

2020-10

Master-field simulations of O(a)-improved lattice QCD: Algorithms, stability and exactness

Francis, A

<http://hdl.handle.net/10026.1/15208>

10.1016/j.cpc.2020.107355

Computer Physics Communications

Elsevier BV

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Master-field simulations of $O(a)$ -improved lattice QCD: Algorithms, stability and exactness

Anthony Francis^a, Patrick Fritsch^a, Martin Lüscher^{a,b} and Antonio Rago^c

^a*CERN, Theoretical Physics Department, 1211 Geneva 23, Switzerland*

^b*Albert Einstein Center for Fundamental Physics
Institute for Theoretical Physics, Sidlerstrasse 5, 3012 Bern, Switzerland*

^c*Centre for Mathematical Sciences, University of Plymouth
Plymouth, PL4 8AA, United Kingdom*

In master-field simulations of lattice QCD, the expectation values of interest are obtained from a single or at most a few representative gauge-field configurations on very large lattices. If the light quarks are included, the generation of these fields using standard techniques is however challenging in view of various algorithmic instabilities and precision issues. Ways to overcome these problems are described here for the case of the $O(a)$ -improved Wilson formulation of lattice QCD and the viability of the proposed measures is then checked in extensive simulations of the theory with $2 + 1$ flavours of quarks.

1. Introduction

Numerical lattice QCD usually proceeds by generating a representative ensemble of gauge-field configurations through a Markov process and estimating the expectation values of the chosen observables through ensemble averages. On very large lattices, translation averages in presence of a single gauge field (the master field) provide an alternative way of calculating the expectation values, the associated statistical errors being determined through translation averages too [1].

In the case of the $SU(3)$ gauge theory in four dimensions, some large-scale simulations of this kind were recently performed and worked out as expected [2]. While the inclusion of the light quarks in these calculations is in principle straightforward,

some technical questions must be answered before the established lattice QCD techniques can be applied with confidence. Global accept-reject steps, for example, can be a source of inexactness as a result of significance losses growing proportionally to the lattice size. Other possible issues include numerical instabilities of the simulation algorithm and unbalanced local inaccuracies of approximately calculated solutions of the Dirac equation.

In the present paper, all these potential obstacles for master-field simulations are addressed and solutions are brought forward in each case. For definiteness Wilson's formulation of lattice QCD is considered [3], with $O(a)$ counterterms added as usual [4,5], but the material covered in sects. 3–5 is not specific to this form of the lattice theory. Moreover, an enhanced stability of the simulations is expected to be beneficial for traditional QCD simulations too.

The proposed stabilizing measures include a slight modification of the standard $O(a)$ -improved lattice Dirac operator (sect. 2) and the use of the Stochastic Molecular Dynamics (SMD) simulation algorithm [6,7] in place of the Hybrid Monte Carlo (HMC) algorithm [8] (sect. 3). Which level of numerical precision is required for the algorithm to simulate the theory exactly, and how sufficient precision can be guaranteed, is discussed in sects. 4 and 5. The results of some representative simulations of QCD with $2 + 1$ flavours of quarks, demonstrating the viability of the proposed framework, are then reported in sect. 6.

2. $O(a)$ -improvement revisited

Accidental near-zero modes of the Wilson–Dirac operator are commonly suspected to cause simulation instabilities and enhanced statistical fluctuations. Experience moreover suggests that the addition of the Pauli term required for $O(a)$ -improvement tends to further these undesirable effects, particularly so on coarse lattices. In this section, a modified improved Dirac operator is introduced, which can be expected to be better behaved in this respect and which has some other advantages too.

2.1 Preliminaries

The lattice theory is set up on four-dimensional hypercubic lattices with time extent T , spatial size L^3 , spacing a and periodic boundary conditions in the space directions. Either periodic (anti-periodic for the quark fields), Schrödinger-functional [9,10] or open boundary conditions [11,12] are imposed in the time direction. The gauge

group is taken to be $SU(3)$ and the notational conventions for the group generators, the link variables, etc., are as in refs. [5,11,12].

In the simulations reported in sect. 6, the tree-level Symanzik-improved gauge action is used [13], but the choice of the gauge action is otherwise unimportant. For notational convenience, lattice units are usually employed, where all dimensionful quantities are expressed in units of the lattice spacing a .

2.2 Lattice Dirac operator

On a lattice with periodic boundary conditions, the $O(a)$ -improved Wilson–Dirac operator is given by

$$D = \frac{1}{2} \{ \gamma_\mu (\nabla_\mu^* + \nabla_\mu) - \nabla_\mu^* \nabla_\mu \} + c_{\text{sw}} \frac{i}{4} \sigma_{\mu\nu} \widehat{F}_{\mu\nu} + m_0, \quad (2.1)$$

where m_0 is the bare quark mass, ∇_μ and ∇_μ^* the forward and backward gauge-covariant difference operators and $\widehat{F}_{\mu\nu}$ the standard (“clover”) lattice expression [5] for the gauge-field tensor. The coefficient c_{sw} is equal to 1 at tree-level of perturbation theory and grows monotonically with the gauge coupling, typically reaching values around 2 on coarse lattices (see ref. [14], for example).

If the lattice points are classified as even or odd depending on the parity of the sum of their coordinates, the Dirac operator assumes the block form

$$D = \begin{pmatrix} D_{ee} & D_{eo} \\ D_{oe} & D_{oo} \end{pmatrix}, \quad (2.2)$$

D_{eo} and D_{oe} being the hopping terms from the odd to the even points and the even to the odd points, respectively, while the Pauli term is included in the diagonal part

$$D_{ee} + D_{oo} = M_0 + c_{\text{sw}} \frac{i}{4} \sigma_{\mu\nu} \widehat{F}_{\mu\nu}, \quad M_0 = 4 + m_0. \quad (2.3)$$

A simple way to accelerate lattice QCD simulations exploits this structure by passing to the even-odd preconditioned form $\widehat{D} = D_{ee} - D_{eo} D_{oo}^{-1} D_{oe}$ of the lattice Dirac operator, which acts on quark spinors on the even lattice sites only.

The Pauli term in these equations can be fairly large, particularly so on coarse lattices, where it may get close to saturating the norm bound

$$\left\| \frac{i}{4} \sigma_{\mu\nu} \widehat{F}_{\mu\nu} \right\|_2 \leq 3. \quad (2.4)$$

Since the positive and negative eigenvalues of the Pauli term are equally distributed, and the bare mass m_0 is usually negative, the diagonal part of the Dirac operator

is then not protected from having arbitrarily small eigenvalues. Even-odd preconditioning is in fact known to occasionally fail for this reason, with probability growing proportionally to the lattice size, which practically excludes its use in master-field simulations.

2.3 Alternative form of the lattice Dirac operator

The fact that the diagonal part of the improved Wilson–Dirac operator is not positive distinguishes the improved from the unimproved operator and could explain why the improvement tends to promote the instabilities mentioned at the beginning of this section.

From the point of view of $O(a)$ -improvement and the continuum limit, the alternative expression

$$D_{\text{ee}} + D_{\text{oo}} = M_0 \exp \left\{ \frac{c_{\text{sw}}}{4 + m_0} \frac{i}{4} \sigma_{\mu\nu} \hat{F}_{\mu\nu} \right\} \quad (2.5)$$

for the diagonal part of the Dirac operator may however do just as well. At leading order of perturbation theory, this expression actually coincides with the traditional one and improvement is achieved by setting $c_{\text{sw}} = 1$. Clearly, this form of the diagonal part of the Dirac operator is positive definite and safely invertible. Even-odd preconditioning is therefore guaranteed to be numerically unproblematic. Moreover, $\det D = \det \hat{D}$ up to a field-independent proportionality constant.

Whether the alternative form (2.5) of the diagonal part of the Dirac operator is a viable choice at all couplings in the scaling regime is an open question at this point, which must ultimately be answered through extensive simulations of the modified theory. In these simulations, the exponential of the Pauli term and the derivative of the exponential with respect to the gauge field must be frequently evaluated, but there are ways to do this with negligible computational effort (see appendix A).

2.4 Including $O(a)$ boundary counterterms

If open or Schrödinger-functional boundary conditions are imposed in the time direction, $O(a)$ -improvement near the lattice boundaries requires boundary counterterms to be added to the quark action. The counterterms amount to replacing the constant M_0 in eq. (2.3) by a diagonal operator that acts on quark fields $\psi(x)$ according to [5,11]

$$M_0 \psi(x) = \{4 + m_0 + (c_{\text{F}} - 1)(\delta_{x_0,1} + \delta_{x_0,T-1})\} \psi(x). \quad (2.6)$$

At tree-level of perturbation theory, the improvement coefficient c_F is equal to 1 and the boundary term in eq. (2.6) thus vanishes at this order.

The alternative expression for the diagonal part of the improved Dirac operator is then again given by eq. (2.5) with M_0 set to the operator (2.6). This way of adding the boundary counterterms preserves all the good properties of the term mentioned above.

3. Stochastic molecular dynamics

The numerical integration of the molecular-dynamics equations in the HMC algorithm is known to occasionally go astray, leading to unbounded violations of energy conservation. In these singular cases, the time-reversibility of the integration is likely to be violated too and the exactness of the simulation is then no longer guaranteed. What causes these instabilities remains unclear. In particular, they occur even if the lattice Dirac operator is rigorously protected from having near-zero modes (through a twisted-mass term, for example).

With respect to the HMC algorithm, the Stochastic Molecular Dynamics (SMD) simulation algorithm [6,7] described in this section tends to be less affected by integration instabilities. One of the reasons for this favourable behaviour is the typically much shorter molecular-dynamics integration time, but the fact that the pseudo-fermion fields are more frequently adapted to the changes of the gauge field may have a stabilizing effect too.

3.1 SMD update cycle

The SMD algorithm is rather similar to the HMC algorithm and like the latter does not assume a particular form of the lattice action. For simplicity the algorithm is here described for two-flavour QCD and no frequency-splitting of the quark determinant, the generalization to other cases of interest being straightforward.

