DESY 03–142
Edinburgh 2003/14
LU-ITP 2003/018
ZIB-Report 03–30

Accelerating Hasenbusch's acceleration of Hybrid Monte Carlo*

A. Ali Khan[a], T. Bakeyev[b], M. Göckeler[c,d], R. Horsley[e], D. Pleiter[f], P.E.L. Rakow[g], A. Schäfer[d], G. Schierholz[f,h] and H. Stüben[i] (QCDSF Collaboration)

[a]Institut für Physik, Humboldt-Universität zu Berlin, 12489 Berlin, Germany

[b]Joint Institute for Nuclear Research, 141980 Dubna, Russia

[c]Institut für Theoretische Physik, Universität Leipzig, 04109 Leipzig, Germany

[d]Institut für Theoretische Physik, Universität Regensburg, 93040 Regensburg, Germany

[e]School of Physics, The University of Edinburgh, Edinburgh EH9 3JZ, UK

[f]John von Neumann-Institut für Computing NIC, 13738 Zeuthen, Germany

[g]Theoretical Physics Division, Department of Mathematical Sciences, University of Liverpool, Liverpool L69 3BX, UK

[h]Deutsches Elektronen-Synchrotron DESY, 22603 Hamburg, Germany

[i]Konrad-Zuse-Zentrum für Informationstechnik Berlin, 14195 Berlin, Germany

Hasenbusch has proposed splitting the pseudo-fermionic action into two parts, in order to speed-up Hybrid Monte Carlo simulations of QCD. We have tested a different splitting, also using clover-improved Wilson fermions. An additional speed-up between 5 and 20% over the original proposal was achieved in production runs.

1. INTRODUCTION

Hybrid Monte Carlo (HMC) [1] is the standard algorithm employed in numerical simulations of full QCD. However, the computational cost of such simulations grows rapidly with decreasing quark mass. At light quark mass (a) the condition number of the fermion matrix increases, which leads to an increased number of iterations in solving the corresponding system of linear equations, (b) the acceptance rate decreases, which has to be compensated by decreasing the integration step size, and (c) the autocorrelation time in units of trajectories increases.

In [2] Hasenbusch has proposed numerical methods to improve conditions (a) and (b) in order to accelerate HMC simulations with dynamical fermions. He suggested splitting the fermion matrix into two pieces both having a smaller condition number than the original matrix. For each factor a pseudo-fermionic field is introduced and the Yang-Mills and fermionic parts of the action are put onto different time-scales in the leap-frog integration. These methods were tested in simulations with clover-improved Wilson fermions and a speed-up of 2 was obtained [3]. The acceleration is greater at lower quark masses [4].

The multiple time-scale approach was initially advocated in [6] where Yang-Mills and pseudo-fermionic terms were put onto different time-scales. The idea was refined in [7] where the following criteria for an efficient splitting of the action $S = S_{\rm UV} + S_{\rm IR}$ were formulated. The force term generated by $S_{\rm UV}$ should be cheap to compute compared to $S_{\rm IR}$. And the splitting should mainly capture the high-frequency modes of the system in $S_{\rm UV}$ and the low-frequency modes in $S_{\rm IR}$. In order to achieve this, a low-order polynomial approximation for mimicking the high-

*Poster presented by H. Stüben at Lattice 2003.

frequency modes was introduced in the fermionic action and the action was split accordingly [7].

In this study the aforementioned methods are combined. The fermion matrix is split according to [2]. Following [7] the two fermionic contributions are put onto different time-scales (this possibility was already mentioned in [3] but no additional advantage was found). In a production run we compared our splitting with the splitting of [3] and found an additional speed-up of about 20% [5]. Here we report on the same comparison for a run at smaller quark mass.

2. NOTATION, TECHNICAL DETAILS

2.1. Actions

We simulated two flavour QCD with clover-improved Wilson fermions employing even/odd preconditioning. The standard action for this model reads

$$S_0[U,\phi^\dagger,\phi] = S_G[U] + S_{\rm det}[U] + \phi^\dagger (Q^\dagger Q)^{-1}\phi \quad (1)$$

where $S_G[U]$ is the standard Wilson plaquette action, $\phi^\dagger$ and ϕ are pseudo-fermion fields, and

$$S_{\rm det}[U] = -2{\rm Tr}\log(1+T_{oo})\,, \quad (2)$$

$$Q = (\mathbf{1}+T)_{ee} - M_{eo}(\mathbf{1}+T)^{-1}_{oo}M_{oe}\,. \quad (3)$$

T_{ee} (T_{oo}) is the clover matrix on even (odd) sites

$$(T)_{a\alpha,b\beta}(x) = \frac{i}{2}c_{sw}\kappa\sigma^{\alpha\beta}_{\mu\nu}\mathcal{F}^{ab}_{\mu\nu}(x)\,. \quad (4)$$

M_{eo} and M_{oe} are Wilson hopping matrices connecting even with odd and odd with even sites, respectively

The standard action is modified [2] by introducing an auxiliary matrix $W = Q + \rho$, $\rho \in \mathbb{R}$, and pseudo-fermion fields $\chi^\dagger$, χ

$$S_1[U,\phi^\dagger,\phi,\chi^\dagger,\chi] = S_G[U] + S_{\rm det}[U] + \phi^\dagger W(Q^\dagger Q)^{-1}W^\dagger\phi + \chi^\dagger(W^\dagger W)^{-1}\chi\,. \quad (5)$$

2.2. Multiple time-scales

One step of the reversible integrator V_n we used is given by [7]

$$V_n(\tau) = V_{\rm IR}\left(\frac{\tau}{2}\right)\times \left[V_{\rm UV}\left(\frac{\tau}{2n}\right)V_Q\left(\frac{\tau}{n}\right)V_{\rm UV}\left(\frac{\tau}{2n}\right)\right]^n \times V_{\rm IR}\left(\frac{\tau}{2}\right) \quad (6)$$

where n is a positive integer and the time-scales are τ and τ/n. The effect of V_Q, $V_{\rm UV}$, $V_{\rm IR}$ on the system coordinates $\{P,Q\}$ is:

$$V_Q(\tau) \;:\; Q \to Q + \tau P \quad (7)$$

$$V_{\rm UV}(\tau) \;:\; P \to P - \tau\partial S_{\rm UV} \quad (8)$$

$$V_{\rm IR}(\tau) \;:\; P \to P - \tau\partial S_{\rm IR} \quad (9)$$

2.3. Splittings of the actions

We performed simulations employing three splittings. The first splitting is based on S_0 (1). The other two are different splittings of S_1 (5).

