One-flavor algorithm for Wilson and domain-wall fermions

Kenji Ogawa, Ting-Wai Chiu, Tung-Han Hsieh (for the TWQCD Collaboration)

Physics Department, and Center for Theoretical Sciences, National Taiwan University, Taipei 10617, Taiwan
Center for Quantum Science and Engineering, National Taiwan University, Taipei 10617, Taiwan
Research Center for Applied Sciences, Academia Sinica, Taipei 115, Taiwan

We construct positive-definite pseudofermion actions for one fermion flavor in lattice field theory, for Wilson and domain-wall fermions respectively. The positive definiteness of these actions ensures that they can be simulated with the Hybrid Monte Carlo (HMC) method. For lattice QCD with optimal domain-wall quarks, we compare the efficiency of HMC simulations of 2-flavor and \((1 + 1)\)-flavor, and find that the efficiency ratio is about 3 : 2.
One-flavor algorithm for Wilson and domain-wall fermions

Kenji Ogawa

1. Introduction

In Monte Carlo simulations of unquenched lattice QCD, a positive-definite fermion determinant is required to ensure the convergence of the Markov chain. For Hybrid Monte Carlo (HMC) simulation \[1\], it also requires that the pseudofermion action is positive-definite such that the pseudofermion field can be simulated with a Gaussian noise. For QCD with two degenerate flavors, this is guaranteed provided that \( \det(D^\dagger D) = \det(D) \). The challenge of simulating one fermion flavor with HMC is that it is nontrivial to write down a positive-definite pseudofermion action with determinant exactly equal to \( \det(D) \). So far, there are several strategies to simulate one fermion flavor with HMC, e.g., to approximate \( D \) by a polynomial of even degree \[2\], or to approximate \( D^\dagger D \) by a rational polynomial function \[3\]. In this paper, we construct a positive-definite pseudofermion action with determinant exactly equal to \( \det(D) \), for Wilson and domain-wall fermions respectively.

The outline of this paper is as follows. In Sec. 2, we construct the pseudofermion action for one-flavor Wilson fermion, as well as that with the Hasenbusch acceleration \[4\]. In Sec. 3, we construct the pseudofermion action for one-flavor domain-wall fermion. In Sec. 4, we construct the pseudofermion action for one-flavor optimal domain-wall fermion. In Sec. 5, we compare the HMC simulations of 2 degenerate flavors with that of one flavor, for lattice QCD with optimal domain-wall fermion. In Sec. 6, we conclude with some remarks.

2. Wilson fermion

In this section, we derive a positive-definite pseudofermion action which gives exactly the determinant of the Wilson-Dirac operator

\[ D_W = W + m + \sum_\mu t_\mu \gamma_\mu = (W + m) 1 \sum_\mu t_\mu \sigma_\mu \sigma_\mu^\dagger (W + m) 1 \; ; \]

where

\[ W = \frac{1}{2} \sum_\mu U_\mu(x) \delta_{x, \mu} + U_\mu^\dagger(x) U_\mu(x) \delta_{x, \mu} + 4 \; ; \]

\[ t_\mu = \frac{1}{2} \left( U_\mu(x) \delta_{x, \mu} \right) U_\mu^\dagger(x) \delta_{x, \mu} + 4 \; ; \]

\[ \gamma_\mu = \begin{pmatrix} 0 & \sigma_\mu^\dagger \\ \sigma_\mu & 0 \end{pmatrix} \; ; \sigma_\mu = (i1_\sigma) ; \]

and \( \sigma_i \) \((i = 1, 2, 3)\) are Pauli matrices. Using the Schur decomposition, the determinant of (2.1) can be written as

\[ \det D_W = \det(W + m)^2 \det W_H ; \]

where \( W_H \) is the Schur complement of \( D_W \), i.e.,

\[ W_H = W + m \sum_\mu t_\mu \frac{1}{W + m} t_\nu \sigma_\mu^\dagger \sigma_\nu ; \]