The fields processed by the SMD algorithm are the gauge field $U(x, \mu)$, the associated momentum field $\pi(x, \mu)$ and a pseudo-fermion field $\phi(x)$ with action

$$S_{\text{pf}}(U, \phi) = (\phi, (D^\dagger D)^{-1} \phi). \quad (3.1)$$

In the case considered, the only other term included in the total action $S(U, \phi)$ of the theory is the gauge action $S_G(U)$. An SMD update cycle then consists of a random rotation of the momentum and the pseudo-fermion field, a short molecular-dynamics

evolution of the momentum and the gauge field and, finally, an accept-reject step that makes the algorithm exact.

3.2 Random field rotation

At the beginning of the update cycle, the momentum and the pseudo-fermion field are refreshed according to

$$\pi \rightarrow c_1\pi + c_2v, \quad c_1 = e^{-\gamma\epsilon}, \quad c_2 = (1 - c_1^2)^{1/2}, \quad (3.2)$$

$$\phi \rightarrow c_1\phi + c_2D^\dagger\eta, \quad (3.3)$$

where the fields $v(x, \mu)$ and $\eta(x)$ are chosen randomly with normal distribution. Both $\gamma > 0$ and $\epsilon > 0$ are fixed parameters of the SMD algorithm, ϵ being equal to the molecular-dynamics integration time (see subsect. 3.3).

The algorithm defined by eqs. (3.2),(3.3) simulates the Gaussian distribution

$$\text{constant} \times \exp\{-\frac{1}{2}(\pi, \pi) - S_{\text{pf}}(U, \phi)\} \quad (3.4)$$

at fixed gauge field. If there are further pseudo-fermion fields, this property must be maintained by including them in the update step and by adapting eq. (3.3) to the chosen pseudo-fermion actions. The “friction parameter” γ determines how quickly the memory of previous field configurations is lost. In principle each field may have its own friction parameter, but for simplicity γ is here taken to be same for all fields.

3.3 Molecular-dynamics evolution and accept-reject step

In the second step of the SMD update cycle, the molecular-dynamics equations for the momentum and the gauge field, which derive from the Hamilton function

$$H(\pi, U) = \frac{1}{2}(\pi, \pi) + S(U, \phi), \quad (3.5)$$

are integrated from the current simulation time t to $t+\epsilon$ using a reversible symplectic integration rule. The fields $\tilde{\pi}$ and \tilde{U} obtained at the end of the integration are then accepted with probability

$$P_{\text{acc}}(\pi, U) = \min\{1, e^{-\Delta H(\pi, U)}\}, \quad \Delta H(\pi, U) = H(\tilde{\pi}, \tilde{U}) - H(\pi, U). \quad (3.6)$$

Otherwise, i.e. if the proposed fields $\tilde{\pi}, \tilde{U}$ are not accepted, the gauge field is set to U and the momentum field to $-\pi$ [7]. Throughout this step, the pseudo-fermion field is held fixed.

The accept-reject step guarantees that the update cycle preserves the distribution e^{-H} of the gauge field, the momentum field and the pseudo-fermion field. Moreover, if ϵ is sufficiently small, the SMD algorithm can be rigorously shown to be ergodic and to asymptotically simulate this distribution, independently of the initial values of the fields [15].

3.4 Continuous simulation-time limit

At fixed ϵ and large γ , the SMD algorithm coincides with the HMC algorithm. On the other hand, if ϵ is taken to zero while γ is held fixed, the acceptance probability (3.5) goes to 1 and the algorithm solves the stochastic molecular-dynamics equations

$$\partial_t \pi = -\gamma \pi - \partial_U S + v, \quad (3.7)$$

$$\partial_t U = \pi U, \quad (3.8)$$

$$\partial_t \phi = -\gamma \phi + D^\dagger \eta. \quad (3.9)$$

A compact notation is used here and the white-noise fields v and η have been scaled by the factor $\sqrt{\epsilon/2\gamma}$ before going to the limit $\epsilon = 0$.

The parameter γ can in principle be set to any positive value. Previous empirical studies of the SU(3) gauge theory however showed that the autocorrelation times of physical quantities have a shallow minimum around $\gamma = 0.3$ [11]. At this value of γ , the memory of old momentum and pseudo-fermion fields lasts for a few molecular-dynamics time units, i.e. for about as long as the trajectory lengths typically chosen in HMC simulations of lattice QCD.

3.5 Choice of parameters and efficiency considerations

SMD step size. Short molecular-dynamics integration times ϵ are beneficial for the stability of the algorithm and do not have a negative impact on the autocorrelation times if γ is held fixed. In practice a few steps of a reversible symplectic integration rule are applied to integrate the fields from time t to $t + \epsilon$ and the acceptance rate $\langle P_{\text{acc}} \rangle$ is set to the desired value by adjusting ϵ .

Acceptance rate. When field configurations are rejected, the momentum is reversed and the algorithm tends to backtrack its trajectory in field space. An efficient sampling thus requires the acceptance rate to be such that rejections only occur at large separations in simulation time, relevant reference time distances being $1/\gamma$ and the autocorrelation times of physical quantities [11].

Large-volume scaling. If all other parameters are held fixed, the energy difference ΔH entering the acceptance probability (3.6) grows roughly like $V^{1/2}$ with the number V of lattice points. In order to preserve a high acceptance rate on large lattices, the integration of the molecular-dynamics equations must thus become more and more accurate. The computational cost per unit of simulation time then increases proportionally to $V^{1/2p}$ if an integration rule of order p is used. Higher-order schemes such as the ones listed in ref. [16] are therefore highly recommended for master-field simulations.

Exploiting continuity in simulation time. Since the SMD algorithm updates the fields in small steps, the computational effort can be significantly reduced by propagating previous solutions of the Dirac equation [17] and a local deflation subspace [18] along the trajectory in field space. Complete regenerations of the deflation subspace are then not even required and regular incremental updates suffice to keep it in good condition at all times [18].

Simulation efficiency. The SMD algorithm tends to consume more computer time per unit of molecular-dynamics time than the HMC algorithm, but the efficiencies of the two algorithms end up being similar once the autocorrelation times are taken into account [11].

4. Required numerical precision of ΔH

Since the Hamilton function (3.5) is an extensive quantity, whose average value grows proportionally to the number of lattice points, important significance losses occur when the energy deficits ΔH are computed at the end of the molecular-dynamics evolution of the gauge and the momentum field. Depending on the lattice size and on how accurately the Hamilton function is obtained, the numerical errors of the calculated values of ΔH may then be such that the correctness of the simulations is compromised by false accept-reject decisions. The goal in this section is to determine at which level the errors can have a statistically noticeable effect.

4.1 Model distribution of ΔH

On large lattices, and if the numerical integration of the molecular-dynamics equations is stable, the probability distribution $\rho(h)$ of $h = \Delta H$ is expected to be well represented by a Gaussian, because ΔH is a sum of uncorrelated contributions from

distant regions on the lattice. The distribution must satisfy

$$\int dh \rho(h) = \int dh e^{-h} \rho(h) = 1 \quad (4.1)$$

and is therefore of the form

$$\rho(h) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(h - \frac{1}{2}\sigma^2)^2}{2\sigma^2}\right\}, \quad (4.2)$$

where the width σ is given by the acceptance rate

$$\langle P_{\text{acc}} \rangle = \int dh \rho(h) \min\{1, e^{-h}\} = 1 - \frac{\sigma}{\sqrt{2\pi}} + \frac{\sigma^3}{24\sqrt{2\pi}} + O(\sigma^5). \quad (4.3)$$

For $\langle P_{\text{acc}} \rangle = 0.80, 0.90$ and 0.95 , for example, σ is equal to $0.51, 0.25$ and 0.125 . The model distribution (4.2) is in fact accurately matched by the empirical distributions measured in simulations like run B₁ discussed in sect. 6, in which the integration of the molecular-dynamics equations never went astray.

The calculated approximate value \tilde{h} of ΔH is distributed slightly differently from the exact value. Assuming the numerical errors are small, randomly distributed and uncorrelated to h , the joint distribution

$$\rho_2(h, \tilde{h}) = \frac{\rho(h)}{\sqrt{2\pi}\delta} \exp\left\{-\frac{(\tilde{h} - h)^2}{2\delta^2}\right\} \quad (4.4)$$

may be expected to describe the situation reasonably well, δ being the mean square deviation of \tilde{h} from h . Integration over h then yields the distribution

$$\tilde{\rho}(\tilde{h}) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}} \exp\left\{-\frac{(\tilde{h} - \frac{1}{2}\sigma^2)^2}{2\tilde{\sigma}^2}\right\}, \quad \tilde{\sigma} = (\sigma^2 + \delta^2)^{1/2}, \quad (4.5)$$

of the actually measured values \tilde{h} of ΔH .

4.2 Kolmogorov–Smirnov test

One may now ask how well the distribution $\tilde{\rho}$ must be sampled to be able to distinguish it from the exact distribution ρ . If a sample size of, say, N trials is insufficient, any sequence of N accept-reject decisions made in an actual simulation is statistically as likely as the one in a (theoretical) simulation without numerical errors.

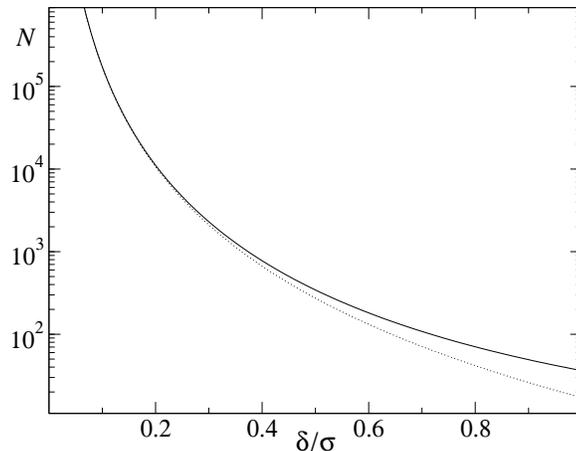


Fig. 1. Number N of accept-reject steps where the difference of the distributions of the calculated and the exact values of ΔH starts to be statistically noticeable. In the limit $\delta \ll \sigma$ of small numerical errors, N is asymptotically equal to $17.1 \times (\sigma/\delta)^4$ (dotted line).