Splitting A (Sexton and Weingarten [6]):

$$\begin{aligned} S_{\rm UV} &= S_G[U] \\ S_{\rm IR} &= S_{\rm det}[U] + \phi^\dagger(Q^\dagger Q)^{-1}\phi \end{aligned} \quad (10)$$

Splitting B (Hasenbusch and Jansen [2,3]):

$$\begin{aligned} S_{\rm UV} &= S_G[U] \\ S_{\rm IR} &= S_{\rm det}[U] + \phi^\dagger W(Q^\dagger Q)^{-1}W^\dagger\phi \\ &\quad + \chi^\dagger(W^\dagger W)^{-1}\chi \end{aligned} \quad (11)$$

Splitting C (our proposal [5]):

$$\begin{aligned} S_{\rm UV} &= S_G[U] + S_{\rm det}[U] + \chi^\dagger(W^\dagger W)^{-1}\chi \\ S_{\rm IR} &= \phi^\dagger W(Q^\dagger Q)^{-1}W^\dagger\phi \end{aligned} \quad (12)$$

Our proposal (12) is motivated by the hypothesis that most of the high-frequency modes of the pseudo-fermion part of the action (5) are located in $\chi^\dagger(W^\dagger W)^{-1}\chi$. We also put the clover determinant $S_{det}[U]$ on the "ultraviolet" time-scale because the force generated by it is computationally cheap. The computationally expensive term $\phi^\dagger W(Q^\dagger Q)^{-1}W^\dagger\phi$ is put on the "infra-red" time-scale.

2.4. Solver

The standard *conjugate gradient* algorithm was used. Starting vectors were obtained from chronological inversion [8] with $N_{\rm guess} = 7$. We checked reversibility by forward and backward integration starting with thermalised configurations, whereupon deviations of energies were less than 10^{-10}.

2.5. Computational gain

The CPU-cost is roughly given by $t_{\rm CPU} \propto (N_Q + N_W)\tau_{\rm int}$ where N_Q and N_W are the numbers of multiplications (per trajectory) with $Q^\dagger Q$

Table 1
Parameters and statistics. (Statistics for each parameter set in Table 2.)

run	V	β	κ	$c_{\rm sw}$	m_π/m_ρ	trajectory length	statistics
(I)	$16^3 \times 32$	5.29	0.13550	1.9192	≈ 0.7	1	300 trajectories
(II)	$24^3 \times 48$	5.25	0.13575	1.9603	≈ 0.6	0.5	100 trajectories

Table 2
Further parameters and performance results. ($N_{\rm steps}$ is the number of integrator steps (6).)

run	splitting	ρ	n	$N_{\rm steps}$	$P_{\rm acc}$	N_Q	N_W	$N_Q + N_W$	$D_{\rm gain}$
(I)	**A**	0	3	140	0.601	139492	0	139492	1
	B	0.5	3	100	0.599	65951	5233	71184	1.95
		0.2	3	70	0.664	47214	7378	54592	2.82
	C	0.5	3	50	0.547	45160	7687	52847	2.40
		0.2	3	40	0.663	32659	12373	45032	3.42
(II)	**A**	0	3	180	0.780	267363	0	267363	1
	B	0.2	3	90	0.891	89517	3242	92759	3.29
		0.1	3	90	0.871	66432	5786	72218	4.13
	C	0.2	3	50	0.799	74002	7967	81969	3.34
		0.1	3	50	0.896	57018	13624	70642	4.35

and $W^\dagger W$, respectively. In order to estimate the computational gain we assume $\tau_{\rm int} \propto 1/P_{\rm acc}$ [3] and calculate the computational gain of splittings B and C compared to A by

$$D_{\rm gain}^{\rm (B,C)} = \frac{N_Q^{\rm (A)}}{N_Q^{\rm (B,C)} + N_W^{\rm (B,C)}} \frac{P_{\rm acc}^{\rm (B,C)}}{P_{\rm acc}^{\rm (A)}} \,. \qquad (13)$$

3. RESULTS

We have tested splittings A, B and C in two production runs. The parameters of the runs are listed in Table 1. Performance results are shown in Table 2. The values for run (I) are old results [5]. The values for run (II) are new. One sees that the speed-up is considerable and that it grows with decreasing quark mass. ρ has to be lowered at smaller quark masses. In run (I) splitting C accelerates the simulation by about 20% better than splitting B. In run (II) the additional gain of using splitting C is only about 5%.

In conclusion, the methods proposed by Hasenbusch work very well. Our variant of his method seems to perform even slightly better. In both cases the choice of the new parameter ρ affects the speed-up noticeably. It would be interesting to know how the number of integrator steps and the trajectory length influence the gain.

ACKNOWLEDGEMENTS

Computations were performed on the APEmille at NIC Zeuthen and on the Hitachi SR8000-F1 at Leibniz-Rechenzentrum Munich.

REFERENCES

1. S. