Dynamical overlap fermions with increased topological tunnelling

Nigel Cundy[*]

University of Regensburg, Universitätstrasse 31, 93040 Regensburg, Germany.
E-mail: nigel.Cundy@physik.uni-regensburg.de

Stefan Krieg

Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany.
E-mail: s.krieg@fz-juelich.de

Thomas Lippert

Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany.
E-mail: th.lippert@fz-juelich.de

Andreas Schäfer

University of Regensburg, Universitätstrasse 31, 93040 Regensburg, Germany.
E-mail: andreas.schaefer@physik.uni-regensburg.de

We present two improvements to our previous dynamical overlap HMC algorithm. We introduce a new method of differentiating the eigenvectors of the Kernel operator, which removes an instability in the fermionic force. Secondly, by simulating part of the fermion determinant exactly, without pseudo-fermions, we are able to increase the rate of topological tunnelling by a factor of more than ten, reducing the auto-correlation.

---

[*]Speaker.

---

1. Introduction

The overlap operator [1] is the only known lattice Dirac operator with an exact lattice chiral symmetry [2]. Since chiral symmetry is important for many low energy observables, it is desirable to use overlap fermions. However, using a full QCD simulation with overlap fermions presents a number of algorithmic challenges. In this paper, we address, and present solutions for, two outstanding algorithmic issues.

The problem of the Dirac δ-function in the fermionic force when changing topological charge has been resolved by using a transmission/reflection algorithm, similar to the case of a classical mechanics particle approaching a potential wall. The original formulation [3, 4] has subsequently been improved in [5] to maximise the rate of topological charge change for a given action jump. Our method is described in these references.

To differentiate the overlap operator in the Hybrid Monte Carlo (HMC) molecular dynamics (MD), it is necessary to differentiate the eigenvectors and eigenvalues of a sparse matrix. Previous methods have led to instabilities when there are degenerate eigenvalues. In section 2 we discuss a new method which avoids these instabilities [6].

The topological auto-correlation depends on the rate of topological activity. To have high rate of topological index change, it is necessary to reduce the discontinuity in the action at this point. In section 3, we outline a new method for reducing this action jump [7].

Our overlap operator is defined as

$$D = (1+\mu) + \gamma_5(1-\mu)\text{sign}(Q). \tag{1.1}$$

In our tests, Q is the standard Wilson operator,

$$Q = \gamma_5 \left[\delta_{xy} - \kappa \left((1-\gamma_\mu)U_\mu(x)\delta_{y,x+\mu} + (1+\gamma_\mu)U_\mu^\dagger(x-\mu)\delta_{y,x-\mu} \right) \right], \tag{1.2}$$

with $\kappa = 0.2$ and, in section 3, two levels of stout smearing [8] at parameter 0.1. Our numerical tests are performed on $8^3 16$ lattices at a lattice spacing of about 0.15fm (measured using r_0), with quark masses $\mu = 0.03, 0.04, 0.05$, corresponding to pion masses in the range $500 - 1000$MeV.

2. Eigenvector mixing

2.1 Differentiating Eigenvectors

The eigenvalues λ_i and eigenvectors $|\psi_i\rangle$ of a matrix Q are defined by

$$Q|\psi_i\rangle = \lambda_i|\psi_i\rangle. \tag{2.1}$$

After a small change to the matrix, δQ, the new eigenvalue equation is

$$(Q + \delta Q)|\psi_i'\rangle = \lambda_i'|\psi_i'\rangle. \tag{2.2}$$

We can expand the new eigenvectors in terms of the old basis

$$|\psi_i'\rangle = |\psi_i\rangle + \sum_j (\cos\theta_{ij} - 1)|\psi_i\rangle + e^{i\phi_{ij}}\sin\theta_{ij}|\psi_j\rangle. \tag{2.3}$$