The Kolmogorov–Smirnov test provides an answer to this question by comparing the exact cumulative distribution

$$F(h) = \int_{-\infty}^h dz \rho(z) \quad (4.6)$$

with the empirical distribution

$$F_N(h) = \{\text{number of cases } k \text{ where } h_k \leq h\}/N \quad (4.7)$$

obtained by drawing N numbers h_1, \dots, h_N randomly with distribution $\rho(h)$ (see ref. [19], for example). In particular, the deviation

$$d = \sup_h |\tilde{F}(h) - F(h)| \quad (4.8)$$

of the modified cumulative distribution \tilde{F} from the exact one can only be resolved, if d is larger than the standard deviation of F_N from F . Since the latter is at most $(4N)^{-1/2}$ at all values of h [19], sample sizes N satisfying

$$N \geq \frac{1}{4d^2} \quad (4.9)$$

are required to detect the deviation (4.8) of the cumulative distributions.

Assuming the model distributions (4.2),(4.5), some algebra shows that

$$d = \frac{1}{2}\{\text{erf}(\tilde{\sigma}r) - \text{erf}(\sigma r)\}, \quad r = \frac{1}{\delta}\sqrt{\ln(\tilde{\sigma}/\sigma)}, \quad (4.10)$$

is a steep function of the ratio δ/σ . The bound (4.9) then implies that huge statistics is usually required before the numerical errors of ΔH can affect the simulations at a statistically significant level (see fig. 1). An SMD simulation with average acceptance rate of 99% and error margin $\delta = 10^{-3}$, for example, is potentially affected only after about 7 million update cycles. The precision requirements on ΔH are thus fairly mild, a fact that was recently confirmed in an empirical study of the SU(2) gauge theory using the HMC algorithm [20].

5. Sources of numerical inaccuracies

Lattice QCD simulations are usually performed on machines complying with the IEEE 754 standard for floating-point data and arithmetic. It is also common to use double-precision (64 bit) data and operations for the basic fields, except perhaps in some intermediate steps, when the accuracy of the final results is provably not affected. Various sources of numerical inaccuracies however require special attention if large lattices are simulated.

5.1 Lattice sums

Scalar products of quark fields and the Hamilton function, for example, involve a sum over all lattice points. Important accumulations of rounding errors in these typically huge sums can be safely avoided using quadruple-precision (128 bit) floating-point arithmetic. Quadruple-precision numbers may conveniently be represented by pairs of standard IEEE 754 double-precision numbers and there exist efficient algorithms that correctly implement the associated arithmetic operations on any machine complying with the standard [21,22].

Clearly, when the energy difference ΔH is computed, the truncation to 64 bit precision should occur only after calculating the difference, so that ΔH is obtained with an absolute numerical error equal to the (now practically exact) sum of the errors of the local contributions to the Hamilton function.

5.2 Spatially non-uniform inaccuracies

In the course of the molecular-dynamics evolution of the gauge and the momentum field, the Dirac equation

$$D\psi(x) = \eta(x) \tag{5.1}$$

must be solved many times for given source fields η . The solution is obtained using some iterative algorithm, the iteration being stopped when the current approximate solution $\tilde{\psi}$ satisfies

$$\|\eta - D\tilde{\psi}\| \leq \omega\|\eta\| \tag{5.2}$$

for some norm $\|\cdot\|$ and specified tolerance ω . In practice the rounding errors involved in the process set a lower limit (about 10^{-14} in the case of 64 bit data and arithmetic) on the tolerances that can be attained by the algorithms.

Traditionally the square norm $\|\cdot\|_2$ is used and the stopping criterion thus requires the sum

$$\|\varrho\|_2^2 = \sum_x \|\varrho(x)\|_2^2, \quad \varrho(x) = \eta(x) - D\tilde{\psi}(x), \tag{5.3}$$

to be small. Since the contributions $\|\eta(x)\|_2$ to the norm of the source fields in lattice QCD simulations are all about equally distributed, the square norm $\|\eta\|_2^2$ is typically a huge number on the order of the number V of lattice points. The bound (5.2) then does not exclude large local imbalances of the accuracy of the approximate solution $\tilde{\psi}(x)$, where, in the worst case, $\|\varrho(x)\|_2^2 = O(V)$ at some points x .

Inaccuracies of the solutions of the Dirac equation propagate to the force terms in the molecular-dynamics equations and thus to the evolved momentum and gauge fields. Whether large local inaccuracies do occur with some appreciable probability is currently not known, nor are their effects on the fields and the energy deficits ΔH . Their possible presence however implies a loss of control over the correctness of the simulations, particularly so if very large lattices are simulated.

Unbalanced inaccuracies are excluded if the uniform norm

$$\|\psi\|_\infty = \sup_x \|\psi(x)\|_2 \tag{5.4}$$

is used in the stopping criterion (5.2). The uniform norm has all properties a norm in a complex linear space must have. Moreover, ample empirical evidence suggests that the iterative algorithms commonly used in lattice QCD are able to deliver accurate

solutions of the Dirac equation satisfying the uniform-norm stopping criterion (see appendix B for further details about the uniform norm and the associated operator norm).

5.3 Precision loss at small quark masses

The deviation of the exact solution ψ of the Dirac equation (5.1) from any approximate solution $\tilde{\psi}$ satisfying the stopping criterion (5.2) is bounded by [23]

$$\|\tilde{\psi} - \psi\| \leq \omega \kappa(D) \|\psi\|, \quad \kappa(D) = \|D\| \|D^{-1}\|. \quad (5.5)$$

At small quark masses m and lattice spacings a , the condition number $\kappa(D)$ of the Dirac operator diverges approximately like $(am)^{-1}$ and reaches values of 10^3 or even 10^4 in practice. Since the computation of the forces in the molecular-dynamics step requires the normal Dirac equation, $D^\dagger D \psi = \eta$, to be solved, the errors propagated to the momentum and the gauge field tend to be enhanced by a factor proportional to $(am)^{-2}$. At the end of the molecular-dynamics evolution, the energy deficit ΔH is then obtained with an absolute numerical error growing like $V^{1/2}(am)^{-2}$ at small am and large V .

A reduction of these potentially catastrophic errors is automatically achieved if the Dirac operator is split in several factors and a corresponding number of pseudo-fermion fields is used to represent the determinants of the factors [24–29]. The effect is particularly transparent in the case of the twisted-mass factorization [24,25]

$$D^\dagger D = (D^\dagger D + \mu_n^2) \prod_{k=0}^{n-1} \frac{D^\dagger D + \mu_k^2}{D^\dagger D + \mu_{k+1}^2}, \quad \mu_n > \mu_{n-1} > \dots > \mu_0 = 0. \quad (5.6)$$

If the largest mass, μ_n , is set to a number of order 1 in lattice units, the computation of the force deriving from the first factor is numerically unproblematic. The actions associated with the rational factors, on the other hand, may be written in the form

$$S_{\text{pf},k}(U, \phi_k) = (\phi_k, \phi_k) + (\mu_{k+1}^2 - \mu_k^2) (\phi_k, (D^\dagger D + \mu_k^2)^{-1} \phi_k), \quad (5.7)$$

which shows that the part involving the inverse of the Dirac operator is suppressed by an explicit twisted-mass factor. In particular, if a log-scale factorization is chosen [12], where μ_1 is on the order of the quark mass and $\mu_{k+1} \simeq 10 \times \mu_k$, the twisted-mass factor provides a suppression equal to μ_{k+1}^2 of the force deriving from eq. (5.7) and its absolute numerical error.

5.4 Synthesis

If all lattice sums are performed with quadruple precision, the approximate solution of the Dirac equation is the dominant source of numerical errors in the simulations. The calculated values of the energy deficit ΔH are directly affected by these inaccuracies, but also indirectly through the inaccuracies of the fields accumulated in the field-rotation and molecular-dynamics steps.

Provided uniform-norm stopping criteria are used in these steps, and if a log-scale factorization of the light-quark determinant is chosen, the total error of ΔH is expected to be reliably controlled by the solver tolerances. The associated sensitivities and tolerances can then be determined empirically so as to meet the precision requirement discussed in sect. 4. Since the error of ΔH grows with the number V of lattice points (roughly like $V^{1/2}$), the tolerances must be tightened with increasing lattice size, but there is still ample room for lattices larger than the ones simulated to date before the tolerances reach the limit set by the machine precision.

6. Numerical studies

The simulations reported in this section mainly serve to check whether the modified $O(a)$ -improved Wilson–Dirac operator introduced in sect. 2 is an attractive choice of Dirac operator for numerical lattice QCD. This is also the first time the SMD algorithm (in the form described here with continuous updates of the pseudo-fermion fields) is used to simulate QCD with light quarks. For these rather technical studies, where a direct comparison with previous work using the standard setup is desirable, the traditional rather than the master-field simulation strategy is chosen.

6.1 Non-perturbative $O(a)$ -improvement

As already mentioned, the theory with tree-level improved gauge action [13] and $2+1$ flavours of quarks, referred to as the up, down and strange quark, is considered. From now on the alternative form of the improved Wilson–Dirac operator will be assumed unless stated otherwise. The complete specification of the theory then still requires the parameter c_{sw} in eq. (2.5) to be determined as a function of the bare gauge coupling g_0 .

The results of a non-perturbative computation of the parameter along the lines of ref. [30] are plotted in fig. 2. No particular difficulty was met in this calculation, where c_{sw} is determined by imposing $O(a)$ -improvement in physically small volumes.

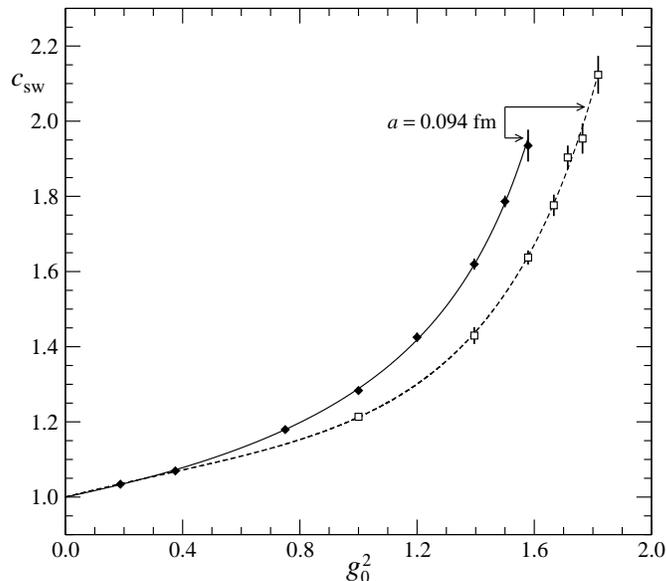


Fig. 2. Non-perturbatively determined values (black diamonds) of the parameter c_{sw} of the modified lattice Dirac operator [cf. eq. (2.5)]. The solid curve represents the rational fit (6.1) of the data. Also shown are the values (open squares) and fit curve (dashed line) in the standard $O(a)$ -improved theory obtained in ref. [14].

