Charm and strange quark in openQCD simulations

Martin Lüscher February 2012; last revised November 2019

1. Introduction

The inclusion of the charm and the strange quark in lattice QCD simulations is not completely trivial. In the `openQCD` package, a version of the RHMC algorithm [1,2] is used for these quarks. The factorization of the quark determinant, the associated pseudo-fermion actions and the computation of the forces deriving from them are briefly described in this note along with further implementation details.

The RHMC algorithm involves a rational approximation and requires reweighting if the approximation error is not completely negligible. Reweighting is discussed here too as well as a strategy of how to proceed in the case of master-field simulations [3], where reweighting is not possible.

2. Quark determinant

The discussion in this section roughly follows the lines of sect. 6.2.6 of ref. [4], which should be consulted for further explanations. Since the charm and the strange quark are treated in the same way, it suffices to consider the latter.

2.1 Factorization of the strange-quark determinant

Let D be the massive O(a)-improved Wilson–Dirac operator with bare mass parameter m_0 set to the bare mass of the strange quark (see ref. [5] for the exact definition of D). The RHMC algorithm implemented in the `openQCD` package makes use of even-odd preconditioning and thus starts from the decomposition

$$\det D = \det(1_{\rm e} + D_{\rm oo}) \det \hat{D} \qquad (2.1)$$

of the strange-quark determinant. In this equation, $\hat{D}$ denotes the even-odd preconditioned Dirac operator, $1_{\rm e}$ the projector to the subspace of quark fields that vanish on the odd sites of the lattice and $D_{\rm oo}$ the odd-odd part of the Dirac operator. As explained in sect. 4 of ref. [6], the first factor in eq. (2.1) can be directly included in the molecular-dynamics Hamilton function.

The other factor is then further decomposed according to

$$\det \hat{D} = W \det R^{-1}, \tag{2.2}$$

where the operator R is a suitable rational approximation to $(\hat{D}^\dagger \hat{D})^{-1/2}$ while the residual factor,

$$W = \det(\hat{D}R), \tag{2.3}$$

is treated as a reweighting factor.

2.2 Zolotarev rational approximation

The Zolotarev rational function

$$R_{n,\epsilon}(y) = A \frac{(y+a_1)(y+a_3)\ldots(y+a_{2n-1})}{(y+a_2)(y+a_4)\ldots(y+a_{2n})} \tag{2.4}$$

of degree $[n,n]$ approximates $1/\sqrt{y}$ in the range $\epsilon \le y \le 1$ with the smallest possible relative deviation

$$\delta = \max_{\epsilon \le y \le 1} |1 - \sqrt{y} R_{n,\epsilon}(y)| \,. \tag{2.5}$$

Somewhat surprisingly, the coefficients $a_1, \ldots, a_{2n}$ of this optimal rational function, the proportionality constant A and the approximation error δ can be determined analytically (see ref. [7]; the results derived there are reproduced in appendix A).

Since the strange quark has a fairly large mass, the eigenvalues of the operator

$$(\hat{D}^\dagger \hat{D})^{1/2} = |\gamma_5 \hat{D}| \tag{2.6}$$

are expected to be separated from zero by a safe spectral gap. Once the simulation has thermalized, a spectral range $[r_a, r_b]$, $0 < r_a < r_b$, can thus be found, which, with probability practically equal to 1, contains all eigenvalues.

In the `openQCD` package, the operator R in eqs. (2.2),(2.3) is taken to be

$$R = r_b^{-1} R_{n,\epsilon}(r_b^{-2} \hat{D}^\dagger \hat{D}), \qquad \epsilon = (r_a/r_b)^2. \tag{2.7}$$

With this choice, the norm bound

$$\|1 - |\gamma_5 \hat{D}| R\| \leq \delta \tag{2.8}$$

holds when the spectrum of $|\gamma_5 \hat{D}|$ is contained in the range $[r_a, r_b]$, i.e. with high probability in a representative ensemble of gauge fields.

