One flavor mass reweighting in lattice QCD

Jacob Finkenrath\textsuperscript{a}, Francesco Knechtli\textsuperscript{a}, Björn Leder\textsuperscript{a,b}

\textsuperscript{a}Department of Physics, Bergische Universität Wuppertal, Gaussstr. 20, D-42119 Wuppertal, Germany
\textsuperscript{b}Department of Mathematics, Bergische Universität Wuppertal, Gaussstr. 20, D-42119 Wuppertal, Germany

Abstract

One flavor mass reweighting can be used in lattice QCD computations to fine tune the quark masses to their physical values. We present a new method that utilizes an unbiased stochastic estimation of the one flavor determinant. The stochastic estimation is based on the integral representation of the determinant of a complex matrix, which we prove. In contrast to other methods it can also be applied in situations where the determinant has a complex phase. The stochastic error is controlled by determinant factorizations based on mass interpolation and Schur decomposition. As an example of an application we demonstrate how the method can be used to tune the up-down quark mass difference.

1. Bare parameter reweighting

Reweighting is an old method introduced in \cite{1} to obtain results for a range of parameters by using a single Monte Carlo simulation at one value of the parameters. Reweighting can also be used to modify the sampling in a Monte Carlo simulation. In both cases the result for the desired ensemble is obtained by including a reweighting factor in the observables. In lattice quantum chromodynamics (QCD) simulations the reweighting factor is typically a determinant of a ratio of Dirac operators, e.g. \cite{2, 3, 4, 5}. We close a gap, which to our knowledge exists in the literature, by proving an integral representation for the determinant of a complex matrix. We prove the efficiency of this method by applying it to the case of one flavor reweighting in the quark mass.

The expectation value of an observable $\mathcal{O}$ in lattice QCD is given by the path integral

$$\langle \mathcal{O} \rangle = \frac{1}{Z_a} \int \mathcal{D}[U] \, P_a(U) \, \mathcal{O}(U),$$

where the gauge link configuration $U$ receives the weight

$$P_a(U) = e^{-S_a(\beta, U)} \prod_{i=1}^{n_f} \det(D(U) + m_i).$$
The expectation value \( \langle \cdot \rangle_a \) is defined at the set of bare parameters
\[
a = \{ \beta, m_1, m_2, \ldots, m_{n_f} \},
\]
where \( \beta = 6/g^2 \) is the bare inverse gauge coupling and \( m_1, m_2, \ldots, m_{n_f} \) are the \( n_f \) bare mass parameters of the sea quarks. Note that in (1) the Grassmann valued quark fields have been integrated out yielding the product of determinants in (2). Therein the “massless” Dirac operator \( D \) is shifted by the corresponding mass parameter. The normalization \( Z \) is fixed by demanding
\[
\int D[U]P(U)/Z = 1.
\]

In numerical computations the path integral in Eq. (1) is evaluated employing Monte Carlo methods. The expectation value is given as the ensemble mean
\[
\langle O \rangle_a = \frac{1}{N} \sum_i O(U_i) + O(1/\sqrt{N}),
\]
over an ensemble of gauge field configurations \( \{U_i, i = 1, \ldots, N\} \) that has been generated according to the probability distribution \( P_a(U) \). Since the generation of such ensembles is a computer time intensive sequential process it is desirable to reuse a generated ensemble to compute the expectation value at a different set of parameters. This is the idea of reweighting.

The expectation value at a set of parameters \( b = \{ \beta', m'_1, m'_2, \ldots, m'_{n_f} \} \) different from \( a \), can be expressed in terms of expectation values at \( a \) via
\[
\langle O \rangle_b = \langle OW_{a,b} \rangle_a, \quad W_{a,b} = \frac{P_b}{P_a}.
\]

The ratio \( W_{a,b} \) is the so-called reweighting factor. Note that the observable \( O \) is the same on both sides of Eq. (4). In particular, if it explicitly depends on the bare parameters they have to be the same on both sides.

In the simplest case the two sets differ only in one parameter and identical terms trivially cancel in the ratio \( W_{a,b} \). Let us distinguish the beta shift reweighting factor, where \( a \) and \( b \) differ in the inverse bare coupling
\[
W_{\beta,\beta'} = e^{-\left( S_g(\beta',U) - S_g(\beta,U) \right)}. \quad (5)
\]

The one flavor reweighting factor, where the two sets differ in the bare mass parameter for quark \( i \), is given by the determinant of the ratio of the corresponding Dirac operators
\[
W_{m_i, m'_i} = \det \frac{D(U) + m'_i}{D(U) + m_i}. \quad (6)
\]

The other cases, where two or more parameters differ, are products of beta shift and one flavor reweighting factors.

Beside the one flavor reweighting Eq. (6) the two flavor reweighting is of practical importance. In the most general case it is the product of two one flavor reweighting factors \( W_{2f} = W_{m_i, m'_i}W_{m_j, m'_j} \). We distinguish the special cases of reweighting two degenerate quarks \( W_{2f} = W_{m_i, m'_i}^2 \), and the isospin
reweighting \( W_{2f} \equiv W_\pm = W_{m,m-\Delta m}W_{m,m+\Delta m} \), where two degenerate quarks are reweighted in opposite directions. The one and two flavor reweighting may be combined with a beta shift.

Additional reweighting factors occur when QED effects or a finite chemical potential are considered. Although we present numerical results exclusively for reweighting in the mass and the coupling, the proof of the integral representation of a determinant (Section 2.1) is an important general result.

2. Stochastic estimation

Since \( D \) is a large sparse matrix the determinant in Eq. 6 has to be computed using stochastic methods. In this section we define an unbiased estimator with controlled convergence.

2.1. Integral representation

Let \( A \) be a complex matrix with eigenvalues \( \lambda(A) \) and \( \eta \) a complex valued vector. In Appendix A we prove

\[
\frac{1}{\det A} = \int D[\eta] e^{-\eta^\dagger A \eta} \quad \text{if} \quad \lambda(A + A^\dagger) > 0. \tag{7}
\]

The condition \( \lambda(A + A^\dagger) > 0 \) is the necessary and sufficient condition for the absolute convergence of the integral in Eq. (7). It is equivalent to the condition that the field of values \( \mathcal{F}(A) \) is in the right half plane, see Appendix A. It implies the weaker condition \( \text{Re} \lambda(A) > 0 \) (see Appendix A) but the two conditions are only equivalent for normal matrices.