In the range $6/g_0^2 \geq 3.8$ of the coupling, the rational function

$$c_{\text{sw}} = \frac{1 - 0.325022 g_0^2 - 0.0167274 g_0^4}{1 - 0.489157 g_0^2} \quad (6.1)$$

provides an excellent fit of the data.

The computation of c_{sw} was pushed up to values of the coupling, where the lattice spacing a in physical units gets close to 0.1 fm and higher-order lattice effects become non-negligible (scale setting is discussed below). Apart from a shift of the coupling to smaller values for a given lattice spacing, the dependence of c_{sw} on the coupling looks similar to the one previously found in the case of the standard $O(a)$ -improved Wilson–Dirac operator.

6.2 Physical observables

Following refs. [31,32], the renormalized parameters of the lattice theory are taken to be the reference gradient-flow time t_0 [33], the pion mass M_π and the kaon mass M_K . The continuum limit, for example, is approached at fixed

$$\phi_2 = 8t_0 M_\pi^2 \quad \text{and} \quad \phi_4 = 8t_0 \left(M_K^2 + \frac{1}{2} M_\pi^2 \right), \quad (6.2)$$

while t_0/a^2 goes to infinity.

The use of the reference flow time is suggested for various technical reasons, including the fact that t_0 is nearly independent of the bare light-quark mass $m_{0,u} = m_{0,d}$ when the gauge coupling and the sum $m_{0,u} + m_{0,d} + m_{0,s}$ of the quark masses are held fixed. Along these curves in parameter space (referred to as “chiral trajectories”), ϕ_4 is nearly constant too and only ϕ_2 varies approximately linearly with the light-quark mass. In particular, $\phi_4 \simeq 1.12$ on the trajectories passing through the physical point [31,32]. For the conversion of lattice to physical units, the physical value of the gradient-flow smoothing range at the reference flow time, $\sqrt{8t_0} = 0.415(4)(2)$ fm, determined by Bruno et al. [32] will be used.

Further observables considered are the pion and kaon decay constants[†]

$$F_\pi = \hat{Z}_A^{ud} F_{0,\pi}, \quad F_K = \hat{Z}_A^{us} F_{0,K}, \quad (6.3)$$

where the renormalization factors \hat{Z}_A^{ud} and \hat{Z}_A^{us} are computed on the same lattices as the bare decay constants, $F_{0,\pi}$ and $F_{0,K}$, by probing the PCAC relation at positive gradient-flow time [34]. Since the calculation includes the mass corrections required for $O(a)$ -improvement [5,35], the values of the renormalization constants are slightly dependent on the flavour channel. The bare decay constants and the renormalization constants also depend on the axial-current improvement coefficient c_A [5], but, as explained in appendix C, the renormalized decay constants (6.3) are insensitive to the value of c_A .

6.3 Simulation algorithm

Apart from the use of the SMD in place of the HMC algorithm, the simulations reported in this paper were set up as described in ref. [12]. In particular, twisted-mass factorizations were chosen for the light-quark determinant, rational approximations for the strange-quark determinant and a hierarchical 4th-order integrator with 2

[†] The normalization convention for the decay constants is the one often used in chiral perturbation theory, i.e. the one where the physical value of the pion decay constant is about 93 MeV.

Table 1. Lattice and run parameters

Run	Lattice	β	κ_u	κ_s	c_{sw}	D	$\langle P_{acc} \rangle$	N_{cycles}
A ₁	96×32^3	3.8	0.1389630	0.1389630	1.955242	m	0.975	10000
A ₂	96×32^3	3.8	0.1391874	0.1385164	1.955242	m	0.986	10000
A ₃	96×32^3	3.8	0.1392888	0.1383160	1.955242	m	0.982	10000
B ₁	96×48^3	4.0	0.1382720	0.1382720	1.783303	m	0.988	24000
X ₁	96×32^3	3.36	0.1366400	0.1366400	2.038765	t	0.996	10000

Table 2. Lattice spacing, lattice size, masses and decay constants in physical units

Run	a [fm]	L [fm]	M_π [MeV]	M_K [MeV]	F_π [MeV]	F_K [MeV]	$M_\pi L$
A ₁	0.094	3.0	408	—	108	—	6.2
A ₂	0.094	3.0	293	454	100	109	4.5
A ₃	0.093	3.0	215	470	96	109	3.3
B ₁	0.064	3.1	409	—	107	—	6.4
X ₁	0.095	3.0	408	—	107	—	6.3

levels for the numerical integration of the molecular-dynamics equations. Following the recommendations in sect. 5, the uniform-norm stopping criterion was used in the computations of the forces deriving from the pseudo-fermion actions, with tolerances far sufficient for the numerical errors of ΔH to be negligible.

All simulations were performed using the publicly available `openQCD` program package [36]. The SMD friction parameter γ was set to 0.3, as suggested in ref. [11], and the SMD step size $\epsilon = 0.31$ was chosen so that a high acceptance rate is attained with 2 steps at the outer level of the molecular-dynamics integration scheme (3 steps in the case of run X₁).

6.4 Lattice parameters

The basic parameters of the simulated lattices are listed in table 1, as usual quoting the values of $\beta = 6/g_0^2$ and the hopping parameters $\kappa = 1/(8 + 2m_0)$ instead of the bare coupling and masses. Periodic boundary conditions (anti-periodic in time for the quark fields) were imposed in all cases and, as indicated by the entries in the column labeled “ D ”, run X₁ is the only one where the traditional form of the Dirac operator was used. In the last two columns, the acceptance rates and numbers of SMD update cycles performed after thermalization are listed.

Table 3. Masses and decay constants in units of t_0

Run	t_0/a^2	ϕ_2	ϕ_4	$\sqrt{8t_0}F_\pi$	$\sqrt{8t_0}F_K$
A ₁	2.4420(36)	0.7366(39)	1.1049(45)	0.22701(60)	–
A ₂	2.4540(27)	0.3807(40)	1.1056(51)	0.21095(72)	0.23018(50)
A ₃	2.4645(35)	0.2039(56)	1.0873(60)	0.20119(86)	0.22912(56)
B ₁	5.2470(57)	0.7414(39)	1.1121(44)	0.22430(66)	–
X ₁	2.3952(34)	0.7361(76)	1.1042(85)	0.22452(55)	–

The simulated lattices are in the large-volume regime of QCD and approximately on the physical chiral trajectories $\phi_4 \simeq 1.12$ [31,32] (see tables 2 and 3). In the case of the lattices A₁–A₃, only the bare quark masses change, keeping their sum exactly constant, while A₁, B₁ and X₁ are all at the SU(3)-symmetric point. In particular, the kinematical situation on the A₁ and X₁ lattices is practically the same. A₁ and B₁, on the other hand, allow the behaviour of the simulations and the lattice effects to be studied as the lattice spacing is reduced at fixed meson masses (i.e. fixed ϕ_2 and ϕ_4). The figures listed in table 2 serve for illustration only and are therefore given without errors.

6.5 Simulation stability

As previously noted, large energy violations ΔH in the molecular-dynamics step of the SMD algorithm are a sign of potentially harmful algorithmic instabilities. On the coarser lattices simulated, A₁–A₃ and X₁, the fractions R_{spk} of update cycles where $|\Delta H| > 1$ (i.e. where ΔH has a “spike”) are one or two per mille and thus fairly small (see table 4). When the lattice spacing is reduced, the stability of the molecular-dynamics integration improves and not a single spike was seen in run B₁.

While the frequency of spikes is similar in run A₁ and X₁, the molecular-dynamics equations had to be integrated with a 1.5 times smaller step size in run X₁ to achieve this. The use of the traditional instead of the alternative form of the Dirac operator also leads to larger fluctuations and a much longer integrated autocorrelation time $\tau(\bar{w}_p)$ of the average plaquette \bar{w}_p (see fig. 3 and table 4). Since the distribution of the latter is a property of the theory, these large fluctuations cannot be attributed to the choice of simulation algorithm and may instead signal the onset of a change of regime at this coarse lattice spacing. The observed autocorrelations of the average plaquette however do suggest the same, given that this quantity is dominated by the high-frequency modes of the gauge field, which are normally rapidly equilibrated by the SMD algorithm. On finer lattices the effect in fact quickly disappears.

Table 4. Observables related to the simulation stability

Run	$R_{\text{spk}} [\%]$	$\tau(\bar{w}_p)$	$\bar{\mu} \times 10^3$	$\sigma \times 10^3$	$\sigma\sqrt{V}$	$\frac{1}{2}\hat{Z}_A^{ud}m_{ud} \times 10^3$
A ₁	0.19(10)	7.6(1.6)	5.251(89)	0.613(57)	1.09(10)	8.508(34)
A ₂	0.19(10)	5.4(1.1)	2.283(51)	0.399(32)	0.708(57)	4.260(32)
A ₃	0.10(7)	7.4(1.4)	1.218(46)	0.351(33)	0.623(59)	2.260(50)
B ₁	0.0	3.5(4)	4.541(13)	0.310(8)	1.011(27)	5.901(18)
X ₁	0.22(5)	> 30	4.071(49)	0.843(30)	1.495(53)	7.760(60)

$\tau(\bar{w}_p)$ is given in units of molecular-dynamics time and the values in columns 4 – 7 in lattice units.