The positive definiteness of \( W_H \) is asserted as follows. For any background gauge field, the eigenvalues of \( W \) and \( (\sigma t) \) satisfy the inequalities: \( 0 \lambda(W) 8 \), and \( 4 \sum_\mu \prod_i \lambda(W) \). It
follows that \( W_H \) is positive-definite for \( m > 4 \). Now, if \( m \) is decreased from 4 to a smaller value, then the smallest eigenvalue of \( (W + m) \) is also decreased. Suppose that at \( m = m_{cr} \) (the largest value), it becomes zero, then \( \det (W + m_{cr}) = 0 \), and \( W_H (m_{cr}) \) becomes singular, thus \( W_H \) is no longer positive-definite. From (2.2), we see that whenever \( W_H (m) \) has a zero eigenvalue, \( D_W (m) \) also has a zero eigenvalue. On the other hand, when \( W_H (m) \) is singular due to a zero eigenvalue of \( (W + m) \), \( D_W (m) \) does not necessarily have a zero eigenvalue, which follows from the inequality, \( \lambda_{min} (W + m) \quad Re(\lambda (D_W (m))) \quad \lambda_{max} (W + m) \). In other words, the largest \( m \) (i.e., \( m_{cr} \) which gives \( \det (W + m) = 0 \) is equal to or larger than the largest \( m \) which gives \( \det D_W (m) = 0 \).

Thus the pseudofermion action for one-flavor Wilson fermion can be written as

\[
S_{PF} = \Phi_1^\dagger (W + m)^2 \Phi_1 + \Phi_2^\dagger (W_H) \Phi_2 = \Phi_1^\dagger (W + m)^2 \Phi_1 \quad 0 \quad \Phi_2^\dagger \quad H_W^{-1} \quad 0 \quad \Phi_2 \tag{2.4}
\]

where \( H_W = \gamma_5 D_W \), \( \Phi_1 \) is a pseudofermion field without Dirac index, and \( \Phi_2 \) is a pseudofermion field with 2 spinor components.

To generate the pseudofermion field \( \Phi_2 \) from a Gaussian random noise field \( \Xi_2 \), we need to take the square root of \( W_H \), i.e., \( \Phi_2 = \sqrt{W_H} \Xi_2 \). Here we use the Zolotarev optimal rational approximation for the square root,

\[
\Phi_2 = \sqrt{W_H} \Xi_2 \quad p_0 + \sum_{l=1}^{N_{app}} p_l + W + m \quad \sum_{\mu, \nu} t_{\mu, \nu} \frac{1}{m_{cr} - m} \sigma_\mu \sigma_\nu \quad \Xi_2 \tag{2.5}
\]

where \( p_0 \), \( p_l \) and \( q_l \) are expressed in terms of Jacobian elliptic functions. At first sight, the operations in (2.5) look formidable. However, since \( W_H \) is the Schur complement of \( D_W \), each term in (2.5) can be obtained by the inversion of \( (D_W + q_l P) \), i.e.,

\[
0 \quad \Phi_2 = P \quad p_0 + \sum_{l=1}^{N_{app}} p_l (D_W + q_l P) \quad 0 \quad \Xi_2 \tag{2.6}
\]

Note that one cannot apply the multi-shift conjugate gradient algorithm in (2.6), since \( P \) does not commute with \( D_W \). However, these \( N_{app} \) number of inversions can be speeded up using chronological inversion method [5].

### 2.1 Wilson fermion with Hasenbusch acceleration

The idea of Hasenbusch’s method [6] is to introduce a heavy pseudofermion field (with mass \( M > m \)) such that the speed of simulating both \( \det (D_W (m) = D_W (M)) \) and \( \det (D_W (M)) \) is faster than that of \( \det (D_W (m)) \). To apply Hasenbusch’s method, we consider

\[
\det \frac{D_W (m_1)}{D_W (m_2)} = \det \frac{W + m_1}{W + m_2}^2 \det \frac{W_H (m_1)}{W_H (m_2)} \tag{2.7}
\]

where

\[
\bar{W}_H (m) = W + m \quad \sum_{\mu, \nu} t_{\mu, \nu} \frac{1}{W + m} \sigma_\mu \sigma_\nu \tag{2.8}
\]
Now, using the Schur decomposition, we can prove that

$$\det (D_W (0) + m_1 P_+ + m_2 P_-) = \det (W + m_1)^2 \det [W_H (m_1) + \Delta_m] = \det (W + m_2)^2 \det [W_H (m_2) + \Delta_m]$$

where $$\Delta_m = m_2$$, $$m_1 > 0$$. Using this relation, we can rewrite (2.7) as

$$\frac{D_W (m_1)}{D_W (m_2)} = \det \frac{W_H (m_1)}{W_H (m_1) + \Delta_m} \det \frac{W_H (m_2)}{W_H (m_2) + \Delta_m} : \quad (2.10)$$