If the integral in Eq. (7) exists it can be evaluated with Monte Carlo methods. Let \( \{\eta_k; k = 1, \ldots, N_\eta\} \) be an ensemble of random complex vectors with probability distribution \( p(\eta) \). The determinant can then be written as

\[
\frac{1}{\det A} = \left\langle \frac{e^{-\eta^\dagger A \eta}}{p(\eta)} \right\rangle_{p(\eta)} = \frac{1}{N_\eta} \sum_{k=1}^{N_\eta} \frac{e^{-\eta_k^\dagger A \eta_k}}{p(\eta_k)} + O(1/\sqrt{N_\eta}), \tag{8}
\]

where \( \left\langle O \right\rangle_{p(\eta)} = \int D[\eta] p(\eta) O(\eta) \) in accordance with Eq. (1). It is convenient to choose a Gaussian distribution \( p(\eta) = \exp(-\eta^\dagger \eta) \), which we will assume from now on. The ensemble mean in Eq. (8) is an estimator of the inverse of the determinant and it converges if the integral in Eq. (7) exists. The variance of this estimator is given by

\[
\sigma^2_\eta = \left\langle \frac{e^{-\eta^\dagger (A + A^\dagger) \eta}}{p(\eta)^2} \right\rangle_{p(\eta)} - \left\langle \frac{e^{-\eta^\dagger A \eta}}{p(\eta)} \right\rangle_{p(\eta)} \left\langle \frac{e^{-\eta^\dagger A^\dagger \eta}}{p(\eta)} \right\rangle_{p(\eta)} \tag{9}
\]

\[
= \frac{1}{\det(A + A^\dagger - I)} - \frac{1}{\det(A A^\dagger)}. \tag{10}
\]
The first integral in Eq. (9) exists if
\[ \lambda(A + A^\dagger - I) > 0 \Leftrightarrow \lambda(A + A^\dagger) > 1 \Rightarrow \text{Re}(\lambda(A)) > 0.5. \] (11)

The condition \( \lambda(A + A^\dagger) > 1 \) automatically implies the existence of the second and third integral, cf. Eq. (7). Therefore, if in an implementation of the estimator in Eq. (8) the variance (9) is monitored, its convergence assures the convergence of the mean.

Often the matrix \( A \) can be written in the form \( A = I + \epsilon B \) with \( \epsilon \| B \| \ll 1 \).

Then a useful approximation of the variance (9) is obtained by expanding in \( \epsilon \)
\[
\frac{\sigma^2}{|\det A|^{-2}} = \det \left( I + \epsilon^2 \frac{BB^\dagger}{I + \epsilon(B + B^\dagger)} \right) - 1 = \epsilon^2 \text{Tr}(BB^\dagger) + O(\epsilon^3),
\] (12)
where we used \( AA^\dagger = I + \epsilon(B + B^\dagger) + \epsilon^2 BB^\dagger \) and \( \det(X) = \exp(\text{Tr} \ln(X)) \) for positive definite \( X \).

### 2.2. One flavor reweighting

The one flavor reweighting factor (6) is given by the determinant of the ratio of two Dirac operators. Dropping the subscripts and explicit dependence on the gauge field from now on, and defining the reweighting distance \( \Delta m = m - m' \) and \( D_m = D + m \), the reweighting factor \( W_{m,m'} \) can be written as
\[
W = \frac{1}{\det M} \quad M_{m^{-1}}D_{m'} = I - \frac{\Delta m}{D_m}.
\] (13)

Because the eigenvalues of \( M + M^\dagger \) converge to two for \( \Delta m \to 0 \), the integral representation of \( W \) in Eq. (13) exists for sufficiently small \( \Delta m \) and non-singular \( D_m \) and \( D_{m'} \). Since \( M = I + \epsilon B \) with \( \epsilon = \Delta m \) and \( B = D_m^{-1} + O(\Delta m) \) the variance (9) of stochastically estimating \( W \) via (8) can be approximated with Eq. (12)
\[
\frac{\sigma^2}{|W|^2} = \Delta m^2 \text{Tr}((D_mD_m^\dagger)^{-1}) + O(\Delta m^3).
\] (14)

In contrast to conventional methods \[6, 4\] the estimator also works in situations where the determinant is negative or complex (for example, it becomes negative when a real eigenvalue of \( D \) crosses zero between the shifts \( m \) and \( m' \)).

### 2.3. Mass interpolation

An estimator where the stochastic variance can be reduced (and its convergence controlled) at fixed \( \Delta m \) is obtained when the matrix \( M \), and thus the determinant, is factorized \( M = \prod_{l=0}^{N-1} M_l \). One possible factorization is given by taking the \( N \)th root \( M_l = M_l^{1/N} \) \[7\]. Another possibility that does not involve the evaluation of a matrix function is an interpolation between \( m \) and \( m' \) \[2\]. In the context of the Hybrid Monte Carlo algorithm a similar method is known as “mass preconditioning” \[8\], in the context of the Partially Stochastic Multi-Step algorithm as “gauge link interpolation” \[9\]. The idea is to introduce
intermediate reweighting factors with a reweighting “distance” that scales with $1/N$. In the case of mass reweighting we set $\delta m = \Delta m / N = (m - m') / N$ and

$$M_l^{-1} = \frac{D_{m_l+1}}{D_{m_l}} = I - \frac{\delta m}{D_{m_l}}, \quad m_l = m - l\delta m.$$  (15)

The estimate of the whole reweighting factor in Eq. (13) is now given as a product of estimates

$$W = \prod_{l=0}^{N-1} W_l, \quad W_l = \left\langle e^{-\eta(l)\eta(l)^\dagger} M_{\eta(l)} \right\rangle_{p(\eta(l))},$$  (16)

where for each factor $W_l$ an independent estimator is used. The variance of each factor can be approximated using Eq. (12). If in the Monte Carlo process for each factor $N_\eta$ random vectors are used (cf. Eq. (8)) the relative stochastic error $\delta_\eta^2 = \sigma_\eta^2 / (|W|^2 N_\eta)$ of the product is given by

$$\delta_\eta^2 = \frac{N}{N_\eta} [\delta m^2 \text{Tr} \left( (D_mD_m^\dagger)^{-1} \right) + O(\delta m^2 \Delta m)] \approx \frac{\Delta m^2}{NN_\eta} \text{Tr} \left( (D_mD_m^\dagger)^{-1} \right),$$  (17)

where in the second step we assumed the higher order terms to be negligible. This can always be achieved by choosing $N$ large enough (i.e., $\delta m$ small enough) and then Eq. (17) states that increasing $N$ or $N_\eta$ has the same effect on decreasing the stochastic error. Note that the above arguments only hold if the estimate for each factor is converging. Again, as long as $D_{m_l}$ does not become singular, convergence can be assured by choosing $N$ large enough, i.e., squeezing $\lambda(M_l)$ in a small region around one.

### 2.4. Schur decomposition

Yet another factorization that reduces the stochastic noise can be employed if the lattice Dirac operator $D$ only involves next-neighbor couplings. In this case the lattice sites can be labeled even and odd, depending on whether the sum of their lattice coordinates is even or odd. Reordering the entries in $D$ such that all even sites come first the determinant of $D$ becomes a product

$$\det(D + m) = \det(D_{ee} + m) \det(D_{oo} + m) \det(\hat{D}_m).$$  (18)

of the determinants of the (block) diagonal $D_{ee} + m$ and $D_{oo} + m$ and the Schur complement $\hat{D}_m = I - (D_{ee} + m)^{-1} D_{oo}(D_{oo} + m)^{-1} D_{ee}$. The two former ones can be evaluated exactly. The latter acts non-trivially only on the even lattice sites and therefore $\dim(\hat{D}_m) = \dim(D) / 2$. The one flavor reweighting factor employing even-odd factorization and stochastic estimation of the Schur complement ratio with mass interpolation reads

$$W = \frac{1}{\det M_{ee} \det M_{oo}} \prod_{l=0}^{N-1} \hat{W}_l, \quad \hat{W}_l = \left\langle e^{-\eta(l)\eta(l)^\dagger} \hat{M}_l \eta(l) \right\rangle_{p(\eta(l))},$$  (19)
where \( M_{ee} = (D_{ee} + m)/(D_{ee} + m') \) and similarly for \( M_{oo} \). Noting that the Schur complement can be expanded \( \hat{D}_{m+\delta m} = \hat{D}_m + \delta m X_m + O(\delta m^2) \) with

\[
X_m = -(D_{ee} + m)^{-1} \left[ \hat{D}_m - I - D_{eo}(D_{oo} + m)^{-2}D_{ee} \right], \tag{20}
\]

the ratio \( \hat{M}_l \) can be cast in a form similar to Eq. (15)

\[
\hat{M}_l^{-1} = \frac{\hat{D}_{m+1}}{\hat{D}_m} = I - \delta m \hat{D}_m^{-1} X_m + O(\delta m^2). \tag{21}
\]

The relative stochastic error of the product \( \prod_l \hat{W}_l \) is then approximated by

\[
\delta^2 \approx \frac{\Delta m^2}{N N_\eta} \text{Tr} \left( X_m X_m^\dagger (\hat{D}_m \hat{D}_m^\dagger)^{-1} \right). \tag{22}
\]

2.5. Zero crossings

In Fig. 1 we show the variance of the one flavor reweighting factor for one configuration of the ensemble D5 with mass \( m \). Along the mass interpolation with \( N = 512 \) steps (dots) to reweight to the mass \( m' \) (of D6) a real eigenvalue of \( D_m \) crosses zero, as is indicated by the “peak” in the variance. We refer to Section 3 for the explanations of the ensemble parameters. For this analysis we used a domain decomposition with \( 6^4 \) blocks and the variance is for the stochastic estimation of the reweighting factor of the Schur complement only (cf. Eq. (19) and comment on domain decomposition in Section 2.6). The number of mass interpolation steps is chosen such that a good resolution of the peak is achieved. A fit of the variances to \( \propto |l - k|^{-p} \) yields a good description
of the data (line) with \( k \approx 322 \) and \( p \approx 1.8 \). An exponent \( p = 2 \) is expected if the trace in Eq. (22) is dominated by one small eigenvalue.

The behavior of the variance can be regularized by a detour in the mass-\( \mu \) plane (squares) as explained in Appendix B. The crucial observation is the well known fact that the singular values of \( D_{m,\mu} = D_m + i \mu \gamma_5 \) are bounded from below by \( |\mu| \). The detour starts at \( l_s = 314 \) and ends at \( l_e = 330 \). The maximal value of \( a\mu \) is 0.00003 and the detour is divided into 48 steps. The reweighting factor is complex for \( \mu \neq 0 \) and we plot the sum of the variance of the real and imaginary parts. As expected by construction the variance is bounded from above and never larger than \( 10^{-1} \) along the detour.

Finally we remark that although the single factors along the detour are complex, their product is real within errors. The imaginary parts add up to a phase equal to minus one giving the correct sign of the one flavor reweighting factor.

### 2.6. Summary

Summarizing, the stochastic estimator of the one flavor reweighting factor in Eq. (6) using factorizations based on even-odd and mass interpolation is given by Eq. (19). The expectation values of the individual factors \( \tilde{W}_l \) with respect to \( p(\eta^{(l)}) \) are obtained via Monte Carlo integration as in Eq. (8) with a fixed \( N_\eta \) for all \( l \). Estimates of the variances of the factors are easily obtained if \( N_\eta \geq 6 \). If a zero crossing is observed (variance diverges) the mass interpolation is modified as described in Section 2.5. Writing the trace in Eq. (22) as \( k_\eta V \) (with the lattice volume \( V \)) and if higher orders in \( \delta m \) are negligible, the relative stochastic error is given by

\[
\delta^2_\eta \approx k_\eta \frac{\Delta m^2 V}{N N_\eta}.
\]

We close this section by listing some practical remarks:

- From the results of [10] we deduce that higher order terms are negligible for \( \delta m ||D_m^{-1}|| \approx \Delta m/(mN) \lesssim 1/16 \), where the smallest eigenvalue of \( D_m \) is approximated with the renormalized quark mass \( \overline{m} \). In Fig. 1 of Ref. [10] higher order terms are negligible for \( N \gtrsim 8 \) and \( \Delta m = \overline{m}/2 \), leading to the stated relation.

- The number of inversions of the global Dirac operator is given by \( N \cdot N_\eta \).

- The value of \( k_\eta \) with even-odd factorization is roughly a factor two smaller than without. Instead of the presented even-odd factorization, a factorization based on a domain decomposition of the lattice is also possible. However, the additional gain is small [10] and has to be compared to the additional cost for the needed (exact) block determinants.

- In the case of Wilson fermions, \( D^\dagger = \gamma_5 D \gamma_5 \) and it follows that the determinant \( \det M^{-1} \) is real. If the determinant is known to be real an improved estimator can be defined by replacing \( \langle \cdot \rangle_{p(\eta)} \to \langle \text{Re}(\cdot) \rangle_{p(\eta)} \) in
Eq. (8). In practice this yields an improvement if $N_\eta$ is small. The real part cannot be taken if we replace $D$ by $D + i\mu \gamma_5$ as discussed in Section 2.5.

- The vectors $A\eta^{(k)}$ that are computed in Eq. (8) also appear in the estimator for $1/\det(AA^\dagger)$. Therefore, if $D^\dagger = \gamma_5 D \gamma_5$, the reweighting factors $W$ and $W^2$ are obtained at once.

3. Ensemble fluctuations

We present results for reweighting of an ensemble generated with the Wilson gauge action and two flavors of $O(a)$ improved Wilson fermions at parameters $\beta = 5.3$, $c_{sw} = 1.90952$, $\kappa = 0.13625$ on a $48 \times 24^3$ lattice. The lattice spacing is $a = 0.0658(10)$ [11]. We use the hopping parameter $\kappa$ which is related to the bare quark mass $m$ by $\kappa = 1/(2am + 8)$. The parameters and boundary conditions are the same as the ones of the D5 simulation of [12]. The pseudoscalar mass is $m_{PS} \approx 440$ MeV and the renormalized quark mass is $\overline{m}(D5) \approx m_s/3$ [11], where $m_s$ is the physical strange quark mass. We generated a statistics of 2012 configurations, separated by 2.0 molecular dynamics units, using mass preconditioned HMC [13]. The target renormalized quark mass $\overline{m}$ is a factor two smaller, i.e., $\overline{m} \approx \overline{m}_s/6$. It corresponds to the parameter $\kappa' = 0.136350$ or $m_{PS}' \approx 310$ MeV, as in the D6 simulation of [3].