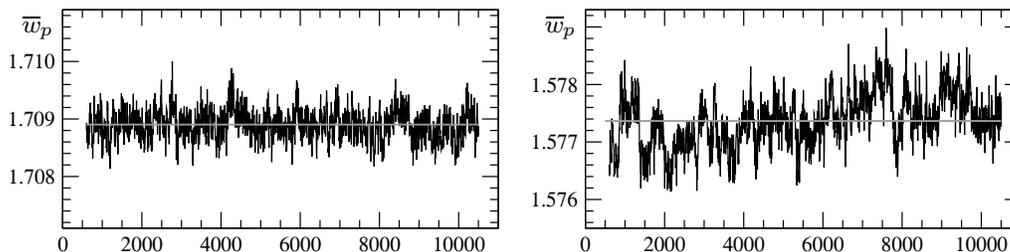


Fig. 3. Histories of the average Wilson plaquette loop \bar{w}_p versus the SMD update cycle number. The values plotted are the ones measured after every 5th cycle in run A₁ (left) and X₁ (right). In both plots the same scale is used on the ordinate.

6.6 Spectral gap

The correctness and efficiency of the simulations may also be negatively affected if the lattice Dirac operator does not have a safe spectral gap [37]. Compared to the ones previously obtained in two-flavour QCD [37,38], the distributions of the lowest eigenvalue of the light-quark operator $(D^\dagger D)^{1/2}$ measured in the runs A₁–X₁ turn out to be rather similar (see table 4 and appendix D). In particular, the distributions are separated from the origin on all these lattices. Their medians $\bar{\mu}$ (fourth column in table 4) are lower than the product of the axial-current renormalization constant times the bare current-quark mass $\frac{1}{2}m_{ud}$ of the light quarks (last column), but the difference decreases toward the continuum limit as has to be the case [37].

Theoretical arguments and numerical evidence produced in refs. [37,38] suggest that the widths σ of the eigenvalue distributions are, in lattice units, roughly equal to $1/\sqrt{V}$ (where V denotes the number of lattice points). The values of $\sigma\sqrt{V}$ listed in table 4 are broadly consistent with this rule and moreover show that the product tends to decrease with $M_\pi L$, an observation previously made in the standard $O(a)$ -

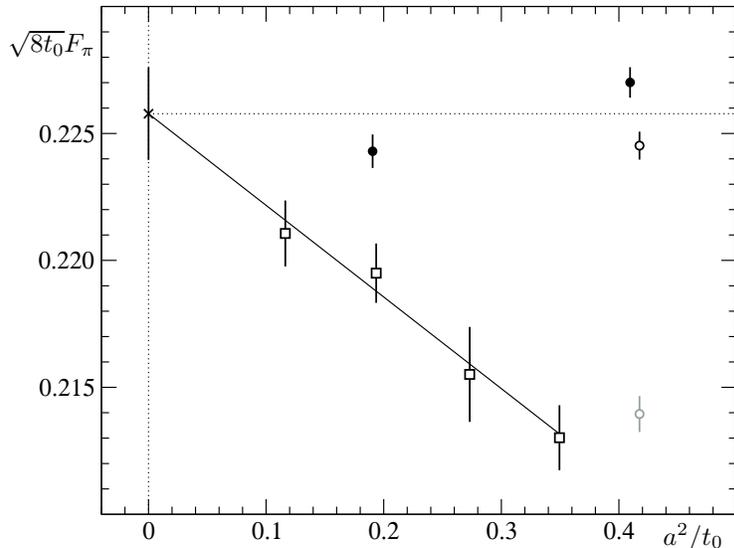


Fig. 4. Lattice-spacing dependence of $\sqrt{8t_0}F_\pi$ at $m_{0,u} = m_{0,d} = m_{0,s}$. Open squares represent results previously obtained in ref. [32] at $\phi_4 = 1.11$ using the traditional setup of the $O(a)$ -improved theory. A linear extrapolation of these data yields a value in the continuum limit (cross), which coincides with the results obtained in the runs A_1 , B_1 (black circles) and X_1 (open circle). The latter moves down if the axial-current renormalization constant is replaced by the one used in ref. [32] (grey open circle).

improved two-flavour theory [38]. With respect to the distribution measured in run A_1 , the one obtained in run X_1 is however noticeably wider, i.e. at this fairly coarse lattice spacing, the use of the modified Dirac operator leads to a narrower eigenvalue distribution.

6.7 Higher-order lattice effects

All simulated lattices are in a range of parameters, where $O(a^2)$ lattice effects cannot be expected to be very small, independently of which Dirac operator is chosen. The values of t_0/a^2 , for example, would change by 15% on the A_1 lattice and by 6% on the B_1 lattice, if defined with the Wilson plaquette instead of the symmetric (“clover”) expression for the Yang–Mills action density.

The dimensionless combination $\sqrt{8t_0}F_\pi$ is potentially more sensitive to the choice of the lattice Dirac operator than gluonic quantities like t_0 . In fig. 4 the results for $\sqrt{8t_0}F_\pi$ obtained at the $SU(3)$ -symmetric point (i.e. in the runs A_1 , B_1 and X_1) are compared with data published in ref. [32]. For a sensible comparison, ϕ_4 should

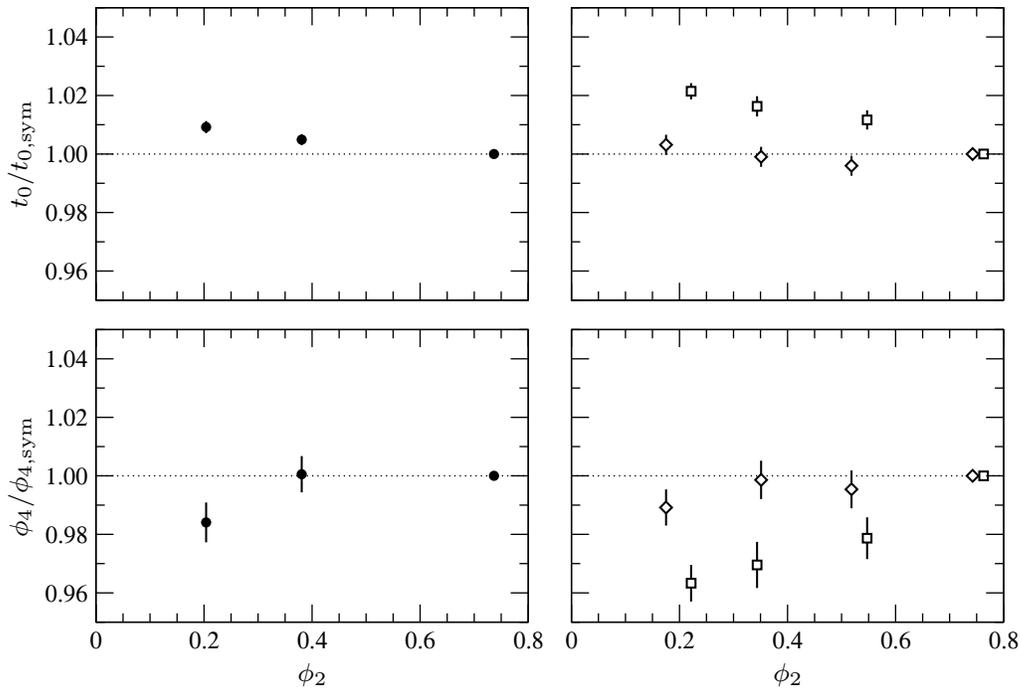


Fig. 5. Quark-mass dependence of t_0 and ϕ_4 along chiral trajectories (approximately) passing through the physical point. All data are normalized by the values measured at the SU(3)-symmetric points on the trajectories. The plots on the right show results reported in ref. [32] from simulations of two series of lattices with spacing $a = 0.086$ fm (open squares) and $a = 0.064$ fm (open diamonds), respectively, in which the traditional form of the improved lattice Dirac operator was used. On the left the results obtained in the runs A₁–A₃, where the lattice spacing is about 0.094 fm, are displayed (black circles).

assume the same value on all lattices. This is not exactly the case, but the variations in ϕ_4 are too small to have a noticeable effect. Different (non-perturbative) strategies to determine the axial-current renormalization constant however lead to values of F_π differing by 6% and more on the coarser lattices simulated (see fig. 4)[†]. The chosen renormalization condition for the axial current thus matters, and it is partly a consequence of the choice made here that the computed values of $\sqrt{8t_0}F_\pi$ coincide with the continuum value determined in ref. [32] within a margin of 0.7%.

[†] In ref. [32] the values of Z_A extracted from the so-called chirally rotated Schrödinger functional [39] were used. Even smaller values of Z_A were previously found by probing a chiral Ward identity on lattices with ordinary Schrödinger-functional boundary conditions [40].

The quark-mass dependence of t_0 and ϕ_4 along the chiral trajectories in parameter space provides another opportunity to study the magnitude of the lattice effects. As can be seen from the data plotted on the right of fig. 5, the dependence on the mass of the light quarks becomes flatter when the continuum limit is approached on the physical trajectories. The corresponding results obtained in runs A₁–A₃ (plots on the left) thus show that the use of the modified lattice Dirac operator leads, in the case of these observables, to significantly reduced lattice effects.

6.8 Miscellaneous remarks

SMD cycle timing. Per unit of molecular-dynamics time, the SMD algorithm updates the pseudo-fermion fields more frequently than the HMC algorithm and the number of accept-reject steps is larger as well. In practice this overhead however accounts for only a small fraction of the required computer time (about 5% in the runs A₁–A₃ and B₁), if the solutions of the Dirac equation are reused whenever this is profitable.

Algorithmic instabilities tend to have a cost in terms of computer time too. In run X₁, for example, the molecular-dynamics integration step size had to be reduced with respect to the one in run A₁ in order to achieve a similarly low rate of spikes in ΔH . The larger fluctuations of the gauge field moreover made it more difficult to obtain the solutions of the Dirac equation to the required precision. As a result, run X₁ consumed about two times the computer time spent for run A₁.

Autocorrelations. The simulations A₁–X₁ are not long enough for an accurate determination of the integrated autocorrelation times of the calculated physical quantities. Significant autocorrelations were however not observed in the runs A₁–A₃ and X₁ among measurements separated by 100 SMD update cycles, while practically decorrelated measurements were obtained in run B₁ at two times larger separations. The behaviour is thus largely the same as in the more extensive SMD simulations of the SU(3) gauge theory reported in ref. [11].

Statistical errors of F_π and F_K . Although the available statistics is rather limited, the renormalized pseudo-scalar decay constants are obtained with statistical errors of a fraction of a percent (cf. table 3). The smallness of these errors partly derives from the fact that the axial-current renormalization constants are determined on the same lattices as the bare decay constants. At the smaller light-quark masses, in particular, the error correlations are then such that the renormalized decay constants are obtained with significantly better statistical precision than the unrenormalized ones.