Next we assert the positive definiteness of both operators on the RHS of (2.10). First, we assume that $$m_2$$ is sufficiently large such that $$(W + m_2)$$, $$W_H (m_2)$$, and $$W_H (m_2)$$ are positive-definite. Then, the second factor on the RHS of (2.10) is positive-definite for $$m_1 > m_1^B$$ where $$m_1^B$$ is the largest value satisfying $$\det [D_W (0) + m_1^B P_+ + m_2 P_-] = 0$$, as a consequence of (2.9). Now, for the operator $$W_H (m_1) (W_H (m_1) + \Delta_m)^{-1} = (1 + \Delta_m W_H (m_1))^{-1}$$, we note that even if $$W_H (m_1)$$ is not well-defined, $$W_H (m_1)^{-1} = P H_W (m_1)^{1/2} P$$ can still be well-defined. From (2.10), $$\det (W_H (m_1) (W_H (m_1) + \Delta_m)^{-1})$$ is positive-definite if both $$\det D_W (m_1)$$ and $$\det [W_H (m_2) + \Delta_m]$$ are positive-definite. It follows that both operators on the RHS of (2.10) are positive-definite for $$m_1 > \max (m_1^A, m_1^B)$$, where $$m_1^B$$ has been defined above, and $$m_1^A$$ is the largest value satisfying $$\det D_W (m_1^A) = 0$$.

Thus the pseudofermion action for one-flavor Wilson fermion with Hasenbusch acceleration can be written as

$$S_{PF} = \Phi_3^\dagger \frac{\Delta_m}{W_H (m_1)} \Phi_3 + \Phi_4^\dagger \frac{\Delta_m}{W_H (m_2)} \Phi_4$$

$$= \Phi_3^\dagger \Phi_3 + \Phi_4^\dagger \Phi_4 \Delta_m 0 \Phi_3^\dagger H_W^{-1} (m_1) 0 \Phi_3 + \Delta_m \Phi_4^\dagger 0 (H_W (0) + m_1 P_+ + m_2 P_-)^{1/2} \Phi_4$$

where $$\Phi_3$$ and $$\Phi_4$$ are pseudofermion fields with 2 spinor components.

3. Domain-wall fermion

The basic idea of domain-wall fermion is to use $$N_s$$ layers of 4-d Wilson-Dirac fermions (with nearest neighbor coupling in the 5-th dimension) such that the exactly chiral fermion fields emerge at the boundary layers in the limit $$N_s \to \infty$$. The domain-wall fermion operator can be defined as

$$D_{dwf} (m) = W m_0 + \sum_\mu \gamma_\mu t_\mu + M (m)$$

$$M (m) = P_+ M_+ (m) + P_- M_- (m)$$

$$M_+ (m)_{s,s'} = \begin{cases} \delta_{s,s'} & 1 < s \leq N_s; \\ \delta_{s,s'} + m \delta_{s,s_{N_s}} & s = 1; \end{cases}$$

$$M_- (m)_{s,s'} = \begin{cases} \delta_{s,s'} & 1 < s \leq N_s; \\ \delta_{s,s'} + m \delta_{s,s_{N_s}} & s = N_s; \end{cases}$$

where $$m$$ is the fermion mass, and $$m_0 \geq 0$$ is a parameter called "domain-wall height". Using the Schur decomposition, the determinant of domain-wall fermion operator can be written as