After analyzing the ensemble fluctuations of the one flavor reweighting factor we show how correlated reweighting factors can be combined in order to reduce the ensemble fluctuations. In particular, we present the combination of one and two (degenerate) flavor reweighting with a beta shift and combinations of different one flavor reweighting factors.

3.1. One flavor reweighting

The one flavor mass reweighting factor Eq. (13) can be written as

$$W = \exp \left[ \text{Tr} \ln \left( I - \frac{\Delta m}{D_m} \right) \right].$$

Expanding for small $\Delta m$, the relative ensemble fluctuations of the reweighting factor can be shown to be

$$\frac{\sigma_1^2}{\langle W \rangle^2} = \frac{\langle W^2 \rangle}{\langle W \rangle^2} - 1 = \Delta m^2 \left[ \langle (\text{Tr}(D_m^{-1}))^2 \rangle - \langle \text{Tr}(D_m^{-1}) \rangle^2 \right] + O(\Delta m^3).$$

In Fig. 2 we plot the history of the one flavor reweighting factor, evaluated on each configuration with the techniques explained in Section 2. The configuration with a very small reweighting factor corresponds to an exceptional configuration [14, 15] (without zero crossing), on which the pseudoscalar propagator has (almost) a pole. This pole is regulated by the reweighting factor.

Numerically we find that the ensemble fluctuations scale proportional to the lattice volume $V$ for the full Dirac operator (there are indications for a weaker
volume dependence for the Schur complement only), cf. Fig. 2 of Ref. [10]. Thus the ensemble fluctuations can be described by

$$\frac{\sigma^2_{1f}}{\langle W \rangle^2} \approx k_{1f} \Delta m^2 V,$$

(26)

and the leading dependence on the lattice volume $V$ and the reweighting distance $\Delta m$ cancels in the ratio of relative stochastic error Eq. (23) and ensemble fluctuations Eq. (26). Eventually we will require that this ratio is at the level of 10%. In order to achieve this we have to determine how many inversions $NN_\eta$ are necessary. Assuming

$$\frac{\langle \delta^2_\eta \rangle}{\sigma^2_{1f}/\langle W \rangle^2} = \frac{\langle k_\eta \rangle}{k_{1f} N N_\eta},$$

(27)

the ratio $\langle k_\eta \rangle/k_{1f}$ can be determined from the measured relative stochastic error and the ensemble fluctuations for different $V$, $\Delta m$ and renormalized quark masses $\overline{m}$. We find it to be independent of the lattice volume and to mildly change between 2 and 3 if $\Delta m$ and $\overline{m}$ are varied in the ranges $0 \leq \Delta m \leq m/2$ and $m_s/6 \leq \overline{m} \leq 4/3 \cdot m_s$. The value of $k_{1f}$, and thus the fluctuations, can be reduced by exploiting the correlation between one flavor and beta shift reweighting factor, which we explain next.

Finally we remark that in the computation of the statistical error of reweighted observables the correlation between the observable and the reweighting factor in Eq. (4) has to be taken into account and this requires some care. For example, we use [16] where we incorporated the analytic derivatives of the quotient in Eq. (4).

$^1$For $\overline{m} \neq \overline{m}(D5) \approx m_s/3$ this means a partially quenched determination of the reweighting factor.
3.2. Beta shift

We introduce a beta shift, cf. Eq. (5), by minimizing the variance of
\[
\ln(W_{\beta,\beta'}W) = -\Delta \beta S_g(U) + \text{Tr} \ln \left( I - \frac{\Delta m}{D_m} \right),
\]  
(28)
as a function of \( \Delta \beta = \beta' - \beta \). For the two (degenerate) flavor case, one can see from Eq. (28), by making an expansion for small mass shifts \( \Delta m \), that \( \Delta \beta \) is to first approximation a function of the sum of the masses of the two flavors. Furthermore, for \( N_f \) mass degenerate flavors \( \Delta \beta \) is proportional to \( N_f \). Numerically we find \( \Delta \beta/N_f \simeq -3 \times 10^{-4} \) for \( \Delta m = m/2 \).

The fluctuations \( \sigma^2_{\Delta \beta}/\langle W_{\Delta \beta} \rangle^2 \) of the reweighting factor \( W_{\Delta \beta} = W_{\beta,\beta'} W^{N_f} \) are reduced by roughly a factor two with respect to \( W^{N_f} \) alone, i.e., we find \( k_{\Delta \beta}/k_{1f} \approx 0.4 \).

Our result means that the ensemble fluctuations are minimized if the lattice spacing is increased when decreasing the quark mass(es). The effect is shown in Fig. 3. In order to quantify it we use the reference scale \( t_0 \) introduced in [18]. The change of the lattice spacing is given by \( a(\beta')/a(\beta) = \sqrt{t_0(\beta)/t_0(\beta')} \) and is plotted as a function of the non-singlet pseudoscalar mass \( m^2_{PS}t_0 \) in units of \( t_0 \), for \( N_f = 1 \) and \( N_f = 2 \). We observe that the lattice spacing \( a \) changes by 1%-2% at most and therefore lies within the accuracy of \( a \) quoted in [11].

Since the lattice spacing grows with the beta shift, the reweighting distance \( \Delta m \) in physical units is less than it is without beta shift. We also observed that the beta shift keeps the value of the plaquette constant within errors.

---

2The change in the clover coefficient \( c_{sw} \) [17] is of the same order of magnitude and can be therefore neglected.
3.3. Two flavor reweighting

Correlations can also be exploited when two quark flavors \( r \) and \( s \) are reweighted simultaneously. Without loss of generality let the masses \( m_r \leq m_s \) be reweighted by \(-\gamma \Delta m\) and \( \Delta m \), respectively. Defining \( m_{\pm} = (m_s \pm m_r)/2 \) the reweighting factor \( W_{2f} = W_{m_r,m_s - \gamma \Delta m}W_{m_r,m_s + \Delta m} \) can be written as

\[
W_{2f}^{(\gamma)} = \det \left[ I - \Delta m \frac{(\gamma - 1)D_{m_+} + (\gamma + 1)m_- + \gamma \Delta m}{D_{m_+}^2 - m_-^2} \right]. \tag{29}
\]

If the two quarks are degenerate \( m_r = m_s = m \) and thus \( m_+ = m \) and \( m_- = 0 \), three special values of \( \gamma \) occur. For \( \gamma = 0 \) and \( \gamma = -1 \) one flavor reweighting \( W_{2f}^{(0)} = W_{m,m'} \) and two degenerate flavor reweighting \( W_{2f}^{(-1)} = W_{m,m'} \) are recovered, respectively (with \( m' = m + \Delta m \)). The case \( \gamma = 1 \) equals the isospin reweighting \( W_{2f}^{(1)} = W_{\pm} \) below, cf. Eq. \((33)\). The ensemble fluctuations in the three cases at the same \( \Delta m \) fulfill \( \sigma_{2f,-1}^2 > \sigma_{2f,0}^2 > \sigma_{2f,1}^2 \). The first relation is obvious and the second is a consequence of the suppression with the fourth power of \( \Delta m \) in Eq. \((34)\).

Imposing \( \gamma = 1 \) in the non-degenerate case means keeping the sum of the bare quark masses constant, but is not the optimal choice in terms of fluctuations. Instead the value of \( \gamma \) might be fixed by minimizing the ensemble fluctuations of \( W_{2f} \). From the preceding discussion we expect the minimum for \( 0 \leq \gamma^* \leq 1 \), depending on \( m_- \): if \( m_- = 0 \) isospin reweighting and thus \( \gamma = 1 \) is optimal, whereas if \( m_s \to \infty \) (and thus \( m_- \to \infty \)) one flavor reweighting of \( m_s \) and thus \( \gamma = 0 \) is optimal. We write Eq. \((29)\) as

\[
W_{2f}^{(\gamma)} = \exp(w), \quad w = -\Delta m[(\gamma - 1)\langle b \rangle + (\gamma + 1)\langle c \rangle] + O(\Delta m^2), \tag{30}
\]

with \( b = D_{m_+}/(D_{m_+}^2 - m_-^2), c = m_-/(D_{m_+}^2 - m_-^2) \). Minimization of the variance \( \text{var}(w) = \langle w^2 \rangle - \langle w \rangle^2 \) with respect to \( \gamma \) yields

\[
\gamma^* \approx \frac{\text{var}(\text{Tr} b) - \text{var}(\text{Tr} c)}{\text{var}(\text{Tr} b + \text{Tr} c)}. \tag{31}
\]

Approximating, for small \( m_- \), \( \text{var}(\text{Tr} b) \approx \text{var}(\text{Tr} D_{m_+}^{-1}) \approx k_{1f}V \) and \( \text{var}(\text{Tr} c) \approx m_-^2 \text{var}(\text{Tr} D_{m_+}^{-2}) \approx m_-^2 k_{3f}V \) and neglecting the covariance \( \text{cov}(\text{Tr} b, \text{Tr} c) \) we obtain

\[
\gamma^* \approx 1 - 2m_-^2 \frac{k_{1f}}{k_{3f}} + O(\Delta m, m_-^3). \tag{32}
\]

We have determined \( \gamma^* \) directly by minimizing the fluctuations of Eq. \((29)\) for \( m_- \approx \overline{m}_s/2 \) and \( m_+ \approx 5/6\overline{m}_s \). Inserting in Eq. \((32)\) \( m_- = 50\text{MeV} \) \(^3\) and the value of \( k_{1f}/k_{3f} \), independently determined below for \( m_+ \approx \overline{m}_s/3 \), results in

\(^3\) Differences of quark masses renormalize with a factor \( Z_m \) which is \( \approx 1.5 \) in our case \(^{11}\) and we neglect it here.
\( \gamma^* \approx 0.86 \). In the direct determination we find the optimal value \( \gamma^* = 0.82(1) \) in good agreement with the prediction.

We now return to the special case \( m_\pi = 0, \gamma = 1 \). It means reweighting two degenerate quarks with mass \( m = m_r = m_s \) by keeping the sum of the bare masses \( m_r + m_s \) constant. The resulting isospin reweighting factor is

\[
W_\pm = W_{m_r, m_r - \Delta m} W_{m_s, m_r + \Delta m} = \det \left[ 1 - \left( \frac{\Delta m}{D_m} \right)^2 \right].
\] (33)

Following the same calculation which led to Eq. (25), we obtain in this case for the fluctuations

\[
\frac{\sigma_\pm^2}{(W_\pm)^2} = \Delta m^4 \left[ \left( \langle \text{Tr}(D_m^{-2}) \rangle \right)^2 - \left( \langle \text{Tr}(D_m^{-2}) \rangle \right)^2 \right] + O(\Delta m^6),
\] (34)

which shows a suppression with the fourth power of \( \Delta m \), as compared to the second power in Eq. (25). If we use mass interpolation with \( \frac{4}{M - 1} \), the same is true for the relative stochastic error, such that the ratio of the both isospin reweighting to the one flavor reweighting of Section 3.1 we find

\[
\frac{k_\pm}{k_{1f}} \approx 2.5 \times 10^{-4} \left( m'_s - m'_r \right)^2.
\] (35)

the same is true for the relative stochastic error, such that the ratio of the both isospin reweighting to the one flavor reweighting of Section 3.1 we find

\[
\frac{k_\pm}{k_{1f}} \approx 2.5 \times 10^{-4} \left( m'_s - m'_r \right)^2.
\]

4. Tuning of the up-down quark mass difference

In lattice QCD computations the \( n_f \) bare mass parameters are usually related to a minimal number \( n_R \) of dimensionless ratios \( R_j, j = 1, \ldots, n_R \) of observables. The minimal number is given as the number of independent mass parameter: For example, if \( n_f = 2 \) and \( m_1 \equiv m_2 \) (a doublet of degenerate quarks) then \( n_R = 1 \). The bare masses are then tuned such that these ratios take their physical value \( R_{j,\text{phys}} \). In practice this amounts to the generation of several ensembles at different values of the bare masses and extrapolation or interpolation to the physical point. Even if the tuning of \( R_j \) is independent of the tuning of \( R_i \) for \( i \neq j \) the numerical cost grows rapidly with \( n_R \). Reweighting can help to reduce the number of necessary ensembles.

We present two tuning strategies for the parameters \( \kappa_u, \kappa_d \) and \( \kappa_s \) using the kaons \( K^0 \) and \( K^\pm \). More precisely, we complement the strategy of Ref. [11] by even-odd preconditioning of \( D_m \).
to fix $\kappa_u$ at degenerate light quarks (therein called strategy 1) by two different strategies to fix the up-down quark mass difference. The dimensionless ratios to be used are

$$R_1 = \frac{m_{K^0}^2}{f_{K^0}^2}, \quad R_2 = \frac{m_{K^0}^2 - m_{K^\pm}^2}{m_{K^0}^2}, \quad R_3 = \frac{m_{\pi^0}^2}{f_{K^0}^2}.$$  \tag{36}

with physical values $R_{1,\text{phys}}^\text{phys} = (494.2)^2/(155)^2$, $R_{2,\text{phys}}^\text{phys} = (497.2^2 - 491.2^2)/(497.2^2)$ and $R_{3,\text{phys}}^\text{phys} = (134.8)^2/(155)^2$, where QED effects have been removed as discussed in [19]. In the limit $m_u = m_d$, all the kaons are the same. A mass difference $m_d - m_u \neq 0$ breaks isospin and splits the kaon masses $m_{K^0} \neq m_{K^\pm}$.

4.1. One flavor reweighting

In the first strategy we use one flavor reweighting to lower the mass of the up quark $m_u$. The outline of the strategy is:

1. For $\kappa_u = \kappa_d$, tune $\kappa_s$ such that $R_1 \equiv R_{1,\text{phys}}$.
2. Keeping $\kappa_d$ and $\kappa_u$ fixed, tune (increase) $\kappa_u$ such that $R_2 \equiv R_{2,\text{phys}}$.
3. Repeat steps 1. and 2. for different values of $\kappa_d$ and extrapolate to the physical quark mass values, e.g., to the point where $R_3 \equiv R_{3,\text{phys}}$.

Step 1 has been performed in [11] with a quenched strange quark and step 3 is beyond the scope of this paper. Fig. 4 shows the result of step 2 performed on the ensemble D5, cf. Section 3. The parameter $\kappa_s = 0.135777$ is taken from Table 1 (ensemble E5) of [11] and the mass $m_u$ of the up quark is reweighted down to 1/2 of the down quark mass. We also have computed the reweighting factors at three intermediate points, allowing for an interpolation. The ratio $R_2$ of Eq. (36) is shown in Fig. 4 and a linear interpolation (corresponding to leading order of chiral perturbation theory) yields $\kappa_u = 0.1362762(6)$ or $a(m_d - m_u) = 0.00071(2)$. This is approximately 1/4 of the maximal reweighting distance.
\( m/2 \) in Fig. 4 or 4 MeV and amounts to 6\% lower average renormalized quark mass.

For this procedure to work, the ratio \( R_1 \) in Eq. (36) should not significantly change when tuning \( \kappa_u \), i.e., the conditions \( R_1 \equiv R_{1,\text{phys}} \) and \( R_2 \equiv R_{2,\text{phys}} \) should be approximately independent. Indeed, we find the changes in \( R_1 \) to be much smaller than its statistical error. Finally we observe that the numerical result for \( a(m_d - m_u) \) is practically the same when the beta shift is included in the tuning, cf. Fig. 4.

4.2. Isospin reweighting

We modify step 2 of the tuning described in Section 4.1 to change simultaneously \( m_u \) and \( m_d \) by keeping their sum constant. In this case we have to replace \( R_1 \) in Eq. (36) by

\[
R_1' = 0.5 \left( \frac{m^2_{K^0}}{f_{K^0}^2} + \frac{m^2_{K^\pm}}{f_{K^\pm}^2} \right),
\]

which is in leading order of chiral perturbation theory a function of the light quark mass sum \( m_u + m_d \). \( R_1' \) is equal to \( R_1 \) for \( m_u = m_d \) and thus step 1 is unchanged. Using this alternative tuning procedure in step 2 we obtain on the ensemble D5 \( \kappa_u = 0.1362633(3) \) and \( \kappa_d = 0.1362367(3) \) or \( a(m_d - m_u) = 0.00072(1) \).

The normalized reweighting factor \( W_\pm/\langle W_\pm \rangle \) close to the tuned mass difference is very close to one: deviations are \( \lesssim 10^{-4} \). This is mainly due to the suppression of ensemble fluctuations with the fourth power of the mass difference in isospin reweighting (Eq. (34)). Consequently, repeating the tuning with \( W \equiv 1 \) gives practically the same result and we conclude that at this average light quark mass isospin breaking can be treated in the quenched approximation [20].

However, this will most probably change as soon as smaller light quark masses are considered. That is because in the tuning step 2 the small eigenvalues of \( D \) will be shifted close to zero or even cross zero (cf. Section 2.5). At this point the inclusion of the reweighting factor will become indispensable to compensate poles in the propagators.

5. Conclusion

The interest in reweighting in the lattice QCD community has been growing in recent years, but a comprehensive analysis was missing. The main achievement of this paper is to close this gap by:

- Giving a proof for the integral representation of the inverse complex determinant of a complex matrix, valid if the field of values of the matrix is in the right half plane.
• Defining an unbiased stochastic estimator based on mass interpolation and optional $\mu$-detour that has a controlled error and works also in situations where the determinant has a complex phase.

• Expansions of the relative stochastic error and the ensemble fluctuations that are shown to correctly describe the data and lead to optimized re-weighting strategies.

Because of Eq. (27) the numerical cost for keeping the relative stochastic error at a level of 10% of the ensemble fluctuations is independent of the lattice volume and the ensemble quark mass: ca. 50 inversions of the global Dirac operator. The reweighting distance is restricted by the growth of the ensemble fluctuations in Eq. (26), also known as “overlap problem”.

As an example of an application we presented two approaches to the tuning of isospin breaking. The first reweights the up quark to a smaller mass, thus lowering the light quark mass average $(m_u + m_d)/2$. This introduces ensemble fluctuations which can be tamed by a simultaneous beta shift. In the second approach the tuning is done at fixed $m_u + m_d$. This leads to a suppression of the ensemble fluctuations at the price of a larger final average quark mass.

Further possible applications are the tuning of the strange quark mass in simulations with dynamical up, down and strange quark, stabilization of the Hybrid Monte Carlo algorithm, treatment of determinants with phases and the inclusion of QED effects.

Acknowledgments. We thank Martin Lüscher for pointing out a mistake in Eq. (7). The program package which we developed is based on the freely available software package DD-HMC \cite{21, 22, 23} and we thank the ALPHA collaboration for providing the MP-HMC code and the CLS consortium for providing the code for measuring the scale $\tau_0/a^2$. This work was partially funded by Deutsche Forschungsgemeinschaft (DFG) in form of Transregional Collaborative Research Centre 55 (SFB/TRR55). The numerical calculations were carried out on the Cheops supercomputer at the RRZK computing centre of the University of Cologne and on the cluster Stromboli at the University of Wuppertal and we thank both Universities.

Appendix A.

Let $A \in \mathbb{C}^{n \times n}$ with eigenvalues $\lambda(A)$ and $\eta \in \mathbb{C}^n$. Let $f(\eta)$ be a function $f: \mathbb{C}^n \to \mathbb{C}$. We define the integral of $f$ over $\eta$ through

$$\int D[\eta] \; f(\eta) \quad \text{with} \quad D[\eta] = \prod_i \frac{d \text{Re}(\eta_i) \; d \text{Im}(\eta_i)}{\pi}.$$  \hspace{1cm} (A.1)

With this definition $\int D[\eta] \; e^{-\eta^\dagger \eta} = 1$. In order to prove Eq. (7), we first note that the integral over $f(\eta) = \exp(-\eta^\dagger A \eta)$ is defined if and only if the integral over the absolute value $|f(\eta)| = \exp(-\text{Re}(\eta^\dagger A \eta))$ is defined. The condition of absolute convergence of the integral is therefore equivalent to the condition that
the field of values $\mathcal{F}(A)$ defined in Eq. (A.7) is in the right half plane or that $\lambda(A + A^\dagger) > 0$ (see below). Now consider the Schur decomposition of $A$, i.e., $A = QUQ^{-1}$ with unitary matrix $Q$ ($Q^{-1} = Q^\dagger$) and upper triangular matrix $U$. The diagonal entries of $U$ are the eigenvalues $\lambda(A)$. Therefore $U$ can be written as $U = D + K$ with diagonal matrix $D = \text{diag}(\lambda(A))$ and strictly upper triangular matrix $K$. Since $Q$ is unitary the substitution $\eta \rightarrow Q\eta$ in the integral leads to

$$
\int \text{D}[\eta] \ e^{-\eta^\dagger A\eta} = \int \text{D}[\eta] \ e^{-\eta^\dagger (D + K)\eta}.
$$

