Heavy Wilson quarks and O(a) improvement: nonperturbative results for b_{g}

Mattia Dalla Brida ⓘ,^{a} Roman Höllwieser,^{b} Francesco Knechtli,^{b} Tomasz Korzec,^{b} Stefan Sint^c and Rainer Sommer ⓘ^{d,e}

^{a}*Theoretical Physics Department, CERN,*
  1211 Geneva 23, Switzerland
^{b}*Deptartment of Physics, University of Wuppertal,*
  Gaußstrasse 20, 42119 Germany
^{c}*School of Mathematics and Hamilton Mathematics Institute, Trinity College Dublin,*
  Dublin 2, Ireland
^{d}*Deutsches Elektronen-Synchrotron DESY,*
  Platanenallee 6, 15738 Zeuthen, Germany
^{e}*Institut für Physik, Humboldt-Universität zu Berlin,*
  Newtonstr. 15, 12489 Berlin, Germany

  E-mail: mattia.dalla.brida@cern.ch, hoellwieser@uni-wuppertal.de,
  knechtli@physik.uni-wuppertal.de, korzec@uni-wuppertal.de,
  sint@maths.tcd.ie, rainer.sommer@desy.de

ABSTRACT: With Wilson quarks, on-shell O(a) improvement of the lattice QCD action is achieved by including the Sheikholeslami-Wohlert term and two further operators of mass dimension 5, which amount to a mass-dependent rescaling of the bare parameters. We here focus on the rescaled bare coupling, $\tilde{g}_0^2 = g_0^2(1 + b_{\mathrm{g}} a m_{\mathrm{q}})$, and the determination of $b_{\mathrm{g}}(g_0^2)$, which is currently only known to 1-loop order of perturbation theory. We derive suitable improvement conditions in the chiral limit and in a finite space-time volume and evaluate these for different gluonic observables, both with and without the gradient flow. The choice of β-values and the line of constant physics are motivated by the ALPHA collaboration's decoupling strategy to determine $\alpha_s(m_Z)$ [1]. However, the improvement conditions and some insight into systematic effects may prove useful in other contexts, too.

Contents

1 Introduction

On-shell O(a) improvement of lattice QCD with Wilson quarks near the chiral limit requires a single new term in the action, the Sheikholeslami-Wohlert term with coefficient c_sw [2]. The latter is a function of the bare coupling g_0^2 and can be determined by imposing continuum chiral symmetry at finite lattice spacing [3]. With N_f degenerate quarks of subtracted bare mass $m_\mathrm{q} = m_0 - m_\mathrm{cr}$, there are 2 further counterterms which can be implemented as a rescaling of the bare parameters, i.e.

$$\tilde{g}_0^2 = g_0^2(1 + b_\mathrm{g}(g_0^2)am_\mathrm{q}), \qquad \widetilde{m}_\mathrm{q} = m_\mathrm{q}(1 + b_\mathrm{m}(g_0^2)am_\mathrm{q})\,. \tag{1.1}$$

In practice, perturbative estimates of these b-coefficients are often sufficient, as am_q is a small parameter for the light quark flavours up, down and strange. For the most common

gauge actions, i.e. Wilson's plaquette or the tree-level $O(a^2)$ improved Lüscher-Weisz actions, the one-loop result for b_g reads [4],

$$b_g(g_0^2) = 0.01200 \times N_f g_0^2 + O(g_0^4). \qquad (1.2)$$

For heavier quarks, am_q is no longer small, and in the recent decoupling project of the ALPHA collaboration, values of am_q up to $0.3 - 0.4$ are included in the analysis [1, 5]. In such situations a non-perturbative determination of the relevant b-coefficients becomes very desirable as, otherwise, uncontrolled systematic errors may ensue. In particular, the decoupling result for $\alpha_s(m_Z)$ [1] includes a sizeable systematic error due to an assumed 100 percent uncertainty on the truncated perturbative result for b_g, eq. (1.2).

Various b-coefficients have been determined non-perturbatively over the past 25 years, including b_m and some of the coefficients required for the $O(a)$ improvement of quark bilinear composite operators (see [6–14] for an incomplete list of references). The coefficient b_g is a notable exception. While there exist some qualitative ideas in the literature as to how b_g could be determined by measurement (see e.g. [15]), no proof of concept regarding their viability and practicality has been given. This is somewhat surprising given the central rôle of b_g for the consistency of $O(a)$ improvement: it is required when the quark masses are varied at fixed lattice spacing such as needed for chiral extrapolations or in studies of decoupling [1, 5]. Furthermore, since constant lattice spacing amounts to keeping $\tilde{g}_0^2$ constant, the renormalization constants in mass-independent schemes depend on $\tilde{g}_0^2$ rather than g_0^2. In fact, for composite operators this often leads to the determination of effective b-coefficients, which include the b_g-contribution from the renormalization constant by a Taylor expansion to first order in am_q [11], even if not mentioned explicitly [12].

In this paper we propose a class of improvement conditions to determine b_g non-perturbatively and then proceed to evaluate these for $N_f = 3$, non-perturbatively $O(a)$ improved lattice QCD [16] with the tree-level $O(a^2)$ improved Lüscher-Weisz gauge action [17]. We perform consistency checks by looking at gradient flow observables as well as Creutz ratios in a finite space time volume and for different lines of constant physics. We also generated some data in the small coupling region in order to compare to perturbation theory. Our parameter choices, in particular the range of lattice spacings, are motivated by the decoupling project [1, 5] and do not overlap with the parameters of CLS [18, 19].

The paper is organized as follows. In section 2 we discuss improvement conditions for b_g. In particular we review the connection to the $O(a)$ improvement term in the flavour singlet scalar density, which clarifies and corrects the discussion in [20]. We then define our set of observables (section 3) and present the chosen line of constant physics and the corresponding simulation parameters (section 4). In section 5, we discuss some details of the data analysis and various consistency checks before we present our results for b_g (section 6) and our conclusions (section 7). A few technical details have been delegated to appendices: appendix A gives some details pertaining to the derivation of the key equations in section 2. Appendix B contains some additional information on the simulations and corresponding parameters, while the simulation results are tabulated in appendix C.

2 Improvement conditions for b_{g}

2.1 Chiral Ward identities and b_{g}

While improvement coefficients such as c_{sw} or c_{A}, multiplying the $\mathrm{O}(a)$ counterterms in the action and to the axial current, respectively, can be determined by chiral Ward identities [3], this is less obvious for the b-coefficients multiplying am_{q}. One problem is that, away from the chiral limit, the on-shell condition is potentially violated by unavoidable contact terms, at least for chiral Ward identities beyond the simple PCAC relation. A systematic analysis for the $N_{\mathrm{f}} = 3$ lattice QCD action and all quark bilinear composite fields has been carried out by Bhattacharya et al. in [20]. By studying the general case of mass non-degenerate quarks, they show how violations of the on-shell condition can be by-passed and they list all combinations of renormalization constants and improvement coefficients which are, at least in principle, determined by chiral Ward identities. The coefficient b_{g} is included in this list, through its relation to the $\mathrm{O}(a)$ counterterm of the flavour singlet scalar density, $S^0(x) = \overline{\psi}(x)\psi(x)$. In the mass-degenerate limit, the corresponding renormalized and $\mathrm{O}(a)$ improved operator takes the form [20, 21],

$$S_{\mathrm{R}}^0 = Z_{\mathrm{S}^0}\left(1 + b_{\mathrm{S}^0}am_{\mathrm{q}}\right)\left(S^0 + c_{\mathrm{S}}a^{-3} + d_{\mathrm{S}}a\operatorname{tr}\left\{F_{\mu\nu}F_{\mu\nu}\right\}\right). \tag{2.1}$$

Using an argument based on Feynman-Hellmann type identities, Bhattacharya et al. derive the identity[1]

$$b_{\mathrm{g}} = -2g_0^2 d_{\mathrm{S}}, \tag{2.2}$$

and d_{S} can, in prinicple, be determined by chiral Ward identities, together with c_{S} and the scale independent ratio $Z_{\mathrm{P}}/Z_{\mathrm{S}^0}$ [21]. A difficulty is posed by the fact that the chiral limit is only defined up to cutoff effects of $\mathrm{O}(a^2)$, which render the subtraction of the cubic power divergence ambiguous. More precisely, with $c_{\mathrm{S}} \equiv c_{\mathrm{S}}(g_0^2, am_{\mathrm{q}})$, an expansion in powers of am_{q} yields a term $\propto am_{\mathrm{q}}/a^3 = m_{\mathrm{q}}/a^2$. Given the intrinsic $\mathrm{O}(a^2)$ ambiguity of m_{q}, this results in an ambiguity of $\mathrm{O}(1)$ in the subtraction term rendering the definition of the renormalized scalar density impossible along these lines. It is therefore important to note that one may combine Ward identities such that only connected correlation functions remain, in which all power divergences cancel. A practical implementation using SF boundary conditions [22, 23] can be obtained following the discussion in [21]. Whether these observations lead to a practical method for a non-perturbative determination of b_{g} remains to be seen. We emphasize that this strategy relies entirely on the identity (2.2). We would therefore like to point out a subtlety that was overlooked in [20]. Given that one relates an improved operator to the improved action parameters, it is natural to expect that the $\mathrm{O}(a)$ counterterm in eq. (2.1) cannot be discretized arbitrarily, but must be related to the derivative of the lattice action density with respect to g_0^2. For definiteness, the $\mathrm{O}(a)$ improved lattice action

$$S = S_g + a^4\sum_x \overline{\psi}(x)\left(D_{\mathrm{W}} + m_0 + c_{\mathrm{sw}}\frac{ia}{4}\sigma_{\mu\nu}F_{\mu\nu}(x)\right)\psi(x), \tag{2.3}$$

---

[1]In the notation of [20], $g_{\mathrm{S}} = -d_{\mathrm{S}}$, due to their convention of hermitian, rather than anti-hermitian $F_{\mu\nu}$.

includes the lattice gauge action $S_g = a^4 \sum_x \mathcal{L}_g(x)$ with $\mathcal{L}_g$ being the lattice version of the (Euclidean) Yang-Mills Lagrangian density in the continuum,

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2g_0^2} \operatorname{tr}\left\{F_{\mu\nu}(x)F_{\mu\nu}(x)\right\}. \tag{2.4}$$

We find that, due to the g_0-dependence of c_{sw}, a contribution from the fermionic part of the action must be included, so that the correct lattice discretization of the counterterm is given by,

$$\operatorname{tr}\left\{F_{\mu\nu}F_{\mu\nu}\right\} \quad \rightarrow \quad -2g_0^2 \left(\mathcal{L}_g - g_0^2 \times \frac{ia}{4}c_{\text{sw}}'\overline{\psi}\sigma_{\mu\nu}F_{\mu\nu}\psi\right), \tag{2.5}$$

with $c_{\text{sw}}' \equiv \frac{d}{dg_0^2}c_{\text{sw}}$. For future reference we provide a detailed derivation in appendix A.