We assume that only one of the mixing angles θ_{ij} is large, so it is unimportant that $|\psi_i'\rangle$ is not normalised[1]. The mixing angles θ_{ij} and ϕ_{ij} are

$$\tan 2\theta_{ij} = \frac{2\sqrt{\langle\psi_i|\delta Q|\psi_j\rangle\langle\psi_j|\delta Q|\psi_i\rangle}}{\lambda_j - \lambda_i + \langle\psi_j|\delta Q|\psi_j\rangle - \langle\psi_i|\delta Q|\psi_i\rangle}$$

$$e^{i\phi_{ij}} = \sqrt{\frac{\langle\psi_j|\delta Q|\psi_i\rangle}{\langle\psi_i|\delta Q|\psi_j\rangle}}. \tag{2.4}$$

Expanding θ and ϕ in $\tau/(\lambda_i - \lambda_j)$ gives to lowest order

$$\delta|\psi_i\rangle = \sum_{j\neq i}|\psi_j\rangle\frac{\langle\psi_j|\delta Q|\psi_i\rangle}{\lambda_j - \lambda_i} = \frac{1}{Q - \lambda_i}(1 - |\psi_i\rangle\langle\psi_i|)\delta Q|\psi_i\rangle, \tag{2.5}$$

which agrees with other methods (see, for example, [5]). Equation (2.5) breaks down when $\lambda_i - \lambda_j$ is small. In this situation, it is necessary to use the exact expressions for the mixing angles (2.4). Defining $T_\mu^n(x)$ as the eight SU(3) generators on one link, we can write

$$\delta Q_{ij} = \langle\psi_i|\delta Q|\psi_j\rangle = \tau\pi_\mu^n(x)\alpha_{ij}^{n,x,\mu}, \tag{2.6}$$

where π represents the MD momentum, and the vectors α_{ij} are given by

$$\alpha_{ij}^{n,x,\mu} = -i\kappa\langle\psi_i|\gamma_5\left[(1-\gamma_\mu)T_\mu^n(x)U_\mu(x)\delta_{y,x+\mu} - (1+\gamma_\mu)U_\mu^\dagger(x)T_\mu^i(x)\right]\delta_{y,x-\mu}|\psi_j\rangle. \tag{2.7}$$

The generalisation to the smeared operator is trivial [6]. The NAC force, F^{NAC}, constructed for the overlap operator in [6], is proportional to α_{ij}, and a function of only $\pi\alpha$ and the gauge field, U. We can construct a reversible algorithm by combining forward and backward half-steps (using a quick and reversible iterative procedure for the backward step). The force is not area conserving, but it is possible to use a non-area conserving force by calculating the Jacobian, J, and including $\log J$ in the HMC accept/reject step (see [5] for an example). J can be written as

$$J = \left|1 + \alpha_{ij}^{n,x,\mu}\frac{\partial F_{ij}}{\partial\pi^{m,y,v}}\right| = \left|1 + A_{ij,kl}\alpha_{ij}\alpha_{kl}\right|. \tag{2.8}$$

It is now a simple task to rewrite α_{ij} in terms of an orthonormal complete basis α', and use

$$\det\left[1 + A'\alpha'\alpha'^\dagger\right] = \det\left[1 + A'\right] \tag{2.9}$$

to calculate J. $\log J$ scales with τ^3, and does not affect the HMC acceptance rate.

2.2 Numerical results

In figure 1, we compare the fermionic forces for the old algorithm and the NAC algorithm across a typical HMC trajectory (using overlap fermions with no smearing). It is clear that the old method gives an unstable force, which cannot be used, while the NAC force is stable. In all our test trajectories, we have not observed a $\log J$ larger than 0.3.