Simulations with open boundary conditions. As a further check on the viability of the modified lattice Dirac operator, a 96×48^3 lattice with open boundary conditions

in time [11] was simulated at $\beta = 4.1$. The run turned out to be completely stable, with zero spikes in ΔH , and no indication of an unexpected behaviour triggered by the presence of the boundaries of the lattice.

7. Conclusions

The alternative $O(a)$ -improved Wilson–Dirac operator and the algorithmic measures put forward in this paper have a stabilizing effect on lattice QCD simulations, particularly so on coarse lattices. With respect to the traditional setup, there do not appear to be any disadvantages. Non-perturbative improvement, for example, works out in much the same way and the observed residual lattice effects turned out to be smaller. Moreover, the simulations tend to be somewhat faster.

Another interesting outcome, unrelated to the stability issues, concerns the renormalization of the axial currents. On coarse lattices, different renormalization conditions can lead to significantly different results for the renormalized pseudo-scalar decay constants. Here the current renormalization factors were determined by probing chiral symmetry at positive gradient-flow time [34] and, with this choice, the calculated decay constants turned out to depend only weakly on the lattice spacing (see fig. 4). The renormalization procedure has further technical advantages, among them the fact that the computation of the renormalization factors does not require additional simulations and that the renormalized decay constants can be shown to be insensitive to the value of the axial-current improvement coefficient c_A .

While master-field simulations with light quarks await to be performed, there is every reason to expect such simulations to work out if the recommendations given in this paper are followed. Making most out of such simulations is however a non-trivial task. The evaluation of hadron propagators, for example, and their statistical errors will have to be reconsidered [1].

P.F. thanks the Institute of Theoretical Physics at the WWU Münster for hospitality and Jochen Heitger for useful discussions. Thanks also go to the ALPHA collaboration for supporting this work and to Mattia Bruno for providing a table of the CLS data plotted in fig. 4. A.R. is supported in part by the STFC Consolidated Grant ST/P000479/1. Many simulations reported in this paper were performed on a dedicated HPC cluster at CERN. This work also used the DiRAC Extreme Scaling service at the University of Edinburgh, operated by the Edinburgh Parallel Computing Centre on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk) and funded by BEIS capital funding via STFC capital

grant ST/R00238X/1 and STFC DiRAC Operations grant ST/R001006/1. DiRAC is part of the UK National e-Infrastructure. The generous support of all these institutions is gratefully acknowledged.

Appendix A. Implementation of the modified improved Dirac operator

If a chiral representation of the Dirac matrices is chosen, the Pauli term assumes the block-diagonal form

$$\frac{c_{\text{sw}}}{4 + m_0} \sum_{\mu, \nu=0}^3 \frac{i}{4} \sigma_{\mu\nu} \hat{F}_{\mu\nu} = \begin{pmatrix} A_+ & 0 \\ 0 & A_- \end{pmatrix} \quad (\text{A.1})$$

and the diagonal part (2.5) of the modified Dirac operator is then given by

$$D_{\text{ee}} + D_{\text{oo}} = \text{constant} \times \begin{pmatrix} e^{A_+} & 0 \\ 0 & e^{A_-} \end{pmatrix}. \quad (\text{A.2})$$

The blocks A_{\pm} in these equations are traceless Hermitian 6×6 matrices acting on the upper and lower two Dirac components of the quark fields, respectively.

Since the Pauli term already appears in the traditional form of the $O(a)$ -improved Dirac operator, its numerical evaluation and differentiation with respect to the gauge field are not discussed here. Instead some efficient methods to compute the exponential e^A and its differential de^A with respect to the parameters of A are described, where A stands for either A_+ or A_- .

A.1 Evaluation of the exponential function

The Cayley–Hamilton theorem asserts that any matrix satisfies its own characteristic equation. Specifically, in the case considered here, the equation

$$A^6 = \sum_{k=0}^4 p_k A^k \quad (\text{A.3})$$

holds, with real coefficients p_k given by

$$p_0 = \frac{1}{6} \text{tr}\{A^6\} - \frac{1}{8} \text{tr}\{A^4\} \text{tr}\{A^2\} - \frac{1}{18} \text{tr}\{A^3\}^2 + \frac{1}{48} \text{tr}\{A^2\}^3, \quad (\text{A.4})$$

$$p_1 = \frac{1}{5}\text{tr}\{A^5\} - \frac{1}{6}\text{tr}\{A^3\}\text{tr}\{A^2\}, \quad (\text{A.5})$$

$$p_2 = \frac{1}{4}\text{tr}\{A^4\} - \frac{1}{8}\text{tr}\{A^2\}^2, \quad (\text{A.6})$$

$$p_3 = \frac{1}{3}\text{tr}\{A^3\}, \quad (\text{A.7})$$

$$p_4 = \frac{1}{2}\text{tr}\{A^2\}. \quad (\text{A.8})$$

An arbitrary polynomial $\sum_{k=0}^N c_k A^k$ in A of degree $N \geq 6$ can therefore be reduced to a polynomial of degree 5 with A -dependent coefficients d_0, \dots, d_5 , which may be calculated through a simple recursion.

The bound (2.4) implies that the eigenvalues λ of A satisfy

$$|\lambda| \leq R, \quad R = \frac{3c_{\text{sw}}}{4 + m_0}. \quad (\text{A.9})$$

At the values of c_{sw} and m_0 of interest, the Taylor series

$$e^A = \sum_{k=0}^N \frac{A^k}{k!} + r_N(A) \quad (\text{A.10})$$

is thus guaranteed to converge rapidly, with a remainder bounded by

$$\|r_N(A)\|_2 \leq \frac{R^{N+1}}{(N+1)!} e^R. \quad (\text{A.11})$$

Using the Cayley–Hamilton theorem, the exponential of the Pauli term can in this way easily be obtained to machine precision.

A.2 Differential of the exponential function

The differential of e^A with respect to the independent parameters of A is given by

$$de^A = \int_0^1 dt e^{tA} dA e^{(1-t)A}. \quad (\text{A.12})$$

Expansion of the exponentials on the right then yields the rapidly convergent series

$$de^A = \sum_{k=0}^N \sum_{l=0}^{N-k} \frac{1}{(k+l+1)!} A^k dA A^l + r_N(A, dA), \quad (\text{A.13})$$

where

$$\|r_N(A, dA)\|_2 \leq \frac{R^{N+1}}{(N+1)!} e^R \|dA\|_2. \quad (\text{A.14})$$

The derivatives of the exponential can thus be obtained to machine precision by truncating the series at the same value of N as the series (A.10). Application of the Cayley–Hamilton theorem and the recursion that derives from it finally leads to the expression

$$de^A = \sum_{k,l=0}^5 C_{kl} A^k dA A^l + r_N(A, dA), \quad (\text{A.15})$$

C being a real symmetric 6×6 matrix that depends on N and A .

In practice only matrix elements $(v, de^A w)$ of the differential need to be computed, where v and w are the upper or lower two Dirac components of quark spinors at a given lattice point. The evaluation of these matrix elements proceeds by computing the spinors $A^k v$ and $A^l w$ for all $k, l = 0, \dots, 5$, then the linear combinations $\sum_l C_{kl} A^l w$ and finally the scalar products of these with $A^k v$.

Appendix B. Properties of the uniform norm

In this appendix, some basic facts about the uniform norm (5.4) are briefly described, omitting the elementary but often lengthy proofs of the statements made.

B.1 Uniform norm of random fields

Let η be a random quark field with normal distribution. Its uniform norm, $r = \|\eta\|_\infty$, is a random variable, whose distribution $p(r)$ on a lattice with V points,

$$p(r) = \frac{d}{dr} \{f(r)^V\}, \quad f(r) = \frac{2}{11!} \int_0^r ds s^{23} e^{-s^2}, \quad (\text{B.1})$$

has an approximately Gaussian shape. When V increases from, say, 10^4 to 10^{12} , the position of the maximum of the distribution slowly moves from 5.42 to 7.37. Since the probability for $\|\eta\|_\infty$ to be larger than these values rapidly gets extremely small, the uniform norm of random fields is, in practice, typically in the range from 5 to 8.

B.2 Associated operator norm

The uniform norm of a linear operator A acting on quark fields $\psi(x)$ is defined by

$$\|A\|_\infty = \sup_{\psi \neq 0} \frac{\|A\psi\|_\infty}{\|\psi\|_\infty}. \quad (\text{B.2})$$

It has all the usual properties of an operator norm and in particular satisfies

$$\|AB\|_\infty \leq \|A\|_\infty \|B\|_\infty \quad (\text{B.3})$$

for any pair A, B of linear operators. The norm of A and its Hermitian conjugate A^\dagger however need not be the same, a notable exception being the γ_5 -Hermitian Wilson-Dirac operator.

In terms of the position-space kernel,

$$(A\psi)(x) = \sum_y A(x, y)\psi(y), \quad (\text{B.4})$$

the uniform norm is given by

$$\|A\|_\infty = \sup_x \left\{ \sup_{\|s\|_2=1} \sum_y \|A(x, y)^\dagger s\|_2 \right\}, \quad (\text{B.5})$$

where the inner supremum is taken over all y -independent spinors s of norm 1.

B.3 Uniform-norm condition number of the Dirac operator D

The uniform-norm condition number

$$\kappa_\infty(D) = \|D\|_\infty \|D^{-1}\|_\infty \quad (\text{B.6})$$

of the Dirac operator can be shown to be at least as large as the condition number $\kappa_2(D)$ defined through the square norm. While the bound

$$\|D\|_\infty \leq \|D_{\text{ee}} + D_{\text{oo}}\|_2 + 4\sqrt{2} \quad (\text{B.7})$$

is easily derived, an estimation of the second factor in eq. (B.6),

$$\|D^{-1}\|_\infty = \sup_x \left\{ \sup_{\|s\|_2=1} \sum_y \|S(x, y)^\dagger s\|_2 \right\}, \quad (\text{B.8})$$

requires the long-distance behaviour of the quark propagator $S(x, y)$ to be known. In the case of free quarks on an infinite lattice, for example, the leading asymptotic behaviour near the chiral limit,

$$\|D^{-1}\|_\infty \underset{m_0 \rightarrow 0}{=} \frac{c_1}{m_0} + \dots, \quad c_1 = 2.60(1), \quad (\text{B.9})$$

turns out to be practically the same as the one of the square norm.

The spontaneous breaking of chiral symmetry in QCD however leads to a more singular chiral behaviour. Assuming mass-degenerate up and down quarks, the norm of the light-quark propagator can be estimated by replacing $\|S(x, y)s\|_2^2$ through its average value, the pion propagator, and by using chiral perturbation theory for the latter. The asymptotic formula

$$\|D^{-1}\|_\infty \underset{m_{ud} \rightarrow 0}{=} \frac{c_2 F_\pi}{M_\pi Z_A^{ud} m_{ud}} + \dots, \quad c_2 = 36.1(1), \quad (\text{B.10})$$

obtained in this way then shows that the condition number $\kappa_\infty(D)$ must be expected to grow more rapidly than $\kappa_2(D)$ in the chiral limit, although in practice the factor $c_2 F_\pi / M_\pi$ never becomes very large.

Appendix C. Axial current renormalization

The goal in this appendix is to show that the renormalized decay constants F_π and F_K are insensitive to the value of the axial-current improvement coefficient c_A if the bare decay constants are extracted from the vacuum-to-meson matrix elements of the axial currents and if the currents are renormalized as described in ref. [34].

C.1 Masses and decay constants

In order to simplify the notation, an unspecified non-singlet flavour channel is considered, the flavour indices are omitted and the lattice spacing is set to unity. Moreover, the time extent T of the lattice is assumed to be sufficiently large that its effects are completely negligible.

The bare $O(a)$ -improved axial current

$$(A_I)_\mu(x) = A_\mu(x) + c_A \overset{\circ}{\partial}_\mu P(x) \quad (\text{C.1})$$

is a linear combination of the bare current $A_\mu(x)$ and the gradient of the axial density $P(x)$, where $\hat{\partial}_\mu$ denotes the symmetric nearest-neighbour difference operator [5]. At large times x_0 and up to exponentially suppressed terms, the two-point functions

$$\sum_{\vec{x}} \langle P(x)P(0) \rangle = -\frac{G^2}{M} e^{-Mx_0} + \dots, \quad (\text{C.2})$$

$$\sum_{\vec{x}} \langle (A_1)_0(x)P(0) \rangle = F_1 G e^{-Mx_0} + \dots, \quad (\text{C.3})$$

are given by the pseudo-scalar meson mass M , the vacuum-to-meson matrix element G of the axial density and the bare improved decay constant F_1 .

Following common practice, the sum m of the current-quark masses in the chosen flavour channel may be defined by requiring the PCAC relation

$$\sum_{\vec{x}} \langle \{\hat{\partial}_0 A_0(x) + c_A \partial_0^* \partial_0 P(x)\} P(0) \rangle = m \sum_{\vec{x}} \langle P(x)P(0) \rangle + \dots \quad (\text{C.4})$$

to hold at large x_0 , up to exponentially decaying terms with exponents larger than M , ∂_0 and ∂_0^* being the standard forward and backward difference operators. Insertion of eqs. (C.2),(C.3) then leads to the exact relation

$$m = \frac{M \hat{M} F_1}{G} - \frac{1}{4} c_A \hat{M}^4, \quad \hat{M} = \sinh(M), \quad \hat{M} = 2 \sinh(M/2). \quad (\text{C.5})$$

The term proportional to c_A in this equation is a lattice artefact of order a^3 , which could be removed by replacing the forward and backward difference operators in eq. (C.4) by $\hat{\partial}_0$.

C.2 Renormalization and dependence on c_A

As explained in ref. [34], the axial-current renormalization constant \hat{Z}_A can be computed by probing the PCAC relation at positive gradient-flow time. The renormalization constant calculated in this way includes the $O(am)$ corrections required for $O(a)$ -improvement [5,35].

The chiral Ward identities used to determine \hat{Z}_A (eqs. (8.8) and (8.10) in ref. [34]) are relations among two-point correlation functions summed over a range $[-d, d]$ of time x_0 . At large d the calculated values of \hat{Z}_A very rapidly become independent of d , the leading corrections decaying exponentially with exponents equal to the next-to-lowest energies in the flavour channel considered. The renormalization constant is then determined in the range of d , where these corrections can be safely neglected.

Table 5. PCAC quark-mass sums and axial-current renormalization constants

Run	c_A	am_{ud}	am_{us}	\hat{Z}_A^{ud}	\hat{Z}_A^{us}
A ₁	-0.039	0.021072(81)	–	0.8075(11)	–
A ₂	-0.039	0.010581(77)	0.026315(65)	0.8054(24)	0.8079(10)
A ₃	-0.039	0.00556(14)	0.028321(99)	0.8132(63)	0.8093(12)
B ₁	-0.035	0.014502(41)	–	0.8138(7)	–
X ₁	-0.0533	0.01924(17)	–	0.8066(22)	–

The improvement coefficient c_A appears in the Ward identities implicitly, through the quark-mass sum m , and explicitly multiplying one of the correlation functions. Since the latter decays exponentially at large d , and since \hat{Z}_A and m only occur in the combination $\hat{Z}_A m$ in this limit, it follows that the product $\hat{Z}_A m$ is independent of c_A . Recalling the PCAC relation (C.5), this implies that the renormalized improved decay constant $\hat{Z}_A F_1$ is insensitive to the value of c_A up to a tiny contribution of order a^3 deriving from the term proportional to c_A in eq. (C.5).

C.3 Table of renormalization factors

The results obtained in the runs A₁–X₁ for the quark-mass sums and the renormalization constants in the ud and us flavour channels are listed in table 5. While F_π and F_K are practically independent of c_A , the numbers quoted in the table are not and a choice of c_A thus had to be made (second column in table 5).

In the case of run X₁, where the traditional form of the Dirac operator was used, c_A was set to the non-perturbatively determined value obtained in ref. [41]. Lacking a similarly systematic determination of the coefficient in the theory with the modified Dirac operator, the other values of c_A listed in the table were estimated directly on the simulated lattices using the so-called LANL method [42,43]. To leading order, the perturbation expansion

$$c_A = -0.00603(3) \times g_0^2 + \mathcal{O}(g_0^4), \quad (\text{C.6})$$

incidentally coincides with the one in the theory with the traditional form of the Dirac operator and tree-level improved gauge action [44,45], since the relevant axial-current vertex diagram is the same. The coefficients of the $\mathcal{O}(am)$ contributions to the axial-current renormalization constant are, to this order, unchanged as well and the remark applies in the case of other composite fields too.

On coarse lattices, the calculated values of c_A may depend quite a bit on the chosen

improvement condition. Varying c_A by ± 0.05 however affects the renormalized decay constants only at a level of a quarter of a per mille, i.e. by amounts roughly an order of magnitude smaller than the statistical errors quoted in table 3.

Appendix D. Eigenvalue distributions

The low-lying eigenvalues of $(D^\dagger D)^{1/2}$ (where D is the light-quark lattice Dirac operator) were computed to a relative precision of 0.5% using the Chebyshev-accelerated subspace iteration described in appendix A of ref. [37]. This method delivers both the eigenvalues and eigenvectors, which allows the accuracy of the eigenvalues to be rigorously controlled.

Some of the simulations reported in sect. 6 use twisted-mass reweighting for the light quarks [12] and all use a rational approximation for the pseudo-fermion representation of the strange-quark determinant [28,29]. The probability densities $\rho(\lambda)$ of the lowest eigenvalue λ of $(D^\dagger D)^{1/2}$ shown in fig. 6 take the associated reweighting factors into account and so do the medians $\bar{\mu}$ and widths σ of the distributions quoted in table 4.

More precisely, the latter are defined as follows. Consider an ensemble of N gauge fields with normalized weights w_1, \dots, w_N and let $\lambda_1, \dots, \lambda_N$ be the computed lowest eigenvalues of $(D^\dagger D)^{1/2}$. The empirical probability for the eigenvalue to be less than or equal to λ is given by

$$P_N(\lambda) = \sum_{k=1}^N w_k (\lambda_k \leq \lambda), \quad (\text{D.1})$$

where the bracket is 1 if the enclosed condition is true and 0 otherwise. $P_N(\lambda)$ is a step function that increases monotonically from 0 to 1. The median of the eigenvalue distribution is then equal to $(u + v)/2$, $u \leq v$ being the largest and smallest step points satisfying $P_N(u) \leq 0.5$ and $P_N(v) \geq 0.5$, respectively. And its width is *half* the size $|v - u|$ of the smallest interval $[u, v]$ such that $\sum_{k=1}^N w_k (u \leq \lambda_k \leq v) \geq 0.683$.