2.2 Restoration of chiral symmetry in a small volume

In this paper we pursue a different approach. It is based on the observation that, in the absence of spontaneous chiral symmetry breaking, physical quark mass effects in gluonic observables are quadratic or higher order in the quark mass. In contrast, the counterterm proportional to b_{g} is designed to cancel terms that are linear in the quark mass, allowing us, in principle, to distinguish these effects. Let us first have a closer look at the physical quark mass dependence. If we consider a finite space time manifold without boundaries, the lattice QCD partition function with Ginsparg-Wilson type quarks [24–28] becomes a finite dimensional, mathematically well-defined ordinary and Grassmann-integral, with exact chiral and flavour symmetries. The absence of spontaneous symmetry breaking in a finite volume, implies that the partition function is an analytic function of m_{q}. For even N_{f}, a change of the fermionic variables by a discrete chiral field transformation[2]

$$\psi \rightarrow \gamma_5\psi, \qquad \overline{\psi} \rightarrow -\overline{\psi}\gamma_5, \tag{2.6}$$

then establishes that this function must be even, implying the absence of terms linear in m_{q} for any gluonic observable. With odd $N_{\text{f}} > 2$, this is no longer true; however, by adapting the argument of appendix D in ref. [29] to the lattice regularization with Ginsparg-Wilson quarks, the discrete chiral field transformation,

$$\psi \rightarrow \exp\left(i\frac{\pi}{N_{\text{f}}}\gamma_5\right)\psi, \qquad \overline{\psi} \rightarrow \overline{\psi}\exp\left(i\frac{\pi}{N_{\text{f}}}\gamma_5\right), \tag{2.7}$$

leads to a change of variables with unit Jacobian that leaves the massless action invariant. Under the same transformation, a single insertion of the flavour singlet scalar density, S^0, into a gluonic correlation function is proportional to itself and must vanish, provided that both the QCD action and the gluonic observable are parity even.[3] Although this does not exclude all odd powers in the quark mass, it means that contributions linear in the quark mass are indeed absent.

---

[2]On the lattice with Ginsparg-Wilson quarks and Neuberger's (γ_5-hermitian) Dirac operator D_N [26], the first of the γ_5 factors is replaced by $\hat{\gamma}_5 \equiv \gamma_5(1 - aD_N)$ [28]. Note that, for even N_{f}, the transformation is part of the *non-singlet* chiral symmetry.

[3]We again use a continuum notation; on the lattice, besides using $\hat{\gamma}_5$ [28] for the transformation of ψ, the scalar lattice density takes the form $S^0 = \overline{\psi}(1 - \frac{a}{2}D_N)\psi$, such that it correctly transforms under the lattice chiral symmetry (see e.g. section 2 in [30]).

We will denote a generic renormalized gluonic observable by O_g. Gradient flow observables [31, 32], such as the gauge action density at finite flow time [32], are natural candidates, but the gradient flow is not a necessary requirement. We choose a finite Euclidean space time volume L^4 with boundary conditions that do not break all the chiral symmetries. A hyper-torus geometry with some kind of periodic boundary conditions for all fields is an obvious possibility, however, other possible options include chirally rotated SF (χSF) boundary conditions (with even $N_{\rm f}$) [33, 34], or a mixture of SF [22, 23] and χSF boundary conditions (with odd $N_{\rm f}$) [35]. For simplicity we also assume that O_g has no mass dimension, which can always be achieved by multiplication with the appropriate power of L. In the continuum limit, the observable thus becomes a function of two dimensionless variables, which we may choose as $z = ML$ and $\Lambda_{\overline{\rm MS}}^{(3)} L$, with M and $\Lambda_{\overline{\rm MS}}^{(3)} \equiv \Lambda$ being the Renormalization Group Invariant (RGI) quark mass and Λ-parameter of the $N_{\rm f} = 3$ theory, respectively. From the previous discussion we then expect that the continuum limit of the lattice expectation value, $\langle O_g \rangle$, becomes a function of $(z, \Lambda L)$, with an expansion in the RGI mass of the form

$$\langle O_g \rangle = \langle O_g \rangle_{z=0} + A(\Lambda L) z^2 + {\rm O}(z^3) \,. \tag{2.8}$$

When considering this observable on the lattice, the mass dependence will have a linear term in z, unless $b_{\rm g}$ is chosen correctly. Hence, a possible improvement condition for $b_{\rm g}$ is given by

$$\left. \frac{\partial \langle O_g \rangle}{\partial z} \right|_{z=0, L\Lambda = {\rm const}} = 0 \,. \tag{2.9}$$

In order to derive an explicit equation for $b_{\rm g}$, we will proceed in two steps. First we keep the lattice size L/a and spacing a fixed. Up to ${\rm O}(a^2)$ effects this can be achieved by working at fixed $\tilde{g}_0^2$. The choice for a line of constant physics only matters later, when $\tilde{g}_0^2$ and thus the lattice spacing is changed. This will be discussed in section 4. Since z is proportional to $\widetilde{m}_{\rm q}$, we may change variables from $(z, \Lambda L, L/a)$ to $(a\widetilde{m}_{\rm q}, \tilde{g}_0^2, L/a)$ and the improvement condition at a given lattice spacing then reads

$$\left. \frac{\partial \langle O_g \rangle}{\partial a\widetilde{m}_{\rm q}} \right|_{\widetilde{m}_{\rm q}=0, \tilde{g}_0^2} = 0 \,. \tag{2.10}$$

We would now like to perform a final change of variables from improved to unimproved bare parameters. In fact, it is technically convenient to do the reverse transformation and then solve a 2×2 linear system. For the sake of readability we have relegated this discussion to appendix A. The improvement condition then implies,

$$b_{\rm g}(g_0^2) = \left. \frac{\partial \langle O_g \rangle}{\partial am_{\rm q}} \right|_{g_0^2, m_{\rm q}=0} \times \left[g_0^2 \left. \frac{\partial \langle O_g \rangle}{\partial g_0^2} \right|_{m_{\rm q}=0} \right]^{-1} \,. \tag{2.11}$$

This is the desired equation for $b_{\rm g}$ in terms of the bare parameters, which are the ones directly controlled in numerical simulations. Note that the condition is again formulated in the chiral limit, $m_{\rm q} = 0$.

At this point we recall that, with Wilson quarks, the massless limit is not unambiguously defined. What is required here is a definition up to an ${\rm O}(a^2)$ ambiguity, which can be

achieved by tuning the improvement coefficients c_{sw} in the action and c_{A} in the improved axial current used for the definition of the PCAC mass [3]. Note that any remnant O(a) ambiguity of the chiral limit would make the quadratic physical mass effect vanish only up to O(a) corrections, which, by virtue of the $1/a$ factor in the am_{q}-derivative would contribute at O(1), leading to a wrong result for the b_{g}-estimate.

A number of choices must still be made. First, one needs to pick suitable gluonic observables, O_g. Second, one needs to find a practical way to implement the derivatives with respect to the bare parameters and then one needs to follow a line of constant physics as $\beta = 6/g_0^2$ and thus the lattice spacing is changed.

3 Choice of observables

We choose a space-time volume with linear extent L in all directions and a hyper-torus topology. We use periodic boundary conditions for the gauge fields, and antiperiodic boundary conditions for the fermions in all 4 space-time directions. The absence of a boundary avoids potential violations of chiral symmetry, while antiperiodic boundary conditions for the fermions allow us to perform simulations around the chiral limit (cf. section 4.2 and appendix B). To finish the description of our set-up it remains to specify our choices of gluonic observables, O_g.

3.1 Gradient flow energy density

First, we consider the action density in terms of the gradient flow field tensor $G_{\mu\nu}(t,x)$. In a continuum language it is given by

$$\sigma(c) = \sum_{\mu,\nu=0}^{3} \frac{t^2 \left\langle \mathrm{tr}\left\{G_{\mu\nu}(t,x)G_{\mu\nu}(t,x)\right\} \delta_{Q(t),0}\right\rangle}{\left\langle \delta_{Q(t),0}\right\rangle}\Bigg|_{8t=c^2L^2}. \tag{3.1}$$

In the numerical evaluation translation invariance w.r.t. x is of course used. The projection to the sector of vanishing topological charge $Q(t)$ is performed to avoid large autocorrelation times as discussed in [36, 37]. While the motivation is algorithmic, it is part of the definition of the observable. The gradient flow observable is finite: proven to all orders of perturbation theory [38] and well established in numerical simulations. The remaining parameter c fixes the ratio between the scale set by the flow time and the box length [39]. On the lattice we discretize the gradient flow equations through the O(a^2) improved Zeuthen flow definition, while the action density of the flow fields is obtained from an O(a^2) improved combination of plaquette and clover discretizations (see [37, 40] for the details).

3.2 Creutz ratios

To check for O(a) ambiguities in b_{g} we also use a second observable. It is very independent from σ, namely it is defined without the gradient flow. We start from rectangular Wilson loops $W(R,T)$ and form Creutz ratios [41]

$$\chi(R,T) = -\tilde{\partial}_R \tilde{\partial}_T \log(W(R,T)), \tag{3.2}$$

where the lattice derivatives $\tilde{\partial}_y f(y) = [f(y+a) - f(y-a)]/(2a)$ are symmetric to avoid linear terms in a. Note that, analogously to the energy density at positive flow time in eq. (3.1), also

the Wilson loops $W(R, T)$ defining the Creutz ratios are projected to the topologically trivial sector. These Creutz ratios are made dimensionless and restricted to the diagonal ones,

$$\hat{\chi} = \left[R^2 \chi(R, R) \right]_{R=L/4} . \tag{3.3}$$

We have fixed the ratio R/L to $1/4$ since this is a practical number for lattices with $L/a = 12, 16, \ldots$, which we will use below. Also $\hat{\chi}$ is a finite pure gauge observable for the following reasons. In continuous space-time, Wilson loops around smooth paths have been shown to be finite up to an overall factor $Z_W(l)$ which depends (apart from coupling and regularisation parameter) only on the length l of the path [42]. When the path contains cusps, such as the rectangular Wilson loops, there is an additional renormalization factor depending on the angles of the cusps. Both factors are removed by the derivatives in eq. (3.2). In our implementation on the lattice, we apply one HYP2 smearing step with parameters $\alpha_1 = 1$, $\alpha_2 = 1$ and $\alpha_3 = 0.5$ to the link variables which form the loop [43–47]. Since this does not change the symmetry properties of the links, $\hat{\chi}$ remains finite and has the same continuum limit as the observable with no smearing. Finally we need also the absence of terms linear in a in the observable. Their absence for the static quark potential has been explained in [48]. Analogously it holds for Creutz ratios defined with symmetric derivatives.