2.3 Further factorizations

The Zolotarev rational function (2.4) may be broken up into two or more factors of the form

$$P_{k,l} = \prod_{j=k}^{l} \frac{y + a_{2j-1}}{y + a_{2j}}. \tag{2.9}$$

If $n = 12$, for example, a possible factorization is

$$R_{n,\epsilon} = A P_{1,5} P_{6,9} P_{10,12}. \tag{2.10}$$

Substituting $y = r_b^{-2} \hat{D}^\dagger \hat{D}$ as before, the associated decomposition

$$\det R^{-1} = \text{constant} \times \det\{P_{1,5}^{-1}\} \det\{P_{6,9}^{-1}\} \det\{P_{10,12}^{-1}\} \tag{2.11}$$

of the second factor in eq. (2.2) effectively achieves a frequency splitting of the quark determinant, because the coefficients $a_1, \ldots, a_{2n}$ are monotonically decreasing,

$$a_1 > a_2 > \ldots > a_{2n} > 0, \tag{2.12}$$

and range over the whole spectral interval from 1 down to ϵ.

3. Pseudo-fermion action and strange-quark force

In the following, $P_{k,l}$ denotes the product (2.9) with y set to $r_b^{-2} \hat{D}^\dagger \hat{D}$. Explicitly, the product is given by

$$P_{k,l} = \prod_{j=k}^{l} \frac{\hat{D}^\dagger \hat{D} + \nu_j^2}{\hat{D}^\dagger \hat{D} + \mu_j^2}, \tag{3.1}$$

where the parameters

$$\mu_k = r_b(a_{2k})^{1/2}, \qquad \nu_k = r_b(a_{2k-1})^{1/2}, \qquad k = 1, 2, \ldots, n, \tag{3.2}$$

will be referred to as "twisted masses".

3.1 Actions and fields

Since the operators $P_{k,l}$ are Hermitian and strictly positive, the determinants appearing in a decomposition such as (2.11) can be taken into account in the simulations by including the pseudo-fermion actions

$$S_{\mathrm{pf},k,l} = (\phi_{k,l}, P_{k,l}\phi_{k,l}) \tag{3.3}$$

in the molecular-dynamics Hamilton function. The fields $\phi_{k,l}$ in this expression are independent pseudo-fermion fields that live on the even sites of the lattice.

Apart from the fact that the product (3.1) may have more than one factor, the pseudo-fermion actions (3.3) are very similar to the actions discussed in sect. 4 of ref. [6]. The partial fraction decomposition

$$P_{k,l} = 1 + \sum_{j=k}^{l} \frac{\rho_{k,l,j}}{\hat{D}^\dagger\hat{D} + \mu_j^2}, \tag{3.4}$$

$$\rho_{k,l,j} = (\nu_j^2 - \mu_j^2) \prod_{m=k, m\neq j}^{l} \frac{\nu_m^2 - \mu_j^2}{\mu_m^2 - \mu_j^2}, \tag{3.5}$$

actually shows that the actions (3.3) are sums of the actions previously considered.

3.2 Forces

The force

$$F_{k,l}^a(x,\mu) = \partial_{x,\mu}^a S_{\mathrm{pf},k,l} \tag{3.6}$$

can therefore be computed following the lines of ref. [6]. In the course of this calculation, the fields

$$\chi_{k,l,j} = (\hat{D}^\dagger\hat{D} + \mu_j^2)^{-1}\phi_{k,l} \tag{3.7}$$

must be computed, which requires the normal even-odd preconditioned Dirac equation to be solved for $j = k, \ldots, l$ and thus possibly many times.

Since the source field $\phi_{k,l}$ is the same for all j, the multi-shift conjugate gradient algorithm [8,9] can be used for the simultaneous solution of the equations. This works well as long as the masses $\mu_k, \ldots, \mu_l$ are not too small. Highly optimized single-shift solvers may otherwise prove to be more efficient. The `openQCD` package includes several solvers and one can choose the solver to be used for each factor $P_{k,l}$ separately.

3.3 Pseudo-fermion field generation

At the beginning of the molecular-dynamics evolution, the pseudo-fermion field $\phi_{k,l}$ is chosen randomly with the proper distribution (or, in the case of the SMD algorithm, gets rotated in direction of such a random field).