$$\det D_{dwf} (m) = \det [W m_0 + M_+ (m)]^2 \det W_H (m) = \det [W m_0 + M_- (m)]^2 \det W_H (m); \quad (3.4)$$
where
\[ \mathcal{W}_H (m) = R_5 \ W \ m_0 + M \ (m) \ \ t_\mu \ \frac{1}{W} \ \ m_0 + M \ (m) \ \ t_\nu \ \sigma_\mu^\dagger \sigma_\nu \] (3.5)
\[ \overline{\mathcal{W}}_H (m) = R_5 \ W \ m_0 + M_+ \ (m) \ \ t_\mu \ \frac{1}{W} \ \ m_0 + M \ (m) \ \ t_\nu \ \sigma_\mu \sigma_\nu^\dagger \] : (3.6)

Here \( R_5 \) is the reflection operator in the 5-th dimension, \( \langle R_5 \rangle_{x=v} = \delta_{5,N_5+v} \), which is introduced such that \( \mathcal{W}_H (m) \) and \( \overline{\mathcal{W}}_H (m) \) are hermitian. After incorporating the contribution of the Pauli-Villars fields, the fermion determinant of domain-wall fermion becomes

\[ \det \overline{\mathcal{D}}_{dwf} (m) = \det W m_0 + M_+ (m) \frac{1}{2} \ \det \mathcal{W}_H (m) \] (3.7)

Using the Schur decomposition of \( [\mathcal{D}_{dwf} (1) \ M_+ (1) \ M_+ (m)] \) we obtain the relation
\[ \det W m_0 + M_+ (m) \frac{1}{2} \ \det [\mathcal{W}_H (m) + \Delta \ (m)] = \det W m_0 + M_+ (1) \frac{1}{2} \ \det \overline{\mathcal{W}}_H (1) \ \Delta_+ (m) ; \] (3.8)

where
\[ \{ \Delta_+ (m) \}_{x=v} = (1 \ m) \delta_{5,N_5} \delta_{5+v,N_5} ; \ \{ \Delta_+ (m) \}_{x=v} = (1 \ m) \delta_{5,v} \delta_{5+v} ; \] (3.9)

Using (3.8), we can rewrite (3.7) as
\[ \det \overline{\mathcal{W}}_H (1) \ \Delta_+ (m) = \overline{\mathcal{W}}_H (1) ; \]

and its inverse
\[ \det \mathcal{W}_H (m) = (\mathcal{W}_H (m) + \Delta_+ (m)) \ \det \overline{\mathcal{W}}_H (1) \ \Delta_+ (m) \]
\[ \det 1 + \Delta_+ (m) \ \frac{1}{\mathcal{W}_H (m)} \] (3.10)

can be used to construct the pseudofermion action. Using (3.9), we can simplify (3.10) to
\[ \det 1 + (1 \ m) \ \frac{1}{\mathcal{W}_H (m)} \] (3.11)

Thus we can write the pseudofermion action for one-flavor domain-wall fermion as
\[ S_{PF} = \Phi_1 \Phi_1^\dagger (1 \ m) (0 \ \Phi_1^\dagger) \ [\gamma_5 R_5 \overline{\mathcal{D}}_{dwf} (m)]^1_{s=s_0=1} \ \Phi_1 \]
\[ + \ \Phi_2 \Phi_2^\dagger (1 \ m) (0 \ \Phi_2^\dagger) \ [\gamma_5 R_5 \overline{\mathcal{D}}_{dwf} (1) \ ] \ \Delta_+ (m) \ ] P_{1 \overline{1}}_{s=s_0=N_i} \ \Phi_2 \]

where \( \Phi_1 \) and \( \Phi_2 \) are pseudofermion fields (on the 4-dimensional lattice) with 2 spinor components. Now we assert that the operators in (3.11) are positive-definite for \( m > 0 \). At \( m = 1 \), they are equal to the identity operator, thus are positive-definite. As \( m \) is decreased to be less than one, their positive definiteness will lose only if either of the determinants in (3.11) becomes zero or singular. Using [\( \mathcal{W}M_+ (m) \)] = [\( \mathcal{W}M_+ (1) \)] = 0, and the fact that the eigenvalues of \( M_+ (1) \) and \( M_+ (m) \) have non-zero imaginary parts, we immediately see that \( (\mathcal{W} \ m_0 + M_+ (1)) \) and \( (\mathcal{W} \ m_0 + M_+ (m)) \) cannot have zero eigenvalue for \( m > 0 \). Thus the operators in (3.11) are well-defined for \( m > 0 \). Furthermore, since (3.10) is equal to the determinant of the massive overlap-Dirac operator with polar approximation for the sign function of \( H = \gamma_5 D_w (2 + D_w) \), thus for \( m > 0 \), it must be positive and it follows that the operators in (3.11) are positive-definite.
4. Optimal domain-wall fermion