(A.2)

Now we introduce the two vectors

$$
r = \frac{1}{2}[\eta + (I + D^{-1}K^T)\eta^*] \quad \text{and} \quad s = -\frac{i}{2}[\eta - (I + D^{-1}K^T)\eta^*].
$$

(A.3)

A few lines of algebra show that $\eta^\dagger (D + K)\eta = r^T Dr + s^T Ds$. In terms of the real part $x = \text{Re}(\eta)$ and imaginary part $y = \text{Im}(\eta)$ the vectors $r$ and $s$ read

$$
r = x + \frac{1}{2}D^{-1}K^T(x - iy) \quad \text{and} \quad s = y + \frac{i}{2}D^{-1}K^T(x - iy).
$$

(A.4)

Because $D^{-1}K^T$ is strictly lower triangular, the Jacobian matrix $M$ of the change of variables

$$
\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} r \\ s \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix}, \quad M = \begin{pmatrix} I + \frac{1}{2}D^{-1}K^T & -\frac{i}{2}D^{-1}K^T \\ \frac{i}{2}D^{-1}K^T & I + \frac{1}{2}D^{-1}K^T \end{pmatrix}
$$

(A.5)

has determinant one. Thus we obtain

$$
\int \text{D}[\eta] \ e^{-\eta^\dagger A\eta} = \int \left( \prod_{i=1}^n \frac{dr_i \ ds_i}{\pi} \right) e^{-r^T Dr - s^T Ds} = \prod_{i=1}^n \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_i}} = \frac{1}{\det A}.
$$

(A.6)

The transformations in Eq. (A.4) change the real variables $x = \text{Re}(\eta)$ and $y = \text{Im}(\eta)$ into the complex variables $r$ and $s$. The domain of the integrals over the components of $r$ and $s$ in Eq. (A.6) can be chosen to be along the real axis only if the integral is absolute convergent. The latter property implies that $\text{Re}(\lambda(A)) > 0$ and therefore the existence of the Gaussian integrals in the last step of Eq. (A.6). We thus complete the proof of Eq. (7).

Note that if $A$ is Hermitian the condition $\lambda(A + A^\dagger) > 0$ is equivalent to $A$ being positive definite and thus the Cholesky decomposition $A = LL^\dagger$ exists. In this case the determinant of $A$ can be shown to appear as the Jacobian determinant of $\eta \rightarrow (L^\dagger)^{-1}\eta$.

We close this Appendix by proving inequality Eq. (11). The field of values of $A$ is a subset of the complex plane and defined by

$$
\mathcal{F}(A) = \{\eta^\dagger A\eta : \eta^\dagger \eta = 1\}.
$$

(A.7)

In the proof we will use that (see for example [24])

$$
\mathcal{F}(A) \supset \lambda(A)
$$

(A.8)

$$
\mathcal{F}(A) = [\min(\lambda(A)), \max(\lambda(A))] \Leftrightarrow A = A^\dagger.
$$

(A.9)
Let $d > 0$, then
\[
\lambda(A + A^\dagger) > d \iff \mathcal{F}(A + A^\dagger) > d \iff \eta^\dagger(A + A^\dagger)\eta > d, \quad \forall \eta^\dagger\eta = 1
\]
\[
\iff 2\text{Re}(\eta^\dagger A\eta) > d, \quad \forall \eta^\dagger\eta = 1
\]
\[
\iff 2\text{Re}(\mathcal{F}(A)) > d \iff 2\text{Re}(\lambda(A)) > d
\]
proves Eq. (11) for $d = 1$. \qed

Appendix B.

The stochastic variance of the factors $W_l$ in Section 2.3 diverges if $D_{m_l}$ becomes singular. In order to derive an approximation of the smallest singular value of $D_{m_l,\mu} = D_{m_l} + i\mu\gamma_5$, let us assume without loss of generality that $D_{m_k}$ is singular for a value $k$ in the range $0 < k \leq N - 1$. The smallest singular value $\sigma_{\text{min}}$ of $D_{m_l,\mu} = D_{m_l} + i\mu\gamma_5$ is the square root of the smallest eigenvalue of $D_{m_l,\mu}^\dagger D_{m_l,\mu} = \mu^2 + D_{m_l}^\dagger D_{m_l}$. Using Eq. (C.2) of Appendix C, we arrive at
\[
\sigma_{\text{min}} \approx \sqrt{\mu^2 + (l - k)^2\delta m^2\chi_0^2}. \tag{B.1}
\]
In practice, the so called chirality (cf. Eq. (C.1)) $\chi_0$ of the zero mode of $\gamma_5 D_{m_k}$ is very close to plus or minus one and we assume here $\chi_0 = 1$.

The stochastic variance is bounded from above if $\sigma_{\text{min}} \geq r$ for some $r > 0$. This is achieved by a detour in the mass-$\mu$ plane. Instead of going from $m$ to $m'$ on a straight line, a half circle parametrized by $\mu_j^2 + (\tilde{m}_j - m_k)^2 = r^2$, $j = 0, \ldots, N_c - 1$ is inserted around $m_k$. The starting point is then given for $\mu_0 = 0$ and $\tilde{m}_0 = m_{l_k}$ with $l_k = k - r/|\delta m|$ and can be fixed by restricting the increase of the variance, say by a factor compared to the variance far away from $k$. Equidistant steps along the half circle are obtained if the reweighting distances $\delta m_c = \tilde{m}_{j+1} - \tilde{m}_j$ and $\delta \mu_c = \mu_{j+1} - \mu_j$ fulfill $\delta m^2_c + \delta \mu^2_c = \delta m^2$ in each step. We note that for a constant $\chi_0^2 < 1$ the circle becomes an ellipse with the staring point given by $l_s = k - r/|\delta m\chi_0|$. Finally, we remark that the detour in the mass-$\mu$ plane can be performed in the same way with the even-odd (or any domain decomposition) Schur complement, i.e., by considering $\tilde{D}_{m_l,\mu} = I - (D_{ee} + m_l + i\mu\gamma_5)^{-1}D_{eo}(D_{oo} + m_l + i\mu\gamma_5)^{-1}D_{oe}$. If $\tilde{D}_{m_l,\mu}$ is assured to be non-singular, the same holds for $\tilde{D}_{m_l,\mu}$.

Appendix C.