---

[1]A more general expression can easily be constructed should this assumption break down.
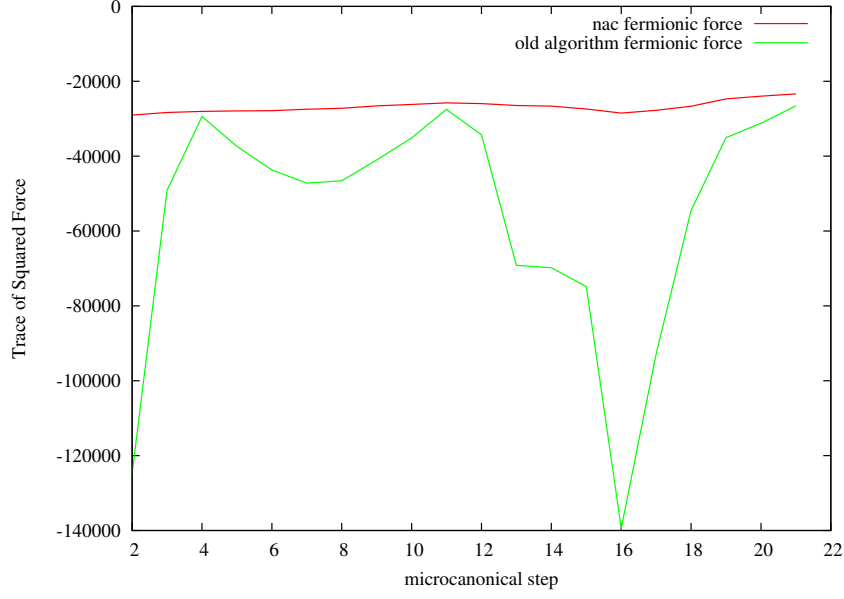
Figure 1: Comparison of the trace of the square of the fermionic forces for the proposed and old algorithms with $\tau = 0.016$ on one of the $\mu = 0.05$ trajectories.

3. Determinant factorisation

3.1 Introduction

The topological auto-correlation is related to the rate of topological activity. We need to be able to measure this rate, from which we can estimate the auto-correlation. Since instanton anti-instanton pairs are difficult to observe, our best way of measuring the auto-correlation is from topological charge changes. The probability of transmission scales as $\min(1, e^{\Delta S})$ [5], where ΔS, the action discontinuity at the topological sector boundary, scales as μ^{-2}. Therefore it is necessary to reduce ΔS.

However, the determinant of the actual Dirac operator, $\log \det(DD^\dagger)$, scales as $\log \mu$ rather than μ^{-2} [9]. The difference is caused the different functional forms of the pseudo-fermion estimate and the actual determinant. It has already been shown that adding additional pseudo-fermions reduces ΔS considerably [10, 11], and this (with one additional pseudo-fermion) is our algorithm A, which we test against our new methods. We propose [7] factorising the determinant into a large continuous part, which can be treated with pseudo-fermions, and a discontinuous determinant small enough to be calculated exactly.

3.2 Algorithm C

Algorithm C factorises the determinant using

$$\det\left[\gamma_5 \frac{1+\mu}{1-\mu} + \varepsilon(Q)\right] = \det\left[\gamma_5 \frac{1+\mu}{1-\mu} + \tilde{\varepsilon}(Q)\right] \det\left[\delta_{ij} + \langle\psi_i| \frac{1}{\gamma_5 \frac{1+\mu}{1-\mu} + \tilde{\varepsilon}(Q)} |\psi_j\rangle(\varepsilon(\lambda_j) - \tilde{\varepsilon}(\lambda_j))\right]$$

$$= \det[D_1] \det[D_2]. \tag{3.1}$$

$\tilde{\varepsilon}$ is a continuous approximation of the sign function, only differing for the smallest n eigenvalues below a cutoff Λ. In our tests, we used a Zolotarev rational approximation. We simulate $\det[D_1]$ using pseudo-fermions; $\det[D_2]$ is calculated by standard methods. We add $-\log \det[D_2]$ to the HMC action. To maintain a high HMC acceptance we calculate $\log \det[D_2]$'s force (unlike the proposal in [12]). Differentiating $\log \det[D_2]$ is straight-forward using the methods of section 2. Algorithm C requires an inversion of D_1 for each eigenvalue projected. While deflation methods [13] reduce this cost, it could lead to difficulties on larger volumes.[2]

