansion is found, and similar arguments can be applied to the β expansion since they share the same origin. In the following, on the other hand, we provide numerical evidence for the first few orders. In our numerical experiments, we have computed the actions $\tilde{S}^{(0)}$ and $\tilde{S}^{(1)}$. Additionally, we can measure the value of $\mathcal{L}_0\tilde{S}^{(2)}$ from its definition, namely,

$$\mathcal{L}_0\tilde{S}^{(2)} = (\partial\mathcal{W}_0, \partial\tilde{S}^{(1)}), \tag{27}$$

which is the scalar product of the forces of $\mathcal{W}_0$ and of $\tilde{S}^{(1)}$, both computed in our experiments. Results for the ratio of the first three orders are given in Fig. 3, in which we analyze the standard deviation of these terms and not their value since the flow action is defined up to a constant (23).

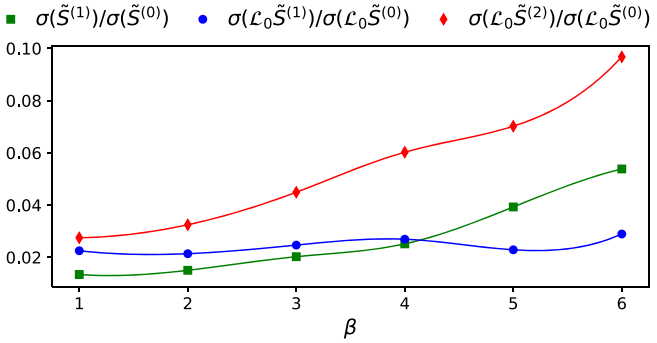

FIG. 3. Study of the convergence of the β-expansion, showing ratios between the standard deviation of the LO, the NLO, and the next-to-next-leading-order (NNLO) of the flow action or its Laplacian. Lines are interpolations for guiding the eye.

While $\sigma(\tilde{S}^{(1)})$ is almost 2 orders of magnitude smaller than $\sigma(\tilde{S}^{(0)})$, the ratio $\sigma(\mathcal{L}_0\tilde{S}^{(2)})/\sigma(\mathcal{L}_0\tilde{S}^{(0)})$ is larger than $\sigma(\mathcal{L}_0\tilde{S}^{(1)})/\sigma(\mathcal{L}_0\tilde{S}^{(0)})$ and increases in β, pointing in the direction that the various orders might have similar magnitude. If this is true, then the β expansion is converging only for $\beta < 6$, while at larger β, possibly all terms contribute to the flow action, making it an extensive operator in the region where the regime of physical interest begins. This is unfortunate for flow-based trivializing maps since flowing at large β becomes very challenging. Indeed, to date, only flowing at very small β has been proven successful in four dimensions [21].

Conclusions. In this work, we have presented a novel approach for computing gradients of observables with respect to action parameters, which, in its simpler form given in Eq. (8), is free of any disconnected contribution or vacuum expectation value. We provide numerical evidence within our case study that when a flow action solution to Eq. (9) is found the new approach provides a significant improvement in the signal-to-noise ratio compared to the commonly used perturbative expansion. We use, as a case study, the calculation of the slope in β of Wilson loops in pure-gauge SU(3) Yang-Mills theories. In this situation, we are able to provide in Eq. (24) an analytic solution to the required flow action in terms of a series expansion, referred to as β expansion. We compute the first two orders of the series and show results at various β for many Wilson loops. At $\beta = 0$, the expansion converges at the leading order, and in this case, we obtain the largest improvement in the errors. At finite β, instead, the expansion converges in orders of $\beta/6$. Therefore, a reduction in the errors is seen only for $\beta < 6$, which is a region of nonphysical interest. On the other hand, these results are only meant as proof of concept, since the approach will change according to the application. Indeed, each case should be studied independently, searching for an appropriate flow action, that satisfies Eq. (9) for a desired action S and parameter of the action θ. Where possible, this can be done analytically, or alternatively, it can be done numerically using a parametric definition of $\tilde{S}$ [20] and machine-learning techniques to find the optimal flow action. In this case, for the training, we suggest minimizing either the variance of $\langle\mathcal{C}\rangle$ [see Eq. (7)] since the final objective is to make it constant, i.e., Eq. (9), or to perform an observable-dependent training by minimizing the value or the variance of $\langle\mathcal{O}\rangle\langle\mathcal{C}\rangle - \langle\mathcal{O}\mathcal{C}\rangle$. We have tested such an approach for the case study presented here, and we have obtained results compatible with the analytical solution, and no further improvement was observed. On the other hand, for all other cases where an analytical solution is not easy to find, this is a new valuable application of machine-learning techniques.

Acknowledgments. The author wishes to thank Pan Kessel, Stefan Schaefer, and Lorenz Vaitl for useful discussions and the enjoyable collaboration on gradient flows. The author acknowledges financial support from the Cyprus Research and Innovation Foundation under the project QC4LGT (Grant No. EXCELLENCE/0421/0019) and 3D-NUCLEON (Grant No. EXCELLENCE/0421/0043). Numerical results were obtained using the Cyclamen cluster of The Cyprus Institute equipped with P100 GPUs. The software is implemented using the Lyncs-API [22], its interface to QUDA [23], and the QUDA library [24].