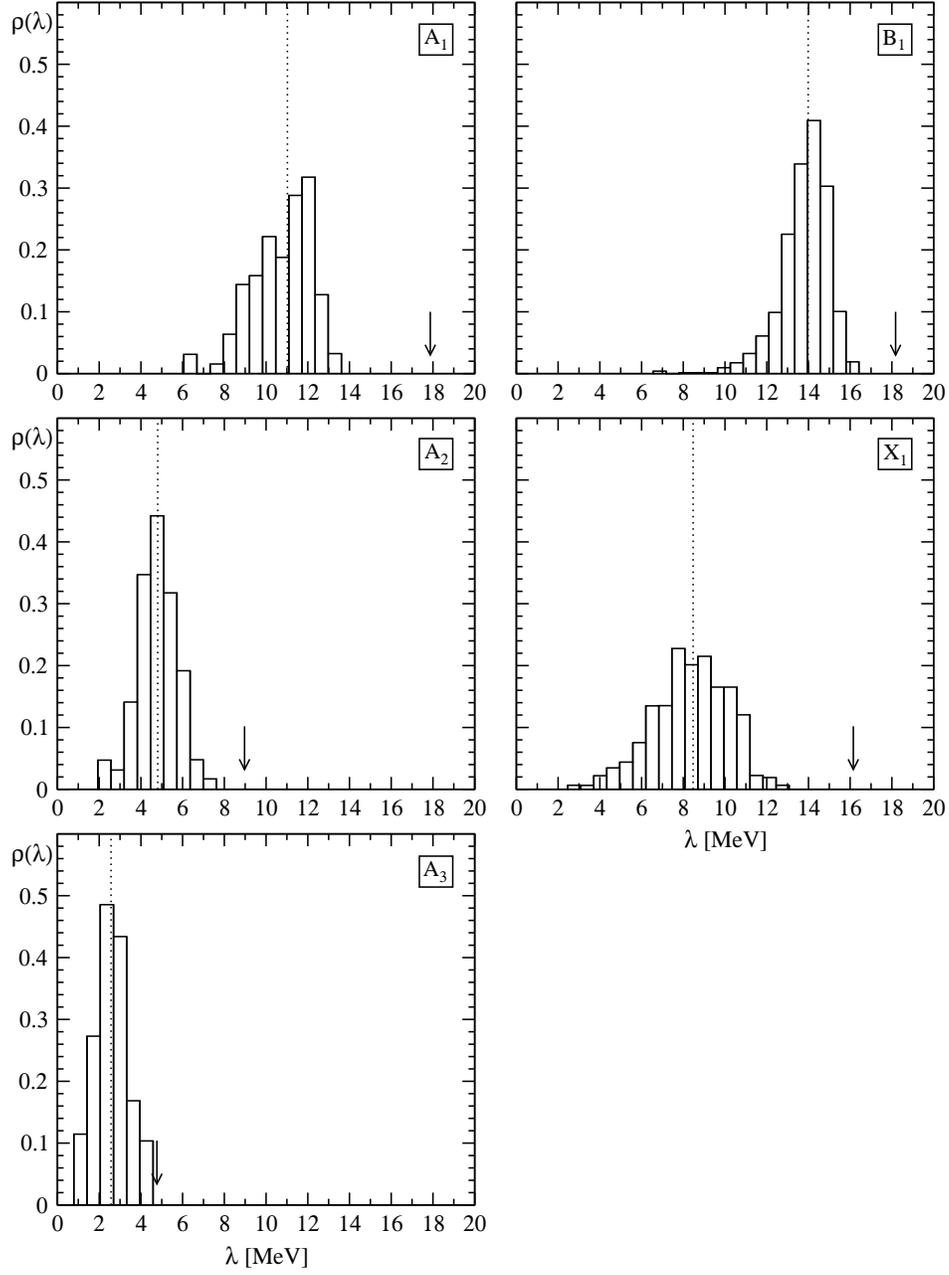


Fig. 6. Normalized distributions of the lowest eigenvalue λ of $(D^\dagger D)^{1/2}$. The dotted lines indicate the position of the median of the distributions and the arrows the one of $\frac{1}{2} \hat{Z}_A^{ud} m_{ud}$ (cf. table 4).

References

- [1] M. Lüscher, *Stochastic locality and master-field simulations of very large lattices*, EPJ Web Conf. 175 (2018) 01002
- [2] L. Giusti, M. Lüscher, *Topological susceptibility at $T > T_c$ from master-field simulations of the $SU(3)$ gauge theory*, Eur. Phys. J. C79 (2019) 207
- [3] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D10 (1974) 2445
- [4] B. Sheikholeslami, R. Wohlert, *Improved continuum limit lattice action for QCD with Wilson fermions*, Nucl. Phys. B259 (1985) 572
- [5] M. Lüscher, S. Sint, R. Sommer, P. Weisz, *Chiral symmetry and $O(a)$ improvement in lattice QCD*, Nucl. Phys. B478 (1996) 365
- [6] A. M. Horowitz, *Stochastic quantization in phase space*, Phys. Lett. 156B (1985) 89; *The second order Langevin equation and numerical simulations*, Nucl. Phys. B280 [FS18] (1987) 510; *A generalized guided Monte Carlo algorithm*, Phys. Lett. B268 (1991) 247
- [7] K. Jansen, C. Liu, *Kramers equation algorithm for simulations of QCD with two flavors of Wilson fermions and gauge group $SU(2)$* , Nucl.Phys. B453 (1995) 375 [E: *ibid.* B459 (1996) 437]
- [8] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, *Hybrid Monte Carlo*, Phys. Lett. B195 (1987) 216
- [9] M. Lüscher, R. Narayanan, P. Weisz, U. Wolff, *The Schrödinger functional: A renormalizable probe for non-Abelian gauge theories*, Nucl. Phys. B384 (1992) 168
- [10] S. Sint, *On the Schrödinger functional in QCD*, Nucl. Phys. B421 (1994) 135
- [11] M. Lüscher, S. Schaefer, *Lattice QCD without topology barriers*, JHEP 1107 (2011) 036
- [12] M. Lüscher, S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput. Phys. Commun. 184 (2013) 519
- [13] P. Weisz, *Continuum limit improved lattice action for pure Yang-Mills theory (I)*, Nucl. Phys. B212 (1983) 1
- [14] J. Bulava, S. Schaefer, *Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action*, Nucl. Phys. B874 (2013) 188
- [15] M. Lüscher, *Ergodicity of the SMD algorithm in lattice QCD*, unpublished notes (2017), <http://cern.ch/luscher/notes/smd-ergodicity.pdf>
- [16] I. P. Omelyan, I. M. Mryglod, R. Folk, *Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations*, Comp. Phys. Commun. 151 (2003) 272

- [17] R. C. Brower, T. Ivanenko, A. R. Levi, K. N. Orginos, *Chronological inversion method for the Dirac matrix in hybrid Monte Carlo*, Nucl. Phys. B484 (1997) 353
- [18] M. Lüscher, *Local coherence and deflation of the low quark modes in lattice QCD*, JHEP 0707 (2007) 081; *Deflation acceleration of lattice QCD simulations*, JHEP 0712 (2007) 011
- [19] D. E. Knuth, *Semi-Numerical Algorithms*, in: The Art of Computer Programming, vol. 2, 2nd ed. (Addison-Wesley, Reading MA, 1981)
- [20] C. Urbach, *Reversibility violation in the Hybrid Monte Carlo algorithm*, Comput. Phys. Commun. 224 (2018) 44
- [21] T. J. Dekker, *A floating-point technique for extending the available precision*, Numer. Math. 18 (1971) 224
- [22] J. R. Shewchuk, *Adaptive precision floating-point arithmetic and fast robust geometric predicates*, Discrete & Computational Geometry 18 (1997) 305
- [23] Y. Saad, *Iterative methods for sparse linear systems*, 2nd ed. (SIAM, Philadelphia, 2003); see also <http://www-users.cs.umn.edu/~saad/>.
- [24] M. Hasenbusch, *Speeding up the Hybrid Monte Carlo algorithm for dynamical fermions*, Phys. Lett. B519 (2001) 177
- [25] M. Hasenbusch, K. Jansen, *Speeding up lattice QCD simulations with clover-improved Wilson fermions*, Nucl. Phys. B659 (2003) 299
- [26] M. Lüscher, *Schwarz-preconditioned HMC algorithm for two-flavor lattice QCD*, Comp. Phys. Commun. 165 (2005) 199
- [27] C. Urbach, K. Jansen, A. Shindler, U. Wenger, *HMC algorithm with multiple time scale integration and mass preconditioning*, Comp. Phys. Commun. 174 (2006) 87
- [28] I. Horváth, A. D. Kennedy, S. Sint, *A new exact method for dynamical fermion computations with non-local actions*, Nucl. Phys. (Proc. Suppl.) 73 (1999) 834
- [29] M. A. Clark, A. D. Kennedy, *Accelerating dynamical fermion computations using the Rational Hybrid Monte Carlo (RHMC) algorithm with multiple pseudo-fermion fields*, Phys. Rev. Lett. 98 (2007) 051601
- [30] M. Lüscher, S. Sint, R. Sommer, P. Weisz, U. Wolff, *Nonperturbative $O(a)$ improvement of lattice QCD*, Nucl.Phys. B491 (1997) 323-343
- [31] M. Bruno et al., *Simulation of QCD with $N_f = 2+1$ flavors of non-perturbatively improved Wilson fermions*, JHEP 1502 (2015) 043
- [32] M. Bruno, T. Korzec, S. Schaefer, *Setting the scale for the CLS 2+1 flavor ensembles*, Phys. Rev. D95 (2017) 074504
- [33] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, JHEP 1008 (2010) 071 [Erratum: *ibid.* 1403 (2014) 092]

- [34] M. Lüscher, *Chiral symmetry and the Yang–Mills gradient flow*, JHEP 1304 (2013) 123
- [35] T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe, J. M. S. Wu, *Improved bilinears in lattice QCD with non-degenerate quarks*, Phys. Rev. D73 (2006) 034504
- [36] <http://cern.ch/luscher/openQCD>
- [37] L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, N. Tantalo, *Stability of lattice QCD simulations and the thermodynamic limit*, JHEP 0602 (2006) 011
- [38] L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, N. Tantalo, *QCD with light Wilson quarks on fine lattices (II): DD-HMC simulations and data analysis*, JHEP 0702 (2007) 082
- [39] M. Dalla Brida, T. Korzec, S. Sint, P. Vilaseca, *High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks*, Eur. Phys. J. C79 (2019) 23
- [40] J. Bulava, M. Della Morte, J. Heitger, C. Wittmeier, *Nonperturbative renormalization of the axial current in $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action*, Phys. Rev. D93 (2016) 114513
- [41] J. Bulava, M. Della Morte, J. Heitger, C. Wittmeier, *Non-perturbative improvement of the axial current in $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action*, Nucl. Phys. B896 (2015) 555
- [42] T. Bhattacharya, R. Gupta, W. Lee, S. Sharpe, *Order “a” improved renormalization constants*, Phys. Rev. D63 (2001) 074505
- [43] S. Collins, C. T. H. Davies, G. P. Lepage, J. Shigemitsu, *A nonperturbative determination of the $O(a)$ improvement coefficient c_A and the scaling of f_π and $m^{\overline{\text{MS}}}$* , Phys. Rev. D67 (2003) 014504
- [44] Y. Taniguchi, A. Ukawa, *Perturbative calculation of improvement coefficients to $O(g^2a)$ for bilinear quark operators in lattice QCD*, Phys. Rev. D58 (1998) 114503
- [45] S. Aoki, R. Frezzotti, P. Weisz, *Computation of the improvement coefficient c_{sw} to 1-loop with improved gluon actions*, Nucl. Phys. B540 (1999) 501