3.3 Algorithm F

To avoid this additional cost, algorithm F projects out just one vector, $|a\rangle$, which is equal to the smallest eigenvector at the moment of crossing. Thus, we write

$$\det[1 + \gamma_5 \varepsilon(Q)] = \det\left[1 + \gamma_5 \varepsilon(Q)(1 - |a\rangle \langle a|)\right] \det\left[1 + \frac{1}{1 + \gamma_5 \varepsilon(Q)(1 - |a\rangle \langle a|)} \gamma_5 \varepsilon(Q) |a\rangle \langle a|\right]$$

$$= \det[D_3] \det[D_4] \tag{3.2}$$

To ensure D_3 is both continuous and practical, $|a\rangle$ must satisfy $|a\rangle = |\psi_i\rangle$ at $\lambda_i = 0$ and $\langle \psi_i | a \rangle = 0$ for $\lambda_i^2 > \Lambda^2$ (where Λ^2 is some suitable eigenvalue cut-off). For the algorithm to remain stable at larger lattice volumes $|a\rangle$ and its differential must be continuous, $d/d\lambda_i |a\rangle$ must be sufficiently small, the eigenvalues and eigenvectors must be differentiated using the results of section 2 and the approximate overlap operator D_3 must not have any exceptional configurations. We construct $|a\rangle$ from n eigenvectors using

$$|a\rangle = \beta_0 |\psi_0\rangle \frac{\langle \psi_0 | \Gamma \rangle}{|\langle \psi_0 | \Gamma \rangle|} + \beta_1 |\psi_1\rangle \frac{\langle \psi_1 | \Gamma \rangle}{|\langle \psi_1 | \Gamma \rangle|} + \dots, \tag{3.3}$$

where Γ is a constant vector used to fix the relative phase of the eigenvectors, and currently we use

$$\tan \frac{\pi}{2} \beta_i = \lambda_i^4 \Lambda^{4(n-1)} / (\Lambda^2 - \lambda_i^2)^2 \prod_{j \neq i} \lambda_j^4, \tag{3.4}$$

although we are still searching for the best function to use for β. Algorithm C has shown stable MD with high HMC acceptance and good reversibility on our test $8^3 16$ ensembles.[3] Algorithm F has encountered large differentials of β with respect to the eigenvalues, which may cause problems on larger volumes, although the HMC acceptance rate is still acceptable on our lattices. We do not expect any scaling of the action jump with the lattice volume.

3.4 Numerical results

Figure 2 and table 1 give preliminary results for the action jump on our $8^3 16$ test configurations. ΔS is much smaller for algorithms C and F than for algorithm A, and the distribution is

---

[2] We cannot decrease Λ without increasing the force; therefore the number of projected eigenvalues will increase as the volume increases.

[3] A variant of algorithm C, not using our method of differentiating the eigenvectors and not treating the small matrix in the molecular dynamics, has encountered problems on $12^3 24$ volumes [14]. We suspect that this was because the range of the approximate sign function was set to small, and a less efficient method was used when differentiating eigenvectors.
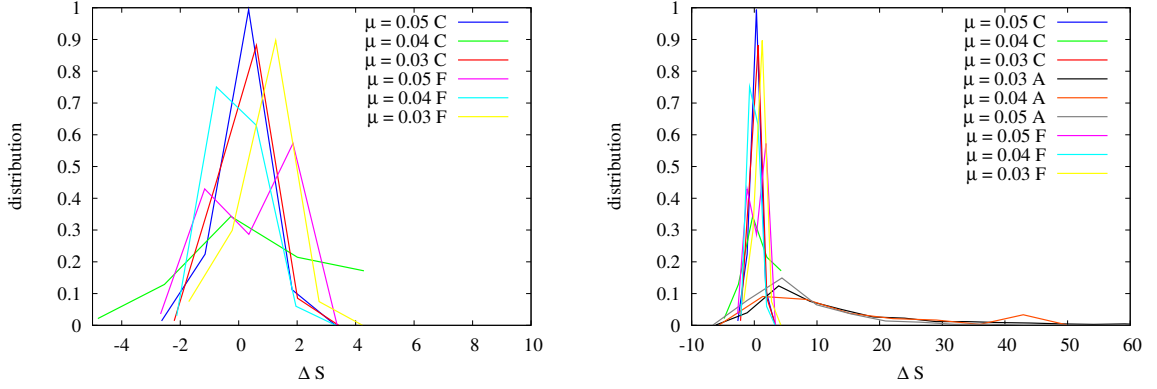
Figure 2: The distribution of the action discontinuity ΔS for algorithms C and F (left), and A, C and F (right), at masses $\mu = 0.03, 0.04$ and 0.05.