[1] M. Lüscher, Trivializing maps, the Wilson flow and the HMC algorithm, Commun. Math. Phys. **293**, 899 (2010).
[2] R. Horsley, R. Millo, Y. Nakamura, H. Perlt, D. Pleiter, P. Rakow, G. Schierholz, A. Schiller, F. Winter, and J. Zanotti, A lattice study of the glue in the nucleon, Phys. Lett. B **714**, 312 (2012).
[3] A. J. Chambers *et al.* (CSSM, QCDSF/UKQCD Collaborations), Feynman-Hellmann approach to the spin structure of hadrons, Phys. Rev. D **90**, 014510 (2014).
[4] A. J. Chambers *et al.*, Disconnected contributions to the spin of the nucleon, Phys. Rev. D **92**, 114517 (2015).
[5] C. Bouchard, C. C. Chang, T. Kurth, K. Orginos, and A. Walker-Loud, On the Feynman-Hellmann theorem in quantum field theory and the calculation of matrix elements, Phys. Rev. D **96**, 014504 (2017).
[6] C. C. Chang *et al.*, A per-cent-level determination of the nucleon axial coupling from quantum chromodynamics, Nature (London) **558**, 91 (2018).
[7] K. U. Can *et al.*, Lattice QCD evaluation of the Compton amplitude employing the Feynman-Hellmann theorem, Phys. Rev. D **102**, 114505 (2020).
[8] G. M. de Divitiis *et al.*, Isospin breaking effects due to the up-down mass difference in lattice QCD, J. High Energy Phys. 04 (2012) 124.
[9] G. M. de Divitiis, R. Frezzotti, V. Lubicz, G. Martinelli, R. Petronzio, G. C. Rossi, F. Sanfilippo, S. Simula, and N. Tantalo (RM123 Collaboration), Leading isospin breaking effects on the lattice, Phys. Rev. D **87**, 114505 (2013).
[10] S. Borsanyi *et al.* (BMW Collaboration), *Ab initio* calculation of the neutron-proton mass difference, Science **347**, 1452 (2015).

[11] D. Giusti, V. Lubicz, C. Tarantino, G. Martinelli, F. Sanfilippo, S. Simula, and N. Tantalo, Leading isospin-breaking corrections to pion, kaon and charmed-meson masses with twisted-mass fermions, Phys. Rev. D **95**, 114504 (2017).
[12] S. Borsanyi *et al.*, Leading hadronic contribution to the muon magnetic moment from lattice QCD, Nature (London) **593**, 51 (2021).
[13] D. Giusti, V. Lubicz, G. Martinelli, C. T. Sachrajda, F. Sanfilippo, S. Simula, N. Tantalo, and C. Tarantino, First lattice calculation of the QED corrections to leptonic decay rates, Phys. Rev. Lett. **120**, 072001 (2018).
[14] M. Di Carlo, D. Giusti, V. Lubicz, G. Martinelli, C. T. Sachrajda, F. Sanfilippo, S. Simula, and N. Tantalo, Light-meson leptonic decay rates in lattice QCD + QED, Phys. Rev. D **100**, 034514 (2019).
[15] C. Alexandrou *et al.* (Extended Twisted Mass Collaboration), Lattice calculation of the short and intermediate time-distance hadronic vacuum polarization contributions to the muon magnetic moment using twisted-mass fermions, Phys. Rev. D **107**, 074506 (2023).
[16] J. Dragos, T. Luu, A. Shindler, J. de Vries, and A. Yousif, Confirming the existence of the strong *CP* problem in lattice QCD with the gradient flow, Phys. Rev. C **103**, 015202 (2021).
[17] C. Alexandrou, A. Athenodorou, K. Hadjiyiannakou, and A. Todaro, Neutron electric dipole moment using lattice QCD simulations at the physical point, Phys. Rev. D **103**, 054501 (2021).
[18] T. Bhattacharya, V. Cirigliano, R. Gupta, E. Mereghetti, and B. Yoon, Contribution of the QCD Θ-term to the nucleon electric dipole moment, Phys. Rev. D **103**, 114507 (2021).
[19] S. Schaefer, R. Sommer, and F. Virotta (ALPHA Collaboration), Critical slowing down and error analysis in lattice QCD simulations, Nucl. Phys. **B845**, 93 (2011).
[20] S. Bacchio, P. Kessel, S. Schaefer, and L. Vaitl, Learning trivializing gradient flows for lattice gauge theories, Phys. Rev. D **107**, L051504 (2023).
[21] R. Abbott *et al.*, Normalizing flows for lattice gauge theory in arbitrary space-time dimension, arXiv:2305.02402.
[22] S. Bacchio, J. Finkenrath, and C. Stylianou, Lyncs-API: A PYTHON API for lattice QCD applications, Proc. Sci. LATTICE2021 (**2022**) 542.
[23] S. Yamamoto, S. Bacchio, and J. Finenrath, Running HMC Simulation with PYTHON via QUDA, Proc. Sci. LATTICE2022 (**2023**) 346.
[24] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, Solving Lattice QCD systems of equations using mixed precision solvers on GPUs, Comput. Phys. Commun. **181**, 1517 (2010).
[25] https://creativecommons.org/licenses/by/4.0/.