For finite $N_s$, the chiral symmetry of the domain-wall fermion is not exact. However, one can attain the (mathematically) maximal chiral symmetry by assigning each layer in the 5-th dimension a different weight $\omega_s$, according to the formula derived in Ref. [7]. Then the effective 4D Dirac operator is exactly equal to the overlap-Dirac operator with Zolotarev optimal rational approximation to the sign function of $H_w$. The optimal domain-wall fermion operator [7] is defined as

$$D^\omega_{\text{opt}}(m) = \omega D_W(0) (1+L(m)) + (1 - L(0)) ; \quad L(m) = P_s L_s(m) + P L(m) ;$$

where

$$L_s(m)_{s\bar{s}} = \begin{cases} \delta_{s\bar{s}} & 1 < s < N_s ; \\ m\delta_{s\bar{s}} & s = 1 ; \\ m\delta_{s\bar{s}} & s = N_s ; \end{cases}$$

Here $\omega$ denotes the diagonal matrix of the weights $\omega_s s = 1; s \neq 0$. Moreover, we can construct $\omega$ such that it satisfies $R_5 \omega = \omega R_5$. To derive a positive-definite pseudofermion action for one flavor, we can start from

$$D^\omega_{\text{opt}}(m) = \omega D_W(0) + (1 - L(0)) = (1 + L(m))$$

which is different from $D^\omega_{\text{opt}}(m)$ by a matrix factor independent of the gauge field. Using the same technique we have used for the case of domain-wall fermion, we obtain the positive-definite pseudofermion action as follows\(^1\)

$$S_{PF} = \Phi_1^\dagger \Phi_1 + \Phi_2^\dagger \Phi_2 + \frac{1}{1+m} \left( \begin{array}{cc} 0 & (1)^s \Phi_1^\dagger \\ (1)^s \Phi_2^\dagger & 0 \end{array} \right) R_S^5 D^\omega_{\text{opt}}(m) \left( \begin{array}{cc} 0 & (1)^s \Phi_1^\dagger \\ (1)^s \Phi_2^\dagger & 0 \end{array} \right)$$

$$+ \frac{1}{1+m} \left( \begin{array}{cc} 0 & (1)^s \Phi_1^\dagger \\ (1)^s \Phi_2^\dagger & 0 \end{array} \right) R_S^5 D^\omega_{\text{opt}}(1) \left( \begin{array}{cc} 0 & (1)^s \Phi_1^\dagger \\ (1)^s \Phi_2^\dagger & 0 \end{array} \right)$$

where $\Phi_1$ and $\Phi_2$ are pseudofermion fields with 2 spinor components on the 4-dimensional lattice.

5. Numerical test

We compare the efficiency of HMC simulation of 2-flavor and $(1+1)$-flavor QCD with optimal domain-wall quarks, on the $12^3 \times 24 \times 16(N_s)$ lattice. For the gluon part, we use Iwasaki gauge action at $\beta = 2.30$. In the molecular dynamics, we use the Omelyan integrator [8], and the Sexton-Weingarten multiple-time scale method [9]. The time step for the gauge field ($\Delta t_{\text{Gauge}}$) is the same for both 2-flavor and $(1+1)$-flavor cases, while the time step ($\Delta t_{PF}$) for the pseudofermion field in the $(1+1)$-flavor case is 4 times larger than that for the 2-flavor case such that the acceptance rate is roughly the same for both cases. We use conjugate gradient (CG) with mixed precision for the inversion of the quark matrix (with even-odd preconditioning). The length of each trajectory is set to 2. After discarding 300 trajectories for thermalization, we accumulate 100 trajectories for the comparison of efficiency. Our results are given in Table.I. We see that acceptance rate is almost the same for $(1+1)$-flavor and 2-flavor simulations. If auto-correlation time is the same, then the efficiency of HMC can be estimated by the total acceptance divided by the CG iteration number, and the efficiency ratio for 2-flavor and $(1+1)$-flavor is about 3 : 2.