4   Line of constant physics (LCP) and simulations

4.1   LCP

The Symanzik expansion applies when the lattice spacing a is changed with all other scales being fixed. Consequently also improvement conditions need to be enforced along a LCP. Since vanishing quark masses are required for our improvement condition, it only remains to fix L in physical units.[4] To this end we choose $\bar{g}_{\mathrm{GF}}^2(L) = 3.949$, where $\bar{g}_{\mathrm{GF}}^2(L)$ is the gradient flow coupling defined in a finite volume with spatial extent L and SF boundary conditions (cf. [1] for a precise definition). For resolutions $L/a \in \{12, 16, 20, 24\}$ the required β-values are found in table 13 of [1] with sufficient precision. They are in the range $\beta \in [4.302, 4.7141]$. For larger L/a, it turned out that a precise computation of b_{g} becomes prohibitively expensive. We therefore deviate from the strict LCP and set $L/a = 24$ for $\beta \geq 4.9$. At that point, the lattice spacing in units of the Λ-parameter is very small and also the only additional scales determining a-ambiguities, $\sqrt{8t}$ as well as R, are large enough to make these ambiguities sufficiently small. We show numerical tests in section 5. Furthermore, the extracted values for b_{g} approach the one-loop perturbative values smoothly. All-in-all it is sufficient to determine b_{g} for the parameters listed in table 1.

---

[4]The values, κ_c, of the hopping parameter where the quark mass vanishes are taken from a fit similar to the one described in appendix A.1.4 of [37]. The data are the same, but instead of separate fits at each value of L/a, we performed a global fit in which the lattice artifacts are parameterized by $(a/L)^3$ times a power series in g_0^2. We include the six lowest powers and such a fit describes the data well, is in good agreement with the previous fit in the range of β and L/a values where we have data, and most importantly allows for inter/extrapolations in L/a. We have checked that the uncertainties of κ_c are negligible in our analysis. This statement refers to the statistical precision as well as to systematics coming from data interpolation and extrapolation to large values of L/a at the larger β's.

L/a	β	$\kappa = \kappa_c$	c_1	c_2	$\hat{c}_2$
12	4.3020	0.135998	0.02391(26)	0.078(18)	0.000191(45)
16	4.4662	0.135607	0.02049(22)	0.108(15)	0.000141(20)
20	4.5997	0.135289	0.01899(79)	0.043(61)	0.000036(50)
24	4.7141	0.135024	0.01796(38)	0.261(33)	0.000144(18)
24	4.9000	0.134601	0.01331(56)	0.163(38)	$-$
24	5.0671	0.134241	0.01166(48)	0.083(33)	$-$
24	5.1719	0.134025	0.00932(43)	0.123(29)	$-$
24	6.0000	0.132575	0.00428(30)	0.053(21)	$-$
24	8.0000	0.130391	0.00134(12)	0.0214(85)	$-$
24	16.000	0.127496	0.00009(3)	0.0016(21)	$-$

Table 1. Bare parameters for which b_{g} has been computed as well as the fit coefficients c_1, c_2 for $\sigma(0.18)$ (cf. eq. (5.1)) and $\hat{c}_2 = (a/L)^2 Z_{\mathrm{m}}^{-2} c_2$, where Z_{m} is the quark mass renormalization in the SF-scheme at the renormalization scale μ_{dec} (see subsection 4.1 for the details).

4.2 Simulations

We simulate QCD with a Lüscher-Weisz gauge action [17] and three flavours of O(a) improved Wilson quarks [2, 16]. To implement the strategy described above, simulations of QCD at zero quark mass or even at slightly negative masses are necessary. A finite volume with suitably chosen boundary conditions is indispensable to prevent negative or very small eigenvalues of the Wilson Dirac operator. With Schrödinger Functional boundary conditions, massless simulations of our action on our LCP were unproblematic [37]. Here, in order to avoid chiral symmetry breaking by the boundaries, we work on a hyper-torus. Choosing anti-periodic boundary conditions for the fermionic fields in all four directions induces a sufficiently large spectral gap.

The simulation setup is similar to previous simulations with $N_{\mathrm{f}} = 3$ degenerate quarks. We use an even-odd preconditioned variant of the HMC algorithm [49] in which the three quarks are separated into a doublet and a third quark. The quark determinant for the doublet is factorized [50] into three factors $\det[\hat{D}^\dagger \hat{D}] = \det[\hat{D}^\dagger \hat{D} + \mu_2^2] \times \det[(\hat{D}^\dagger \hat{D} + \mu_1^2)(\hat{D}^\dagger \hat{D} + \mu_2^2)^{-1}] \times \det[(\hat{D}^\dagger \hat{D})(\hat{D}^\dagger \hat{D} + \mu_1^2)^{-1}]$ and pseudo fermion fields are introduced for each factor. Here $\hat{D}$ denotes the even-odd preconditioned clover Wilson operator. For the third quark it is required that all eigenvalues of $\hat{D}$ have real parts that are larger than some $r_a > 0$. The operator in the determinant $\det[\hat{D}] = \det[\sqrt{\hat{D}^\dagger \hat{D}}]$ can then be approximated by a rational function of $\hat{D}^\dagger \hat{D}$ [51, 52]. The inexactness of this approximation is accounted for by including a stochastically determined reweighting factor in the observables. The molecular dynamics equations derived from the various contributions to the action are solved using multi level integration schemes with two levels [53]. The gauge action is integrated with the finest resolution while forces from different pseudo-fermion actions are evaluated less frequently. We use a modified version of `openQCD-1.6` [54] for our simulations and refer to its documentation for further details. Simulation parameters as well as tests concerning the spectral gap and the ergodicity of the simulations are collected in appendix B.
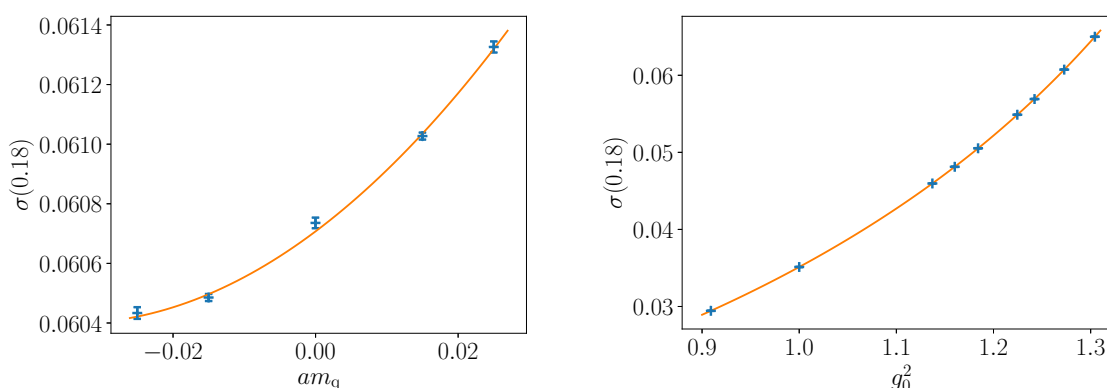
Figure 1. Quadratic fit of $\sigma(0.18)$ vs. am_{q} for $L/a = 24, \beta = 4.7141$ (left), and nine-point fit with $n_g = 5$ of $\sigma(0.18)$ vs. g_0^2 for $L/a = 24$ and $4.6 \leq \beta \leq 6.6$ (cf. table 5) (right).

5 Data analysis

5.1 Estimates for the derivatives entering eq. (2.11)

In principle the derivatives can be obtained from insertions of $\partial_p S$, where p denotes the bare quark mass and the bare coupling. However, such measurements have large statistical errors. Instead, for a given L/a, we simulate at neighboring values in g_0^2 and m_{q}, fit the data as a function of m_{q} at fixed g_0^2 and vice versa, and then obtain the derivative as the derivative of the fit functions. In detail we fit

$$\langle O_g \rangle|_{L/a, g_0^2} = \sum_{k=0}^{n_m - 1} c_k \, (am_{\mathrm{q}})^k, \qquad \langle O_g \rangle|_{L/a, m_{\mathrm{q}}=0} = \sum_{k=1}^{n_g} d_k \, g_0^{2k}, \qquad (5.1)$$

to the data in tables 5 and 6 and mostly take the results with $n_m = n_g = 3$ to three data points.

Let us discuss this choice as well as the exceptions from it. For the determination of the quark mass derivative the fitted data are generically at $am_{\mathrm{q}} \approx 0, \pm 0.025$. We need to check that this is small enough to obtain the first derivative $\partial_{am_{\mathrm{q}}}\langle O_g \rangle \approx c_1$, namely that we are not affected by higher order terms in am_{q} beyond $n_m = 3$. Indeed, note that the variable am_{q} in eq. (5.1) is natural for the linear term in am_{q} that we are seeking. However, already the quadratic term in the mass exists in the continuum limit. It is then naturally written in the form $\hat{c}_2(L\overline{m})^2 = c_2(am_{\mathrm{q}})^2$ with some renormalized quark mass $\overline{m}$ and a coefficient $\hat{c}_2(a/L)$ which has a finite limit $\hat{c}_2(0)$. Therefore, as we increase L/a, the coefficient c_2 increases in magnitude roughly at the rate $(L/a)^2$ (up to logarithmic renormalization effects). At $\beta = 4.7141$ we simulated five quark mass points with $am_{\mathrm{q}} \approx 0, \pm 0.015, \pm 0.025$. The corresponding data for $\sigma(0.18)$ are shown in figure 1. One sees significant curvature, but also that a purely quadratic fit works well. In table 1 we list both c_2 and $\hat{c}_2 = (a/L)^2 Z_{\mathrm{m}}^{-2} c_2$ for the strict LCP.[5] Apart from the small statistics $L/a = 20$ results, the expected behavior is clearly seen. It implies that the largest contamination of our data by higher order effects in am_{q}

---

[5]We define the quark mass renormalization in the SF-scheme, $Z_{\mathrm{m}} = Z_{\mathrm{m}}^{\mathrm{SF}} = Z Z_{\mathrm{A}}/Z_{\mathrm{P}}$, at the renormalization scale μ_{dec}, exactly as in [1], and also use the numerical values for $Z_{\mathrm{A}}, Z_{\mathrm{P}}$ found in that reference.

L/a	β	$g_0^2 \partial\sigma/\partial g_0^2$		b_{g}	
		$n_g = 5$	$n_g = 3$	$n_g = 5$	$n_g = 3$
24	4.7141	0.16474(46)	0.16531(58)	0.1090(23)	0.1086(24)
24	4.9000	0.13907(20)	0.1379(11)	0.0957(40)	0.0965(42)
24	5.0671	0.12132(23)	0.12195(63)	0.0961(40)	0.0957(40)
24	5.1719	0.11211(20)	0.11246(53)	0.0831(39)	0.0828(39)
24	6.0000	0.06835(16)	0.068680(86)	0.0627(43)	0.0624(43)

Table 2. Comparison of b_{g} and corresponding g_0^2 derivatives of $\sigma(0.18)$ from 3-point $n_g = 3$ and 9-point $n_g = 5$ fits, for $L/a = 24$ and different values of β (cf. table 5).

is present for $L/a = 24$. Therefore we checked explicitly for $\beta = 4.7141, L/a = 24$, that we obtain the same derivative by fitting the available five points with $n_m = 3$ and also by taking just the symmetric derivative, $\tilde{\partial}_{am_{\mathrm{q}}} f = [f(am_{\mathrm{q}}) - f(-am_{\mathrm{q}})]/(2am_{\mathrm{q}})$ with either $am_{\mathrm{q}} = 0.015$ or $am_{\mathrm{q}} = 0.025$. The resulting estimates for the derivative agree well within errors:

$$\tilde{\partial}_{am_{\mathrm{q}}}\sigma\big|_{am_{\mathrm{q}}=0.025} = 0.01786(54)\,, \qquad \tilde{\partial}_{am_{\mathrm{q}}}\sigma\big|_{am_{\mathrm{q}}=0.015} = 0.01806(55)\,, \qquad (5.2)$$

$$\partial_{am_{\mathrm{q}}}\sigma\big|_{\mathrm{5pts}}^{\mathrm{fit}} = 0.01796(38)\,. \qquad (5.3)$$

Altogether we conclude that systematic errors due to this step are negligible and use the five-point fit at $\beta = 4.7141$ and the three-point fit otherwise.

The g_0^2 derivatives are obtained from the fits with $n_g = 3$ to the data at the three closest values of g_0^2 of the data compiled in table 5. For the case of $L/a = 24$, we compared these with the estimates from $n_g = 5$, fitting all nine available values of g_0^2 in the range specified by $4.6 \leq \beta \leq 6.6$ (cf. table 5). The comparison, listed in table 2 for $\sigma(0.18)$, shows that the estimates for the g_0^2 derivative agree within uncertainties and the changes in the resulting b_{g} are negligible compared to the overall errors, which are dominated by the ones of the quark mass derivative. The nine-point fit with $n_g = 5$ is illustrated in figure 1.

The data for $\hat{\chi}$ are analyzed in the same way. Results for b_{g} from both observables are listed in table 3. As checks on the dependence on c, R/L, and L/a-values deviating from the lines of constant physics, there are additional data not reported in the tables.

5.2 Relaxation of LCP condition for $\beta \geq 4.9$

For our application to decoupling [1], we also need bare couplings below $g_0^2 = 6/4.7$. Strictly following the LCP then requires prohibitively large L/a, in particular as we seek the connection to perturbation theory below $g_0^2 = 1$. We therefore keep $L/a = 24$ fixed for all $\beta \geq 4.7$. In principle this means that there is an ambiguity of order $a/L = 1/24$ which does not vanish as we increase β and therefore decrease a. We follow a skewed trajectory where L decreases.

The magnitude of the ambiguity can be tested by computing how b_{g} varies as one changes L. At fixed a, we therefore added computations with $L/a = 16, 32$. In figure 2, we show b_{g} from σ, for $L/a = 16, 24, 32$ and $\beta = 4.9$. Within a small margin, a universal behavior, independent of L/a, is seen when b_{g} is considered at fixed $\sqrt{8t}/a$. Therefore it is

L/a	β	κ	$g_0^2\partial\langle O_g\rangle/\partial g_0^2$	$\partial\langle O_g\rangle/\partial am_{\mathrm{q}}$	b_{g}
12	4.3020	0.135998	0.17712(17)	0.02392(26)	0.1350(15)
			1.1560(31)	0.1661(47)	0.1437(41)
16	4.4662	0.135607	0.16429(23)	0.02049(22)	0.1247(13)
			1.1825(51)	0.1518(48)	0.1283(41)
20	4.5997	0.135289	0.1631(11)	0.01898(79)	0.1163(49)
			1.269(24)	0.167(19)	0.132(15)
24	4.7141	0.135024	0.16531(58)	0.01796(39)	0.1086(24)
			1.268(15)	0.1398(95)	0.1103(76)
24	4.9000	0.134601	0.1379(11)	0.01331(56)	0.0965(42)
			1.085(31)	0.099(14)	0.091(13)
24	5.0671	0.134241	0.12195(63)	0.01167(48)	0.0957(40)
			0.948(17)	0.098(13)	0.104(14)
24	5.1719	0.134025	0.11246(53)	0.00931(43)	0.0828(39)
			0.871(15)	0.078(11)	0.090(13)
24	6.0000	0.132575	0.068680(86)	0.00428(30)	0.0624(43)
			0.5497(25)	0.0272(83)	0.049(15)
24	8.0000	0.130391	0.033139(33)	0.00134(12)	0.0405(36)
			0.2739(11)	0.0126(37)	0.046(14)
24	16.000	0.127496	0.0101876(67)	0.000091(30)	0.0089(30)
			0.08898(21)	0.00143(99)	0.016(11)

Table 3. Results for b_{g}, eq. (2.11), and corresponding derivatives w.r.t. the bare parameters from $\langle O_g\rangle = \sigma(0.18)$ and $\langle O_g\rangle = \hat{\chi}$, in the first and second line for each value of L/a and β, respectively.

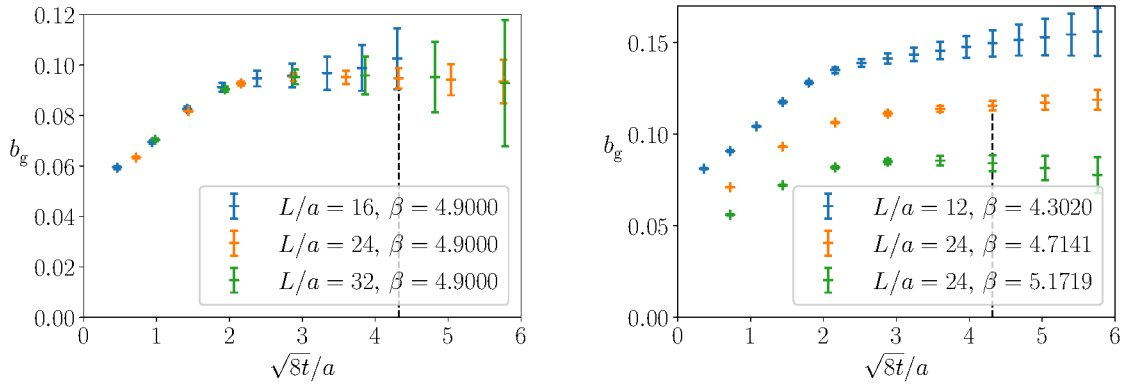**Figure 2.** b_{g}, eq. (2.11), from $\langle O_g\rangle = \sigma$ vs. $\sqrt{8t}/a$ at $\beta = 4.9$ for $L/a = 16$, 24, and 32 (left) and at $\beta = 4.3020$, 4.7141, 5.1719 for $L/a = 12$, 24, and 24 (right). The vertical dashed line indicates $\sqrt{8t}/a = 0.18 \times 24$, as used for all $\beta \geq 4.7$ in the determination of b_{g}.
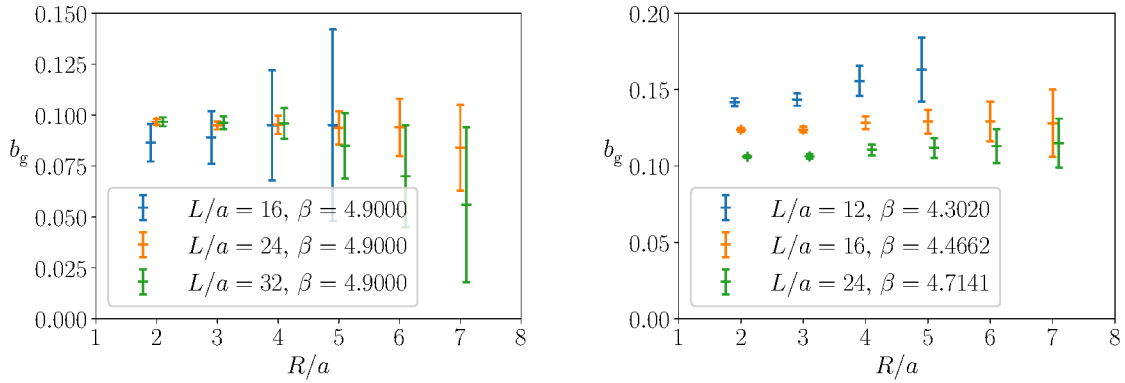
Figure 3. b_{g}, eq. (2.11), from $\langle O_g \rangle = R^2 \chi(R, R)$ vs. R/a at $\beta = 4.9$ for $L/a = 16$, 24, and 32 (left) and $\beta = 4.3020$, 4.4662, 4.7141 for $L/a = 12$, 16, and 24 (right).

actually not very relevant that we deviate from the fixed L trajectory. Rather we have to be concerned about the change in t (which is linked to L by $t = c^2 L^2$ along the strict trajectory). Fortunately, it is seen that b_{g} becomes approximately independent of t for $\sqrt{8t} \gtrsim 3a$. Since $c = 0.18$ puts us into the flat region for $L/a = 24$, there are good indications that our deviation from the strict LCP is irrelevant at the level of our uncertainties. In figure 3, we show the same investigation for b_{g} from $R^2 \chi(R, R)$, where the variable $\sqrt{8t}$ is replaced by R. The precision is not as good as in figure 2, but the statements made before hold, exchanging $\sqrt{8t} \to R$.

In summary, we may relax the strict LCP as long as we keep $\sqrt{8t}$ and R larger than about three in lattice units and have $L/a \geq 24$. The resulting ambiguity is small. We emphasize that above $g_0^2 = 6/4.9$, where the improvement by b_{g} is most relevant, we keep to the strict LCP.

6 Results

Our central results from the gradient flow energy density $\sigma(0.18)$ (cf. table 3) are displayed in figure 4. As an interpolation of the data from the perturbative regime up to our maximum bare coupling, we fit the results to

$$b_{\mathrm{g}}^\sigma = g_0^2 \left[b_{\mathrm{g}}^{(1)} + \sum_{i=1}^{n_k} k_i\, g_0^{2i} \right], \quad b_{\mathrm{g}}^{(1)} = 0.036\,. \tag{6.1}$$

A good fit is obtained already for $n_k = 2$ with parameters

$$k_1 = -0.0151\,, \quad k_2 = 0.0424\,. \tag{6.2}$$

This interpolation deviates little from others, obtained with e.g. $n_k = 3$ or from ones where we add a term proportional to a/L. The maximum (absolute) difference is 4×10^{-4} in the important range of $4.30 \leq \beta \leq 5.17$, and it is 1.1×10^{-3} overall. Apart from this interpolation uncertainty, it is relevant to check for the always present ambiguity of O(a). We display the difference to the determination from $\langle O_g \rangle = \hat{\chi}$ in figure 4. The ambiguity is small and consistent with $b_{\mathrm{g}}^\sigma - b_{\mathrm{g}}^{\hat{\chi}} = \mathrm{const.} \times a/L$. The uncertainty of the difference is dominated by $b_{\mathrm{g}}^{\hat{\chi}}$, but especially in the region of larger g_0^2, the test is rather significant and
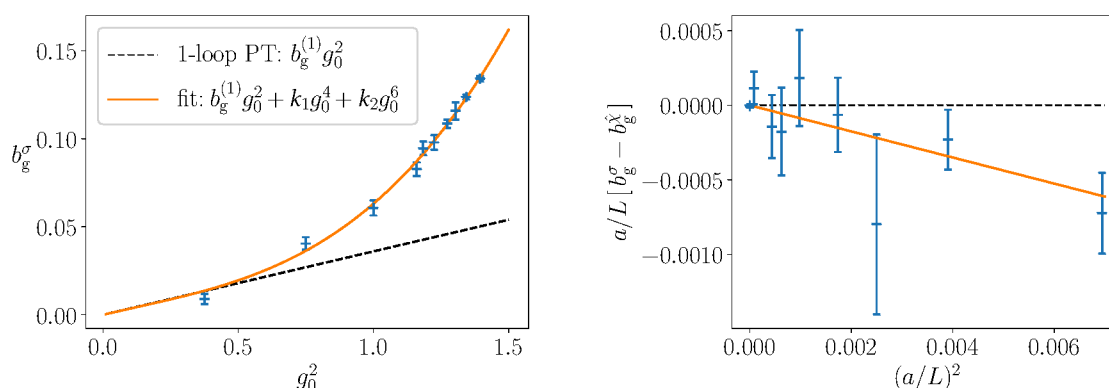
Figure 4. Final result for b_g from $\sigma(0.18)$ (left) and ambiguity $[b_\mathrm{g}^\sigma - b_\mathrm{g}^{\hat\chi}] \times \frac{a}{L}$ (right). The extra factor a/L is added because b_g always enters into observables with an explicit factor of a; we also include a linear fit to all points constrained to go to zero. The values of a/L considered in the plot have been obtained by fitting the values of L/a as a function of β_LCP in table 1 of ref. [1], and enforcing the universal 2-loop running of the lattice spacing for $g_0^2 \to 0$.

reassuring. The large difference of b_g from the one-loop approximation is not due to our particular choice of improvement condition.

7 Conclusions

For the renormalization-by-decoupling strategy [1, 5], massive and massless QCD have to be connected with high precision. In such problems the improved bare coupling $\tilde g_0^2$ needs to be kept fixed as one approaches the continuum limit and the improvement coefficient b_g is needed. The relevant combination $am_\mathrm{q}b_\mathrm{g}$ reaches up to 0.06 in the continuum extrapolation of the massive coupling in [1]. We have shown that b_g can be determined requiring the absence of a term linear in the quark mass around $m_\mathrm{q} = 0$ when boundary conditions do not break chiral symmetry and one is not effectively in the large volume region where chiral symmetry is broken spontaneously. We have further seen that with $L \approx 0.25\,\mathrm{fm}$, periodic boundary conditions for the gluon fields and anti-periodic ones for the quarks, the Dirac operator has a sufficiently large spectral gap such that the theory can be simulated around the chiral limit. Furthermore, the quark mass derivative could be extracted precisely and we obtained b_g with good precision. The result, eqs. (6.1) and (6.2), can be used without uncertainty, since subsequent assumptions made on the form of the continuum limit will be more relevant than our errors. Still, we have cited a systematic error which can be used if desired.

We emphasize that eq. (6.1) has not been shown to be valid below $\beta = 4.3$. Firstly, we simulated only starting at that β-value, and secondly the LCP would quickly demand very small $\sqrt{8t}/a$ where ambiguities are potentially very large. Extending the present method to the region of CLS simulations [18, 19] is therefore not entirely straightforward. One alternative are higher order perturbative estimates; for recent steps towards a 2-loop result cf. [55], and one of us (MDB) has derived a preliminary result for the Wilson gauge action using the code developed for the 2-loop calculation of the SF coupling [56, 57]. Another alternative

is the chiral Ward identity technique, applied to the O(a) improved flavour singlet scalar density, eq. (2.1), cf. [21]. The connection to $b_{\rm g}$ relies on the identity, eq. (2.2), which was first derived by Bhattacharya et al. in [20]. In section 2, we have pointed out an oversight in this original derivation. The identity can however be rescued, provided one uses the particular discretization of the O(a) counterterm to the scalar density, eq. (2.5), which includes a contribution from the Sheikholeslami-Wohlert term.

Acknowledgments

A   Derivation of eqs. (2.2) and (2.11)

In this appendix we first derive the relation eq. (2.2) using renormalized mass insertions into the expectation value of a gradient flow observable. The gradient flow makes this derivation transparent, as the on-shell condition for O(a) improvement is never violated. Equation (2.11), which is key to our $b_{\rm g}$-determination, is then easily established as a corollary.

A.1   Derivation of eq. (2.2)

At finite lattice spacing a we consider some gluonic gradient flow observable, $\langle O_g(t) \rangle$, which is renormalized and O(a) improved once expressed in terms of the renormalized and O(a) improved lattice QCD parameters [38]. Introducing the latter through

$$m_{\rm R} = Z_{\rm m}(\tilde{g}_0^2, a\mu)\widetilde{m}_{\rm q}, \qquad g_{\rm R}^2 = Z_{\rm g}(\tilde{g}_0^2, a\mu)\tilde{g}_0^2, \tag{A.1}$$

we can *define* the insertion of the renormalized and O(a) improved flavour singlet scalar density, $S_{\rm R}$, by taking the mass-derivative of the gradient flow observable,

$$a^4 \sum_x \langle O_g(t) S_{\rm R}^0(x) \rangle_{\rm con} \equiv a^4 \sum_x \left\{ \langle O_g(t) S_{\rm R}^0(x) \rangle - \langle O_g(t) \rangle \langle S_{\rm R}^0(x) \rangle \right\} = - \left. \frac{\partial \langle O_g(t) \rangle}{\partial m_{\rm R}} \right|_{g_{\rm R}^2}. \tag{A.2}$$

At fixed lattice spacing (i.e. fixed $\tilde{g}_0^2$) we have,

$$a^4 \sum_x \langle O_g(t) S_{\rm R}^0(x) \rangle_{\rm con} = - Z_{\rm m}^{-1}(\tilde{g}_0^2, a\mu) \frac{\partial \langle O_g(t) \rangle}{\partial \widetilde{m}_{\rm q}}, \tag{A.3}$$

with the improved bare mass insertion given by,

$$-\frac{\partial \langle O_g(t) \rangle}{\partial \widetilde{m}_{\mathrm{q}}} = \left\langle O_g(t) \frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}} \right\rangle_{\mathrm{con}} \equiv \left\langle O_g(t) \frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}} \right\rangle - \left\langle O_g(t) \right\rangle \left\langle \frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}} \right\rangle . \qquad (A.4)$$

Note that the insertion of the scalar density into the gradient flow observable does not give rise to contact terms, despite the Euclidean space-time summation in eq. (A.2). This means that no O(a) counterterms are needed in eq. (A.2) beyond those of the scalar density, eq. (2.1). Furthermore, the cubic divergence cancels in the connected correlation function, so that we can ignore the counterterm $\propto c_{\mathrm{S}}$.

We now use the chain rule to relate the derivative of the action with respect to O(a) improved bare parameters (depending on b_{g} and b_{m}) to the derivatives with respect to the bare parameters. One may think of this as a change of variables,

$$(g_0^2, am_{\mathrm{q}}) \rightarrow (\tilde{g}_0^2, a\widetilde{m}_{\mathrm{q}}) , \qquad (A.5)$$

where the latter are functions of the unimproved bare parameters. As is customary in the physics literature, we do not use different function names. Straightforward application of the chain rule gives,

$$\left. \frac{\partial S}{\partial m_{\mathrm{q}}} \right|_{g_0^2} = \left. \frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}} \right|_{\tilde{g}_0^2} \left. \frac{\partial \widetilde{m}_{\mathrm{q}}}{\partial m_{\mathrm{q}}} \right|_{g_0^2} + \left. \frac{\partial S}{\partial \tilde{g}_0^2} \right|_{\widetilde{m}_{\mathrm{q}}} \left. \frac{\partial \tilde{g}_0^2}{\partial m_{\mathrm{q}}} \right|_{g_0^2} , \qquad (A.6)$$

$$\left. \frac{\partial S}{\partial g_0^2} \right|_{m_{\mathrm{q}}} = \left. \frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}} \right|_{\tilde{g}_0^2} \left. \frac{\partial \widetilde{m}_{\mathrm{q}}}{\partial g_0^2} \right|_{m_{\mathrm{q}}} + \left. \frac{\partial S}{\partial \tilde{g}_0^2} \right|_{\widetilde{m}_{\mathrm{q}}} \left. \frac{\partial \tilde{g}_0^2}{\partial g_0^2} \right|_{m_{\mathrm{q}}} . \qquad (A.7)$$

We thus obtain a 2×2 system of equations,

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} , \qquad (A.8)$$

in the unknowns $X_1 = \partial S / \partial \widetilde{m}_{\mathrm{q}}$ and $X_2 = \partial S / \tilde{g}_0^2$. Working out the derivatives with respect to the unimproved bare parameters we further have,

$$M_{11} = \left. \frac{\partial \widetilde{m}_{\mathrm{q}}}{\partial m_{\mathrm{q}}} \right|_{g_0^2} = 1 + 2b_{\mathrm{m}} am_{\mathrm{q}} , \qquad (A.9)$$

$$M_{21} = \left. \frac{\partial \widetilde{m}_{\mathrm{q}}}{\partial g_0^2} \right|_{m_{\mathrm{q}}} = am_{\mathrm{q}}^2 b_{\mathrm{m}}' , \qquad (A.10)$$

$$M_{12} = \left. \frac{\partial \tilde{g}_0^2}{\partial m_{\mathrm{q}}} \right|_{g_0^2} = ag_0^2 b_{\mathrm{g}} , \qquad (A.11)$$

$$M_{22} = \left. \frac{\partial \tilde{g}_0^2}{\partial g_0^2} \right|_{m_{\mathrm{q}}} = 1 + (b_{\mathrm{g}} + g_0^2 b_{\mathrm{g}}') am_{\mathrm{q}} , \qquad (A.12)$$

where we use the prime notation for derivatives w.r.t. g_0^2, e.g. $b_{\mathrm{g}}' \equiv db_{\mathrm{g}}/dg_0^2$. The derivatives of the lattice action are given by,

$$A_1 = \frac{\partial S}{\partial m_{\mathrm{q}}}\bigg|_{g_0^2} = a^4 \sum_x \overline{\psi}(x)\psi(x)\,, \tag{A.13}$$

$$A_2 = \frac{\partial S}{\partial g_0^2}\bigg|_{m_{\mathrm{q}}} = a^4 \sum_x \overline{\psi}(x)\frac{ia}{4}c_{\mathrm{sw}}'\sigma_{\mu\nu}F_{\mu\nu}\psi(x) - \frac{S_g}{g_0^2}\,. \tag{A.14}$$

Inverting the matrix M,

$$M^{-1} = \frac{1}{\det M}\begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix}\,, \tag{A.15}$$

solving for X_1, and inserting the explicit expressions for the matrix elements, we obtain,

$$\frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}}\bigg|_{\tilde{g}_0^2} = \frac{(1 + (b_{\mathrm{g}} + g_0^2 b_{\mathrm{g}}')am_{\mathrm{q}})\dfrac{\partial S}{\partial m_{\mathrm{q}}}\bigg|_{g_0^2} - ag_0^2 b_{\mathrm{g}}\dfrac{\partial S}{\partial g_0^2}\bigg|_{m_{\mathrm{q}}}}{\left(1 + (b_{\mathrm{g}} + g_0^2 b_{\mathrm{g}}')am_{\mathrm{q}}\right)\left(1 + 2b_{\mathrm{m}}am_{\mathrm{q}}\right) - a^2 m_{\mathrm{q}}^2 g_0^2 b_{\mathrm{g}} b_{\mathrm{m}}'}\,. \tag{A.16}$$

Expanding this expression to first order in am_{q} and neglecting $\mathrm{O}(a^2)$ terms we then get,

$$\frac{\partial S}{\partial \widetilde{m}_{\mathrm{q}}}\bigg|_{\tilde{g}_0^2} = (1 - 2b_{\mathrm{m}}am_{\mathrm{q}})\left(\frac{\partial S}{\partial m_{\mathrm{q}}}\bigg|_{g_0^2} - ag_0^2 b_{\mathrm{g}}\frac{\partial S}{\partial g_0^2}\bigg|_{m_{\mathrm{q}}}\right)\,. \tag{A.17}$$

Upon insertion into the connected correlation function, eq. (A.4), and comparing with eqs. (2.1), (A.3) we thus see that we can identify,

$$d_{\mathrm{S}} = -\frac{b_{\mathrm{g}}}{2g_0^2}\,, \qquad b_{\mathrm{S}^0} = -2b_{\mathrm{m}}\,, \qquad Z_{\mathrm{S}^0} = Z_{\mathrm{m}}^{-1}\,, \tag{A.18}$$

provided that the d_{S}-counterterm is realized on the lattice in the particular form of eq. (2.5). We conclude that it is this lattice discretized form of the $\mathrm{O}(a)$ counterterm that one must use in order to exploit the identity $b_{\mathrm{g}} = -2g_0^2 d_{\mathrm{S}}$.

A.2 Derivation of eq. (2.11)

We now have a closer look at eq. (A.17) in the chiral limit, $m_{\mathrm{q}} = 0$. According to the discussion in section 2, we impose the additional requirement that terms linear in the quark mass are absent. We thus obtain, in terms of derivatives of expectation values,

$$\frac{\partial \langle O_g(t) \rangle}{\partial m_{\mathrm{q}}}\bigg|_{g_0^2, m_{\mathrm{q}}=0} - ag_0^2 b_{\mathrm{g}}\frac{\partial \langle O_g(t) \rangle}{\partial g_0^2}\bigg|_{m_{\mathrm{q}}=0} = 0\,. \tag{A.19}$$

Solving for b_{g} yields eq. (2.11) for gradient flow observables. In order to generalize this result to any $\mathrm{O}(a)$ improved gluonic observable, O_g, we recall that the gradient flow simply removed the complication from contact terms in the insertion of the $\mathrm{O}(a)$ improved scalar density, thereby allowing us to relate b_{g} to d_{S}. However, to derive eq. (2.11), we only need to know that derivatives of $\mathrm{O}(a)$ improved expectation values with respect to the improved bare parameters are again $\mathrm{O}(a)$ improved. Hence the very same steps as carried out above establish eq. (2.11) for arbitrary $\mathrm{O}(a)$ improved gluonic observables.
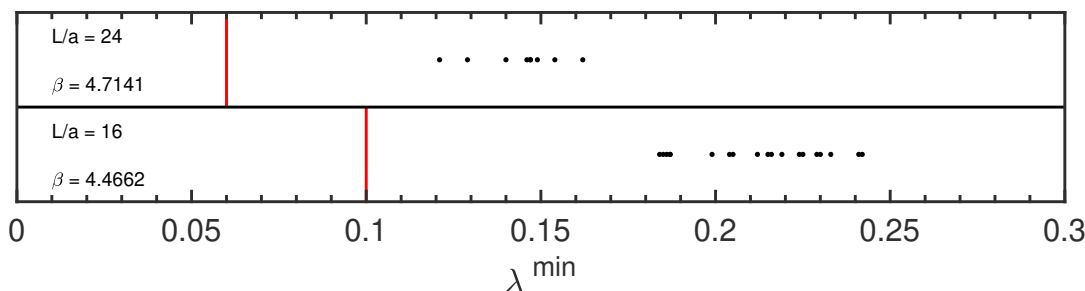
Figure 5. The lowest eigenvalue of $|\gamma_5 \hat{D}|$ on several thermalized configurations at two different lattice spacings. The upper part corresponds to the finer lattice spacing ($L/a = 24, \beta = 4.7141$), while the lower one to the coarser one ($L/a = 16, \beta = 4.4662$). Both are at approximately zero quark mass. The vertical red lines correspond to the lower interval boundaries of the rational approximations that were used in the simulations.

B Simulations

In the following we use the notation and conventions of [54] and the openQCD documentation. The simulation setup is very similar for all runs. The trajectory length is always 2 MDUs. Fermionic forces are integrated with a 4th order Omelyan-Mryglod-Folk (OMF4) [58] integrator with 8 integration steps per trajectory. The gauge force is evaluated five times more frequently, as every interval between fermionic force evaluations is integrated with 1 OMF4 step.

The factorization of the determinant of the doublet into three determinants, as described in section 4.2, has in all cases the masses $a\mu_1 = 0.1$ and $a\mu_2 = 1.5$. The degrees and ranges of the rational approximations for the third quark's determinant vary and can be found in table 4. This table also shows the acceptance rates. For solving linear systems, either a standard conjugate gradient solver, or a variant that solves systems with different mass shifts simultaneously, is used and stopped when a relative residuum norm below 10^{-12} is reached.

More elaborate solvers are not necessary, because the small volume together with our choice of boundary conditions results in quite well conditioned Dirac operators. We monitor the largest and the smallest eigenvalues of $|\gamma_5 \hat{D}|$ in order to decide on the ranges for the rational approximation, and in no case these were dangerously close to zero. Two examples are shown in figure 5.

Data analysis is performed using ADerrors.jl [59, 60] for the gradient flow energy density and pyerrors [61] for the Wilson loops, keeping track of the correct errors and auto-correlations via the Γ-method [62, 63].

At last a remark on the ergodicity of the HMC algorithm. There are two potential difficulties with our setup. First, it is known [63] that with decreasing lattice spacing the correct sampling of topological sectors becomes increasingly difficult. This does not affect our calculation because at the volume corresponding to our LCP the sectors with non-zero charge are suppressed so strongly, that they play virtually no rôle. The second is a difficulty related to the center symmetry. In pure gauge theory, in the large volume limit, the center symmetry is spontaneously broken at high enough temperature. In our small volume the symmetry is restored, but the trace of a Polyakov line exhibits a tri-modal distribution and many algorithms, including the HMC, tend to get stuck in one of the sectors. Although there

is no center symmetry to break in the case of full QCD, similar algorithmic issues have been observed. We investigated the distribution of Polyakov lines and observed that at the volume corresponding to our LCP the distribution is unimodal, and in the perturbative region (e.g. $L/a = 12, \beta = 16$) there is enough movement between the "sectors" to ensure ergodicity.

C Tables of simulation results

In tables 5 and 6 we list the results of our simulations, which are used to obtain the derivatives of $\sigma(c = 0.18)$ and $\hat{\chi}$ with respect to the quark mass and bare coupling.

L/a	β	κ	degree	r_a	r_b	acceptance
12	4.1199	0.136421625496234	8	0.1	6.0	98.8%
12	4.302	0.1350792062069652	8	0.1	6.0	98.9%
12	4.302	0.135997729474131	8	0.1	6.0	99.0%
12	4.302	0.1369288299729244	8	0.1	6.0	99.0%
12	4.4662	0.135602310355609	8	0.1	6.0	98.9%
16	4.302	0.136002904601631	9	0.1	6.0	98.2%
16	4.4662	0.1346938733548286	9	0.1	6.0	98.2%
16	4.4662	0.135607145941904	9	0.1	6.0	98.1%
16	4.4662	0.1365328877033368	9	0.1	6.0	98.0%
16	4.6	0.135286532423878	9	0.1	6.0	98.4%
20	4.4662	0.1356082	10	0.06	6.0	97.1%
20	4.5997	0.1343798945939137	10	0.06	6.0	96.7%
20	4.5997	0.1352889	10	0.06	6.0	98.0%
20	4.5997	0.1362102869948106	10	0.06	6.0	97.7%
20	4.7141	0.1350206	10	0.06	6.0	97.1%
24	4.6	0.135290	10	0.06	6.0	95.8%
24	4.7141	0.1341184114470437	10	0.06	6.0	95.9%
24	4.7141	0.1344791340035503	10	0.06	6.0	96.0%
24	4.7141	0.1350238708	10	0.06	6.0	95.7%
24	4.7141	0.1355730386938167	10	0.06	6.0	95.6%
24	4.7141	0.1359416391158157	10	0.06	6.0	95.5%
24	4.83	0.134758	10	0.06	6.0	96.5%
24	4.9	0.133700810592398	10	0.06	6.0	96.0%
24	4.9	0.134600621200435	10	0.06	6.0	95.9%
24	4.9	0.1355126253782558	10	0.06	6.0	95.9%
24	5.0671	0.1333461402387992	10	0.06	6.0	96.4%
24	5.0671	0.134241167314929	10	0.06	6.0	96.1%
24	5.0671	0.135148290528904	10	0.06	6.0	96.1%
24	5.1719	0.1331332571760407	10	0.06	6.0	96.6%
24	5.1719	0.134025419206206	10	0.06	6.0	96.3%
24	5.1719	0.1349296191446659	10	0.06	6.0	95.9%
24	5.2767	0.133817232093422	10	0.06	6.0	96.1%
24	6.0	0.131701815057449	10	0.06	6.0	96.4%
24	6.0	0.132574832360087	10	0.06	6.0	95.8%
24	6.0	0.1334595009080744	10	0.06	6.0	95.7%
24	6.6	0.131762512060963	10	0.06	6.0	95.8%
24	7.3	0.131001797051939	10	0.06	6.0	94.6%
24	8.0	0.1295464517539626	10	0.06	6.0	94.7%
24	8.0	0.130391036560308	10	0.06	6.0	94.9%
24	8.0	0.1312467062640528	10	0.06	6.0	94.6%
24	8.8	0.129827754035436	10	0.06	6.0	94.0%
24	13.5849	0.127974784827748	10	0.06	6.0	89.0%
24	16.0	0.1266880690255924	10	0.06	6.0	83.6%
24	16.0	0.127495678088902	10	0.06	6.0	83.1%
24	16.0	0.1283136498788704	10	0.06	6.0	83.0%
24	19.4595	0.12702996843205	10	0.06	6.0	75.3%

Table 4. The degrees and ranges of the rational approximations used in the simulations. The last column is the rounded average acceptance rate.

L/a	β	κ	N_{rep}	N_{ms}	$\sigma(0.18)$	$\hat{\chi}$
12	4.1199	0.136422	6	23880	0.076801(11)	0.53993(20)
12	4.3020	0.135998	6	23760	0.0683942(93)	0.48549(17)
12	4.4662	0.135602	6	23880	0.0622623(79)	0.44520(14)
16	4.3020	0.136003	6	24120	0.069545(12)	0.53586(25)
16	4.4662	0.135607	6	42840	0.0628847(77)	0.48830(17)
16	4.6000	0.135287	6	24120	0.0583230(85)	0.45526(20)
20	4.4662	0.135608	6	2292	0.066489(43)	0.53472(95)
20	4.5997	0.135289	6	3360	0.061430(32)	0.49651(79)
20	4.7141	0.135021	6	2244	0.057588(39)	0.46590(80)
24	4.6000	0.135290	10	9600	0.065001(22)	0.53076(54)
24	4.7141	0.135024	10	13800	0.060736(18)	0.49761(42)
24	4.8300	0.134758	10	9600	0.056919(18)	0.46874(49)
24	4.9000	0.134601	10	8200	0.054882(20)	0.45275(52)
24	5.0671	0.134241	10	8200	0.050530(17)	0.41831(46)
24	5.1719	0.134025	10	8440	0.048125(15)	0.39982(39)
24	5.2767	0.133817	10	8800	0.045962(14)	0.38294(39)
24	6.0000	0.132575	10	6600	0.035124(11)	0.29651(29)
24	6.6000	0.131763	10	6600	0.0294367(84)	0.25085(25)
24	7.3000	0.131002	10	9600	0.0247770(52)	0.21282(17)
24	8.0000	0.130391	10	9600	0.0214180(44)	0.18482(13)
24	8.8000	0.129828	10	9600	0.0185570(35)	0.16133(11)
24	13.5849	0.127975	10	7800	0.0103696(21)	0.091990(61)
24	16.000	0.127496	10	16800	0.0084930(11)	0.075731(35)
24	19.4595	0.127030	10	7140	0.0067437(13)	0.060316(43)

Table 5. Results for $\sigma(0.18)$ and $\hat{\chi}$ used to obtain their g_0^2-derivative. All measurements are carried out at vanishing quark mass and originate from N_{rep} independent simulations with a total of N_{ms} measurements.

L/a	β	κ	$N_{\rm rep}$	$N_{\rm ms}$	$\sigma(0.18)$	$\hat{\chi}$
12	4.3020	0.135079	6	23880	0.0690411(94)	0.49040(17)
12	4.3020	0.136929	6	23880	0.0678453(92)	0.48210(17)
16	4.4662	0.134694	6	42840	0.0634640(78)	0.49311(17)
16	4.4662	0.136533	6	42840	0.0624397(76)	0.48552(17)
20	4.5997	0.134380	6	4092	0.061932(29)	0.50028(69)
20	4.5997	0.136210	6	4008	0.060983(27)	0.49194(66)
24	4.7141	0.134118	10	13200	0.061326(19)	0.50417(45)
24	4.7141	0.134479	10	33000	0.061027(12)	0.50067(29)
24	4.7141	0.135573	10	33000	0.060485(12)	0.49659(29)
24	4.7141	0.135942	10	13200	0.060433(20)	0.49701(49)
24	4.9000	0.133701	10	8260	0.055317(21)	0.45685(53)
24	4.9000	0.135513	10	7760	0.054652(19)	0.45191(46)
24	5.0671	0.133346	10	8720	0.050873(17)	0.42154(45)
24	5.0671	0.135148	10	8600	0.050290(17)	0.41663(45)
24	5.1719	0.133133	10	8580	0.048435(16)	0.40215(40)
24	5.1719	0.134930	10	8620	0.047969(15)	0.39823(36)
24	6.0000	0.131702	10	6600	0.035265(11)	0.29760(30)
24	6.0000	0.133460	10	6600	0.035050(10)	0.29624(28)
24	8.0000	0.129546	10	9520	0.0214650(42)	0.18502(13)
24	8.0000	0.131247	10	9600	0.0213978(43)	0.18565(13)
24	16.000	0.126688	10	16800	0.0084963(11)	0.075731(35)
24	16.000	0.128314	10	16800	0.0084918(11)	0.075659(35)

Table 6. Simulation parameters for the $m_{\rm q}$-fits. Results for $\sigma(0.18)$ and $\hat{\chi}$ used to obtain their $m_{\rm q}$-derivative. All measurements are carried out at $am_{\rm q} \approx \pm 0.025$; for $\beta = 4.7141$ also at $am_{\rm q} \approx \pm 0.015$. We carried out $N_{\rm rep}$ independent simulations for a total of $N_{\rm ms}$ measurements. The measurements with $am_{\rm q} \approx 0$ are listed in table 5.

References

[1] ALPHA collaboration, *Determination of $\alpha_s(m_Z)$ by the non-perturbative decoupling method*, *Eur. Phys. J. C* **82** (2022) 1092 [`arXiv:2209.14204`] [INSPIRE].

[2] B. Sheikholeslami and R. Wohlert, *Improved continuum limit lattice action for QCD with Wilson fermions*, *Nucl. Phys. B* **259** (1985) 572 [INSPIRE].

[3] M. Lüscher, S. Sint, R. Sommer and P. Weisz, *Chiral symmetry and $O(a)$ improvement in lattice QCD*, *Nucl. Phys. B* **478** (1996) 365 [`hep-lat/9605038`] [INSPIRE].

[4] S. Sint and R. Sommer, *The running coupling from the QCD Schrödinger functional: a one loop analysis*, *Nucl. Phys. B* **465** (1996) 71 [`hep-lat/9508012`] [INSPIRE].

[5] ALPHA collaboration, *Non-perturbative renormalization by decoupling*, *Phys. Lett. B* **807** (2020) 135571 [`arXiv:1912.06001`] [INSPIRE].

[6] G.M. de Divitiis and R. Petronzio, *Nonperturbative renormalization constants on the lattice from flavor nonsinglet Ward identities*, *Phys. Lett. B* **419** (1998) 311 [`hep-lat/9710071`] [INSPIRE].

[7] ALPHA collaboration, *Nonperturbative results for the coefficients b_m and $b_A - b_P$ in $O(a)$ improved lattice QCD*, *Nucl. Phys. B* **595** (2001) 44 [`hep-lat/0009021`] [INSPIRE].

[8] T. Bhattacharya, R. Gupta, W.-J. Lee and S.R. Sharpe, *Order a improved renormalization constants*, *Phys. Rev. D* **63** (2001) 074505 [`hep-lat/0009038`] [INSPIRE].

[9] T. Bhattacharya, R. Gupta, W. Lee and S.R. Sharpe, *Scaling behavior of discretization errors in renormalization and improvement constants*, *Phys. Rev. D* **73** (2006) 114507 [`hep-lat/0509160`] [INSPIRE].

[10] P. Fritzsch, J. Heitger and N. Tantalo, *Non-perturbative improvement of quark mass renormalization in two-flavour lattice QCD*, *JHEP* **08** (2010) 074 [`arXiv:1004.3978`] [INSPIRE].

[11] P. Korcyl and G.S. Bali, *Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2 + 1$ lattice QCD*, *Phys. Rev. D* **95** (2017) 014505 [`arXiv:1607.07090`] [INSPIRE].