A moment of thought shows that this is achieved by setting

$$\phi_{k,l} = A_{k,l}\eta_{k,l}, \qquad A_{k,l} = \prod_{j=k}^{l} \frac{\gamma_5\hat{D} + i\mu_j}{\gamma_5\hat{D} + i\nu_j}, \tag{3.8}$$

where $\eta_{k,l}$ is a random field on the even sites of the lattice with normal distribution. Since

$$A_{k,l} = 1 + i\sum_{j=k}^{l} \frac{\sigma_{k,l,j}}{\gamma_5\hat{D} + i\nu_j}, \qquad \sigma_{k,l,j} = (\mu_j - \nu_j)\prod_{m=k,m\neq j}^{l} \frac{\mu_m - \nu_j}{\nu_m - \nu_j}, \tag{3.9}$$

the application of $A_{k,l}$ to the source field $\eta_{k,l}$ amounts to solving the Dirac equation $l-k+1$ times. Again the multi-shift CG solver can be used here for the simultaneous solution of these equations, but in the case of the few smallest masses ν_j the use of a highly efficient single-shift solver may be preferable.

4. Stochastic estimation of the reweighting factor *W*

The Hermiticity properties of the lattice Dirac operator guarantee that the reweighting factor (2.3) is real, but the factor may, in principle, change sign from one gauge-field configuration to the next. Sign changes are however practically excluded when the quark mass is set to values as large as the physical strange-quark mass (see ref. [4] for a more extensive discussion of the issue). In the following, the reweighting factor is therefore assumed to be positive.

4.1 Stochastic estimator

Let $\eta_j(x)$, $j = 1, \ldots, N$, be a set of independent random quark fields on the even lattice sites with normal distribution. As in the case of the light-quark reweighting factors discussed in ref. [10], a stochastic estimator for W is given by

$$W_N = \frac{1}{N} \sum_{j=1}^{N} \exp\{-(\eta_j, [(1+Z)^{-1/2} - 1]\eta_j)\}, \tag{4.1}$$

where

$$Z = \hat{D}^\dagger \hat{D} R^2 - 1. \tag{4.2}$$

Recalling the bound (2.8), the inequality

$$\|Z\| \leq \delta(2+\delta) \tag{4.3}$$

is easily established, and since the approximation error δ is, in practice, much smaller than 1, the inverse square root of $1+Z$ in eq. (4.1) is well defined.

4.2 Power series expansion

Actually, the series

$$(1+Z)^{-1/2} = 1 - \tfrac{1}{2}Z + \tfrac{3}{8}Z^2 - \tfrac{5}{16}Z^3 + \tfrac{35}{128}Z^4 - \ldots \tag{4.4}$$

is rapidly convergent in the operator norm. The exponents in eq. (4.1) can therefore be computed by evaluating the first few terms in the expansion

$$(\eta_j, [(1+Z)^{-1/2} - 1]\eta_j) = -\tfrac{1}{2}(\eta_j, Z\eta_j) + \tfrac{3}{8}(\eta_j, Z^2\eta_j) - \ldots \tag{4.5}$$

It is possible to estimate the size of these terms by noting that $\|\eta_j\|^2$ is very nearly equal to 12 times the number $N_{\rm e}$ of even lattice points. Taking the bound (4.3) into account, the matrix element $(\eta_j, Z^p\eta_j)$ is thus expected to be less than $12N_{\rm e}(2\delta)^p$.

4.3 Statistical fluctuations

The statistical fluctuations of the exponents in eq. (4.1) derive from those of the gauge field and those of the random sources η_j. For a given gauge field, the variance of the exponents is equal to

$$\mathrm{Tr}\{[(1+Z)^{-1/2} - 1]^2\} = \tfrac{1}{4}\mathrm{Tr}\{Z^2\} - \tfrac{3}{8}\mathrm{Tr}\{Z^3\} + \ldots \tag{4.6}$$

Since the traces $\mathrm{Tr}\{Z^p\}$ are at most $12N_\mathrm{e}(2\delta)^p$, these fluctuations are guaranteed to be small if, say, $12N_\mathrm{e}\delta^2 \leq 10^{-4}$, and one can then just as well set $N = 1$ in eq. (4.1).