\(^1\)Detailed derivation will be given in a forthcoming paper.
Table 1: Comparison of HMC efficiency for the 2-flavor and \(1 + 1\)-flavor QCD with optimal domain-wall quarks. The step size for the gauge field \(\Delta \tau_{\text{Gauge}}\) is 0.007 (0.010) for \(m = 0.019 (0.038)\), while the step size \(\Delta \tau_{\text{PF}}\) for \(1 + 1\)-flavor pseudofermions is 0.014 (0.020) for \(m = 0.019 (0.038)\), which is 4 times larger than that for the 2-flavor case. Here, \(N_{\text{iter}}^{(\text{HB})}\), \(N_{\text{iter}}^{(\text{MD})}\), and \(N_{\text{iter}}^{(\text{Total})}\) are the average CG iterations for one trajectory (for generating initial pseudofermion fields, molecular dynamics, and their sum respectively).

6. Concluding remark

We have constructed a positive-definite pseudofermion action for one fermion flavor in lattice gauge theory, with determinant exactly equal to that of the lattice Dirac operator, for the Wilson-Dirac operator, and the (optimal) domain-wall fermion operator respectively. Evidently, these one-flavor positive-definite actions are useful for the HMC simulations in nonperturbative studies of the Standard Model (SM), as well as those beyond the SM. Our numerical tests show that the step size \(\Delta \tau_{\text{PF}}\) (for the pseudofermion field in the molecular dynamics) for the \(1 + 1\)-flavor QCD can be 4 times larger than that for the 2-flavor case, yet with the same acceptance rate. The efficiency ratio for HMC with 2-flavor and \(1 + 1\)-flavor is about \(3 : 2\). With the positive-definite pseudofermion action (4.4) for the strange quark, TWQCD Collaboration is performing the HMC simulation of \((2+1)\)-flavor QCD on the \(16^3 \times 32\times 16\) lattice, using a GPU cluster [10].

Acknowledgment

This work is supported in part by the National Science Council (Nos. NSC96-2112-M-002-020-MY3, NSC96-2112-M-001-017-MY3, NSC98-2119-M-002-001), and NTU-CQSE (Nos. 98R0066-65, 98R0066-69).

References

[1] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, Phys. Lett. B 195 (1987) 216.
[2] A. Borici and P. de Forcrand, Nucl. Phys. B 454, 645 (1995) [arXiv:hep-lat/9505021].
[3] M. A. Clark and A. D. Kennedy, Phys. Rev. Lett. 98, 051601 (2007) [arXiv:hep-lat/0608015].
[4] M. Hasenbusch, Phys. Lett. B 519, 177 (2001) [arXiv:hep-lat/0107019].
[5] R. C. Brower, T. Ivanenko, A. R. Levi and K. N. Orginos, Nucl. Phys. B 484, 353 (1997) [arXiv:hep-lat/9509012].
[6] D. B. Kaplan, Phys. Lett. B 288, 342 (1992) [arXiv:hep-lat/9206013].
[7] T. W. Chiu, Phys. Rev. Lett. 90, 071601 (2003) [arXiv:hep-lat/0209153].
[8] T. Takaishi and P. de Forcrand, Phys. Rev. E 73, 036706 (2006) [arXiv:hep-lat/0505020].
[9] J. C. Sexton and D. H. Weingarten, Nucl. Phys. B 380, 665 (1992).
[10] T. W. Chiu, T. S. Guu, T. H. Hsieh, C. H. Huang, Y. Y. Lee, Y. Y. Mao, K. Ogawa, P. K. Tseng [TWQCD Collaboration], PoS LAT2009, 034 (2009) [arXiv:0911.5029 [hep-lat]].