Let $D$ be a lattice Dirac operator with the property $\gamma_5D\gamma_5 = D^\dagger$. Then $Q(m) = \gamma_5(D + m)$ is Hermitian with real eigenvalues $\lambda_i(m)$ and corresponding orthonormal eigenvectors $v_i(m)$. The first derivative of the spectral flow is given by $d\lambda_i/dm = \chi_i$ with the so called chirality of the eigenvector
\[
\chi_i \equiv v_i^\dagger\gamma_5v_i. \tag{C.1}
\]
A Taylor expansion of the spectral flow around \( m_0 \) at first order is then given by
\[
\lambda_i(m) \approx \lambda_i(m_0) + (m - m_0) \chi_i(m_0) \quad [25, 26].
\]
Let us assume without loss of generality that \( Q(m) \) has a zero eigenvalue for \( m = m_0 \), i.e., \( \min_i |\lambda_i(m_0)| = 0 \). Assuming the order of the eigenvalues not to change in the range of interest, the smallest eigenvalue of \( (D + m)^\dagger (D + m) = Q^2 \) is
\[
\lambda_{\min} = (\min_i |\lambda_i(m)|)^2 \approx (m - m_0)^2 \chi_0^2 , \quad (C.2)
\]
where \( \chi_0 \) is the chirality of the zero mode of \( Q(m_0) \).

References

[1] A. Ferrenberg, R. Swendsen, New Monte Carlo Technique for Studying Phase Transitions, Phys.Rev.Lett. 61 (1988) 2635–2638. doi:10.1103/PhysRevLett.61.2635.

[2] A. Hasenfratz, R. Hoffmann, S. Schaefer, Reweighting toward s the chiral limit, Phys.Rev. D78 (2008) 014515. arXiv:0805.2369, doi:10.1103/PhysRevD.78.014515.

[3] M. Lüscher, S. Schaefer, Lattice QCD with open boundary conditions and twisted-mass reweighting, Comput.Phys.Commun. 184 (2013) 519–528. arXiv:1206.2809, doi:10.1016/j.cpc.2012.10.003.

[4] S. Aoki, K. Ishikawa, N. Ishizuka, K. Kanaya, Y. Kuramashi, et al., 1+1+1 flavor QCD + QED simulation at the physical point, Phys.Rev. D86 (2012) 034507. arXiv:1205.2961, doi:10.1103/PhysRevD.86.034507.

[5] Y. Aoki, et al., Continuum Limit Physics from 2+1 Flavor Domain Wall QCD, Phys.Rev. D83 (2011) 074508. arXiv:1011.0892, doi:10.1103/PhysRevD.83.074508.

[6] S. Aoki, et al., Physical Point Simulation in 2+1 Flavor Lattice QCD, Phys.Rev. D81 (2010) 074503. arXiv:0911.2561, doi:10.1103/PhysRevD.81.074503.

[7] A. Hasenfratz, A. Alexandru, Evaluating the fermionic determinant of dynamical configurations, Phys.Rev. D65 (2002) 114506. arXiv:hep-lat/0203026, doi:10.1103/PhysRevD.65.114506.

[8] M. Hasenbusch, Speeding up the hybrid Monte Carlo algorithm for dynamical fermions, Phys.Lett. B519 (2001) 177–182. arXiv:hep-lat/0107019, doi:10.1016/S0370-2693(01)01102-9.

[9] J. Finkenrath, F. Knechtli, B. Leder, Fermions as Global Correction: the QCD Case, Comput.Phys.Commun. 184 (2013) 1522–1534. arXiv:1204.1305, doi:10.1016/j.cpc.2013.01.020.
[10] J. Finkenrath, F. Knechtli, B. Leder, Application of Domain Decomposition to the Evaluation of Fermion Determinant Ratios, PoS LATTICE2012 (2012) 190. [arXiv:1211.1214]

[11] P. Fritzsch, F. Knechtli, B. Leder, M. Marinkovic, S. Schaefer, et al., The strange quark mass and Lambda parameter of two flavor QCD, Nucl.Phys. B865 (2012) 397–429. [arXiv:1205.5380] doi:10.1016/j.nuclphysb.2012.07.026.

[12] 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. [arXiv:hep-lat/0701009] doi:10.1088/1126-6708/2007/02/082.

[13] M. Marinkovic, S. Schaefer, Comparison of the mass preconditioned HMC and the DD-HMC algorithm for two-flavour QCD, PoS LATTICE2010 (2010) 031. [arXiv:1011.0911]

[14] W. A. Bardeen, A. Duncan, E. Eichten, G. Hockney, H. Thacker, Light quarks, zero modes, and exceptional configurations, Phys.Rev. D57 (1998) 1633–1641. [arXiv:hep-lat/9705008] doi:10.1103/PhysRevD.57.1633.

[15] T. A. DeGrand, A. Hasenfratz, T. G. Kovacs, Instantons and exceptional configurations with the clover action, Nucl.Phys. B547 (1999) 259–280. [arXiv:hep-lat/9810061] doi:10.1016/S0550-3213(99)00126-1.

[16] U. Wolff, Monte Carlo errors with less errors, Comput.Phys.Commun. 156 (2004) 143–153. [arXiv:hep-lat/0306017] doi:10.1016/S0010-4655(03)00467-3,10.1016/j.cpc.2006.12.001.

[17] K. Jansen, R. Sommer, O(α) improvement of lattice QCD with two flavors of Wilson quarks, Nucl.Phys. B530 (1998) 185–203. [arXiv:hep-lat/9803017] doi:10.1016/S0550-3213(98)00396-4.

[18] M. Lüscher, Properties and uses of the Wilson flow in lattice QCD, JHEP 1008 (2010) 071. [arXiv:1005.4518] doi:10.1007/JHEP08(2010)071.

[19] G. Colangelo, S. Dürr, A. Jüttner, L. Lellouch, H. Leutwyler, et al., Review of lattice results concerning low energy particle physics, Eur.Phys.J. C71 (2011) 1695. [arXiv:1011.4408] doi:10.1140/epjc/s10052-011-1695-1.

[20] G. de Divitiis, R. Frezzotti, V. Lubicz, G. Martinelli, R. Petronzio, et al., Leading isospin breaking effects on the lattice [arXiv:1303.4896]

[21] M. Lüscher, Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD, Comput.Phys.Commun. 165 (2005) 199–220. [arXiv:hep-lat/0409106] doi:10.1016/j.cpc.2004.10.004.

[22] M. Lüscher, Deflation acceleration of lattice QCD simulations, JHEP 0712 (2007) 011. [arXiv:0710.5417] doi:10.1088/1126-6708/2007/12/011.
[23] M. Lüscher, DD-HMC algorithm for two-flavour lattice QCD, http://luscher.web.cern.ch/luscher/DD-HMC/index.html.

[24] W. V. Parker, Characteristic roots and the field of values of a matrix, Duke Math. J. Volume 15, Number 2 (1948), 439-442.

[25] R. G. Edwards, U. M. Heller, R. Narayanan, The Hermitian Wilson-Dirac operator in smooth SU(2) instanton backgrounds, Nucl.Phys. B522 (1998) 285–297. arXiv:hep-lat/9801015, doi:10.1016/S0550-3213(98)00270-3

[26] H. Simma, D. Smith, Low lying eigenvalues of the improved Wilson-Dirac operator in QCD arXiv:hep-lat/9801025.