[12] P. Fritzsch, *Mass-improvement of the vector current in three-flavor QCD*, *JHEP* **06** (2018) 015 [`arXiv:1805.07401`] [INSPIRE].

[13] A. Gerardin, T. Harris and H.B. Meyer, *Nonperturbative renormalization and $O(a)$-improvement of the nonsinglet vector current with $N_f = 2 + 1$ Wilson fermions and tree-level Symanzik improved gauge action*, *Phys. Rev. D* **99** (2019) 014519 [`arXiv:1811.08209`] [INSPIRE].

[14] ALPHA collaboration, *Non-perturbative determination of improvement coefficients b_m and $b_A - b_P$ and normalisation factor $Z_m Z_P/Z_A$ with $N_f = 3$ Wilson fermions*, *Eur. Phys. J. C* **79** (2019) 797 [`arXiv:1906.03445`] [INSPIRE].

[15] G. Martinelli et al., *Nonperturbative improvement of composite operators with Wilson fermions*, *Phys. Lett. B* **411** (1997) 141 [`hep-lat/9705018`] [INSPIRE].

[16] J. Bulava and S. Schaefer, *Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action*, *Nucl. Phys. B* **874** (2013) 188 [`arXiv:1304.7093`] [INSPIRE].

[17] M. Lüscher and P. Weisz, *On-shell improved lattice gauge theories*, *Commun. Math. Phys.* **98** (1985) 433 [*Erratum ibid.* **98** (1985) 433] [INSPIRE].

[18] M. Bruno et al., *Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions*, *JHEP* **02** (2015) 043 [arXiv:1411.3982] [INSPIRE].

[19] *Status of CLS configurations for $N_f = 2 + 1$ flavours*, https://www-zeuthen.desy.de/alpha/public-cls-nf21/.

[20] T. Bhattacharya et al., *Improved bilinears in lattice QCD with non-degenerate quarks*, *Phys. Rev. D* **73** (2006) 034504 [hep-lat/0511014] [INSPIRE].

[21] S. Sint, *Renormalization of the iso-singlet scalar density in lattice QCD with Wilson quarks*, *Nucl. Phys. B Proc. Suppl.* **73** (1999) 282 [INSPIRE].

[22] M. Lüscher, R. Narayanan, P. Weisz and U. Wolff, *The Schrödinger functional: a renormalizable probe for non-Abelian gauge theories*, *Nucl. Phys. B* **384** (1992) 168 [hep-lat/9207009] [INSPIRE].

[23] S. Sint, *On the Schrodinger functional in QCD*, *Nucl. Phys. B* **421** (1994) 135 [hep-lat/9312079] [INSPIRE].

[24] P.H. Ginsparg and K.G. Wilson, *A remnant of chiral symmetry on the lattice*, *Phys. Rev. D* **25** (1982) 2649 [INSPIRE].

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

[26] H. Neuberger, *Exactly massless quarks on the lattice*, *Phys. Lett. B* **417** (1998) 141 [hep-lat/9707022] [INSPIRE].

[27] M. Lüscher, *Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation*, *Phys. Lett. B* **428** (1998) 342 [hep-lat/9802011] [INSPIRE].

[28] M. Lüscher, *Topology and the axial anomaly in Abelian lattice gauge theories*, *Nucl. Phys. B* **538** (1999) 515 [hep-lat/9808021] [INSPIRE].

[29] M. Dalla Brida, L. Giusti and M. Pepe, *Non-perturbative definition of the QCD energy-momentum tensor on the lattice*, *JHEP* **04** (2020) 043 [arXiv:2002.06897] [INSPIRE].

[30] ALPHA collaboration, *Lattice QCD with a chirally twisted mass term*, *JHEP* **08** (2001) 058 [hep-lat/0101001] [INSPIRE].

[31] R. Narayanan and H. Neuberger, *Infinite N phase transitions in continuum Wilson loop operators*, *JHEP* **03** (2006) 064 [hep-th/0601210] [INSPIRE].

[32] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, *JHEP* **08** (2010) 071 [*Erratum ibid.* **03** (2014) 092] [arXiv:1006.4518] [INSPIRE].

[33] S. Sint, *The chirally rotated Schrödinger functional with Wilson fermions and automatic O(a) improvement*, *Nucl. Phys. B* **847** (2011) 491 [arXiv:1008.4857] [INSPIRE].

[34] M. Dalla Brida, S. Sint and P. Vilaseca, *The chirally rotated Schrödinger functional: theoretical expectations and perturbative tests*, *JHEP* **08** (2016) 102 [arXiv:1603.00046] [INSPIRE].

[35] M. Dalla Brida, T. Korzec, S. Sint and P. Vilaseca, *High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks*, *Eur. Phys. J. C* **79** (2019) 23 [arXiv:1808.09236] [INSPIRE].

[36] P. Fritzsch, A. Ramos and F. Stollenwerk, *Critical slowing down and the gradient flow coupling in the Schrödinger functional*, *PoS* **Lattice2013** (2014) 461 [arXiv:1311.7304] [INSPIRE].

[37] ALPHA collaboration, *Slow running of the gradient flow coupling from* 200 *MeV to* 4 *GeV in* $N_f = 3$ *QCD*, *Phys. Rev. D* **95** (2017) 014507 [arXiv:1607.06423] [InSPIRE].

[38] M. Lüscher and P. Weisz, *Perturbative analysis of the gradient flow in non-Abelian gauge theories*, *JHEP* **02** (2011) 051 [arXiv:1101.0963] [InSPIRE].

[39] Z. Fodor et al., *The Yang-Mills gradient flow in finite volume*, *JHEP* **11** (2012) 007 [arXiv:1208.1051] [InSPIRE].

[40] A. Ramos and S. Sint, *Symanzik improvement of the gradient flow in lattice gauge theories*, *Eur. Phys. J. C* **76** (2016) 15 [arXiv:1508.05552] [InSPIRE].

[41] M. Creutz, *Monte Carlo study of renormalization in lattice gauge theory*, *Phys. Rev. D* **23** (1981) 1815 [InSPIRE].

[42] R.A. Brandt, F. Neri and M.-A. Sato, *Renormalization of loop functions for all loops*, *Phys. Rev. D* **24** (1981) 879 [InSPIRE].

[43] A. Hasenfratz and F. Knechtli, *Flavor symmetry and the static potential with hypercubic blocking*, *Phys. Rev. D* **64** (2001) 034504 [hep-lat/0103029] [InSPIRE].

[44] ALPHA collaboration, *Static quarks with improved statistical precision*, *Nucl. Phys. B Proc. Suppl.* **129** (2004) 346 [hep-lat/0309080] [InSPIRE].

[45] M. Della Morte, A. Shindler and R. Sommer, *On lattice actions for static quarks*, *JHEP* **08** (2005) 051 [hep-lat/0506008] [InSPIRE].

[46] A. Grimbach, D. Guazzini, F. Knechtli and F. Palombi, $O(a)$ *improvement of the HYP static axial and vector currents at one-loop order of perturbation theory*, *JHEP* **03** (2008) 039 [arXiv:0802.0862] [InSPIRE].

[47] M. Donnellan, F. Knechtli, B. Leder and R. Sommer, *Determination of the static potential with dynamical fermions*, *Nucl. Phys. B* **849** (2011) 45 [arXiv:1012.3037] [InSPIRE].

[48] S. Necco and R. Sommer, *The* $N_f = 0$ *heavy quark potential from short to intermediate distances*, *Nucl. Phys. B* **622** (2002) 328 [hep-lat/0108008] [InSPIRE].

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

[50] M. Hasenbusch, *Speeding up the hybrid Monte Carlo algorithm for dynamical fermions*, *Phys. Lett. B* **519** (2001) 177 [hep-lat/0107019] [InSPIRE].

[51] A.D. Kennedy, I. Horvath and S. Sint, *A new exact method for dynamical fermion computations with nonlocal actions*, *Nucl. Phys. B Proc. Suppl.* **73** (1999) 834 [hep-lat/9809092] [InSPIRE].

[52] M.A. Clark and A.D. Kennedy, *Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields*, *Phys. Rev. Lett.* **98** (2007) 051601 [hep-lat/0608015] [InSPIRE].

[53] J.C. Sexton and D.H. Weingarten, *Hamiltonian evolution for the hybrid Monte Carlo algorithm*, *Nucl. Phys. B* **380** (1992) 665 [InSPIRE].

[54] M. Lüscher and S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, *Comput. Phys. Commun.* **184** (2013) 519 [arXiv:1206.2809] [InSPIRE].

[55] M. Costa, D. Gavriel, H. Panagopoulos and G. Spanoudes, *Mass effects on the QCD* β-*function*, in the proceedings of the 40$^{\text{th}}$ *international symposium on lattice field theory*, (2023) [arXiv:2311.15796] [InSPIRE].

[56] ALPHA collaboration, *Two loop computation of the Schrödinger functional in pure* SU(3) *lattice gauge theory*, *Nucl. Phys. B* **540** (1999) 491 [hep-lat/9809175] [INSPIRE].

[57] ALPHA collaboration, *Two loop computation of the Schrödinger functional in lattice QCD*, *Nucl. Phys. B* **576** (2000) 517 [*Erratum ibid.* **608** (2001) 481] [*Erratum ibid.* **600** (2001) 453] [hep-lat/9911018] [INSPIRE].

[58] I. Omelyan, I. 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.

[59] A. Ramos, *Automatic differentiation for error analysis*, *PoS* **TOOLS2020** (2021) 045 [arXiv:2012.11183] [INSPIRE].

[60] A. Ramos, *Automatic differentiation for error analysis of Monte Carlo data*, *Comput. Phys. Commun.* **238** (2019) 19 [arXiv:1809.01289] [INSPIRE].

[61] F. Joswig, S. Kuberski, J.T. Kuhlmann and J. Neuendorf, *pyerrors: a python framework for error analysis of Monte Carlo data*, *Comput. Phys. Commun.* **288** (2023) 108750 [arXiv:2209.14371] [INSPIRE].

[62] ALPHA collaboration, *Monte Carlo errors with less errors*, *Comput. Phys. Commun.* **156** (2004) 143 [*Erratum ibid.* **176** (2007) 383] [hep-lat/0306017] [INSPIRE].

[63] ALPHA collaboration, *Critical slowing down and error analysis in lattice QCD simulations*, *Nucl. Phys. B* **845** (2011) 93 [arXiv:1009.5228] [INSPIRE].