Once the statistical fluctuations of W_N at fixed gauge field are small, the variance of W_N in the simulation is practically the same as the one of W. Noting

$$W = \det(1 + Z)^{1/2} = \mathrm{e}^F, \tag{4.7}$$

$$F = \tfrac{1}{2}\mathrm{Tr}\{\ln(1 + Z)\} = \mathrm{Tr}\bigl\{\tfrac{1}{2}Z - \tfrac{1}{4}Z^2 + \tfrac{1}{6}Z^3 - \ldots\bigr\}, \tag{4.8}$$

the moment-cumulant transformation implies

$$\langle W\rangle = \exp\bigl\{\langle F\rangle + \tfrac{1}{2}\langle F^2\rangle_c + \tfrac{1}{6}\langle F^3\rangle_c + \ldots\bigr\}, \tag{4.9}$$

where $\langle \ldots\rangle_c$ stands for the connected expectation value, e.g. $\langle F^2\rangle_c = \langle F^2\rangle - \langle F\rangle^2$. For the normalized variance of W, the expression

$$\frac{\langle W^2\rangle}{\langle W\rangle^2} - 1 = \exp\bigl\{\langle F^2\rangle_c + \langle F^3\rangle_c + \ldots\bigr\} - 1 \tag{4.10}$$

is then obtained.

In position space the operator Z and its dependence on the gauge field are quasi-local with a localization range roughly equal to the Compton wavelength of the pseudo-scalar meson composed of two valence strange quarks (about 0.2 fm at the physical point). One can show this by expanding Z in partial fractions and noting that the position-space kernels of the terms coincide with pseudo-scalar meson propagators. As a consequence, the connected expectation values of F^p scale proportionally to $N_\mathrm{e}\delta^p$ on large lattices rather than $(N_\mathrm{e}\delta)^p$. The variances of the normalized reweighting factors W and its stochastic approximation W_N are thus both of order $N_\mathrm{e}\delta^2$ in the large-volume regime of the theory.

5. Master-field simulations with heavy quarks

Since reweighting is not possible in master-field simulations, the reweighting factor W must be such that the deviations

$$d_W(\mathcal{O}) = \frac{\langle \mathcal{O}W\rangle}{\langle W\rangle} - \langle \mathcal{O}\rangle \tag{5.1}$$

of the reweighted from the unreweighted expectation values of the observables $\mathcal{O}$ of interest are negligible with respect to the statistical errors. Here and below the expectation values $\langle\ldots\rangle$ are the ones defined in the theory with the approximate heavy-quark actions. Noting

$$d_W(\mathcal{O}) = \frac{\big\langle(\mathcal{O}-\langle\mathcal{O}\rangle)(W-\langle W\rangle)\big\rangle}{\langle W\rangle}, \tag{5.2}$$

the application of the Cauchy–Schwarz inequality

$$|\langle AB\rangle| \leq \langle|A|^2\rangle^{1/2}\langle|B|^2\rangle^{1/2} \tag{5.3}$$

leads to the bound

$$|d_W(\mathcal{O})| \leq \sigma(\mathcal{O})\frac{\sigma(W)}{\langle W\rangle}, \tag{5.4}$$

where $\sigma(\mathcal{O})$ denotes the standard deviation of the observable $\mathcal{O}$.

In master-field simulations, the observables $\mathcal{O}$ are translation averages of local (or quasi-local) fields and the statistical errors coincide with $\sigma(\mathcal{O})$. Sometimes the observables are additionally averaged over a small ensemble of N_m master fields, in which case the errors are reduced by the factor $1/\sqrt{N_m}$ if the fields are statistically independent and otherwise by less than that. The bound (5.4) then implies that master-field simulations obtain the correct results up to the estimated statistical errors if, say,

$$\frac{\sigma(W)}{\langle W\rangle} \leq \frac{0.1}{\sqrt{N_m}}. \tag{5.5}$$

In order to guarantee that this condition is met, the standard deviation of W must be measured on smaller lattices and the approximation error δ must then be chosen appropriately based on the scaling law $\sigma(W) \propto \delta\sqrt{N_{\rm e}}$ (cf. discussion at the end of sect. 4).