μ	A	C	F	A	C	F
0.03	14.0(7)	0.28(8)	0.70(21)	21.9	0.78	1.13
0.04	13.8(10)	0.48(32)	-0.06(11)	19.5	2.4	0.86
0.05	7.8(4)	0.23(7)	0.22(22)	13.6	0.72	1.3

Table 1: The mean values of ΔS for algorithms A,C and F at three quark masses μ (left) and the standard deviations (right).

much narrower. The results for both algorithms C and F give a high transmission rate for all μ. We see little variation with the quark mass with these methods, unlike algorithm A. This suggests transmission may be possible with much smaller μ.

4. Conclusion

We have developed two additions to the overlap HMC algorithm designed to significantly improve simulations at large volume and small fermion mass. By employing a non area conserving algorithm to differentiate the eigenvectors and eigenvalues of the Dirac operator, we remove various instabilities that otherwise are encountered in the fermionic force. By factorising the fermion determinant, we decrease the action jump at the topological sector boundary considerably (we have observed transmission rates of up to 80%), reducing the auto-correlation by an order of magnitude. Thus, with dynamical overlap fermions, it is possible to accurately sample all topological sectors efficiently. It remains an open question whether other lattice formulations will, at small lattice spacing, suffer from the large auto-correlations which we have now avoided.

Acknowledgements

References

[1] R. Narayanan and H. Neuberger, Chiral fermions on the lattice, Phys. Rev. Lett. 71 (1993) 3251–3254, hep-lat/9308011.

[2] M. Lüscher, Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation, Phys. Lett. B428 (1998) 342–345, hep-lat/9802011.

[3] Z. Fodor, S. D. Katz, and K. K. Szabo, Dynamical overlap fermions, results with hybrid Monte-Carlo algorithm, JHEP 08 (2004) 003, hep-lat/0311010.

[4] N. Cundy et al., Numerical methods for the QCD overlap operator. IV: Hybrid Monte Carlo, hep-lat/0502007.

[5] N. Cundy, Current status of dynamical overlap project, Nucl. Phys. Proc. Suppl. 153 (2006) 54–61, hep-lat/0511047.

[6] N. Cundy, Small Wilson Dirac operator eigenvector mixing in dynamical overlap hybrid Monte-Carlo, arXiv:0706.1971 [hep-lat].

[7] N. Cundy, S. Krieg, T. Lippert, and A. Schäfer, in preparation.

[8] C. Morningstar and M. J. Peardon, Analytic smearing of SU(3) link variables in lattice QCD, Phys. Rev. D69 (2004) 054501, hep-lat/0311018.

[9] G. I. Egri, Z. Fodor, S. D. Katz, and K. K. Szabo, Topology with dynamical overlap fermions, JHEP 01 (2006) 049, hep-lat/0510117.

[10] M. Hasenbusch, Speeding up the Hybrid-Monte-Carlo algorithm for dynamical fermions, Phys. Lett. B519 (2001) 177–182, hep-lat/0107019.

[11] T. A. DeGrand and S. Schäfer, Physics issues in simulations with dynamical overlap fermions, Phys. Rev. D71 (2005) 034507, hep-lat/0412005.

[12] S. Schäfer, Algorithms for dynamical overlap fermions, hep-lat/0609063.

[13] N. Cundy, S. Krieg, and T. Lippert, Improving the dynamical overlap algorithm, PoS LAT2005 (2006) 107, hep-lat/0511044.

[14] T. Degrand and S. Schäfer, personal Communication.

7