6. Parameter tuning

When a new simulation is started, the required approximation error δ and the appropriate spectral range $[r_a, r_b]$ may not be known. Reasonable initial choices of δ are such that $12N_e\delta^2 \simeq 10^{-4}$, while for the spectral interval one may take $[\frac{1}{2}am, 8]$, for example, where m is an estimate of the bare current-quark mass of the quark considered.

In the course of the thermalization phase, the parameters will then need to be adjusted by calculating the reweighting factor W, following the lines of sect. 4, and the true spectral range of $|\gamma_5\hat{D}|$ for a subset of the generated gauge field configurations. The `openQCD` package includes two main programs, `main/ms1.c` and `main/ms2.c`, that can be used for this purpose.

The computer time required for the simulation increases with the degree of the Zolotarev rational function. A compromise thus needs to be found, where the number of poles is as small as possible while the fluctuations of the reweighting factor remain tolerable. Compromises should however not be made in the case of the spectral range, since the correctness of the simulation may otherwise be difficult to guarantee. Adding a safety margin of 10% to the ends of the measured spectral range is therefore recommended.

Appendix A

The analytic expressions for the coefficients of the rational function (2.4) that minimizes the approximation error (2.5) involve the Jacobi elliptic functions $\mathrm{sn}(u,k)$, $\mathrm{cn}(u,k)$ and the complete elliptic integral $K(k)$ (see ref. [11], for example, for the definition of these functions). Explicitly, they are given by

$$a_r = \frac{\mathrm{cn}^2(rv,k)}{\mathrm{sn}^2(rv,k)}, \quad r = 1, 2, \ldots, 2n, \tag{A.1}$$

where

$$k = \sqrt{1-\epsilon}, \qquad v = \frac{K(k)}{2n+1}. \tag{A.2}$$

The formulae for the amplitude A and the error δ,

$$A = \frac{2}{1+\sqrt{1-d^2}} \frac{c_1 c_3 \ldots c_{2n-1}}{c_2 c_4 \ldots c_{2n}}, \tag{A.3}$$

$$\delta = \frac{d^2}{\left(1+\sqrt{1-d^2}\right)^2}, \tag{A.4}$$

involve the coefficients

$$c_r = \mathrm{sn}^2(rv, k), \quad r = 1, 2, \ldots, 2n, \tag{A.5}$$

$$d = k^{2n+1} \left(c_1 c_3 \ldots c_{2n-1}\right)^2. \tag{A.6}$$

All these expressions are free of singularities and can be programmed using well-known methods for the numerical evaluation of the Jacobi elliptic functions.

References

[1] 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

[2] M. A. Clark, A. D. Kennedy, Accelerating dynamical fermion computations using the Rational Hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields, Phys. Rev. Lett. 98 (2007) 051601

[3] M. Lüscher, *Stochastic locality and master-field simulations of very large lattices*, EPJ Web Conf. 175 (2018) 01002

[4] M. Lüscher, *Computational strategies in lattice QCD*, in: *Modern perspectives in lattice QCD*, eds. L. Lellouch et al. (Oxford University Press, New York, 2011) [arXiv: 1002.4232]

[5] M. Lüscher, *Implementation of the lattice Dirac operator*, `doc/dirac.pdf`

[6] M. Lüscher, *Molecular-dynamics quark forces*, `doc/forces.pdf`

[7] N. I. Achiezer, *Theory of Approximation* (Dover Publications, New York, 1992)

[8] B. Jegerlehner, *Krylov space solvers for shifted linear systems*, preprint IUHET-353 (1996) [arXiv: hep-lat/9612014]

[9] M. Lüscher, *Multi-shift conjugate gradient algorithm*, `doc/mscg.pdf`

[10] M. Lüscher, F. Palombi, *Fluctuations and reweighting of the quark determinant on large lattices*, PoS (LATTICE2008) 049

[11] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions* (Dover Publications, New York, 1972)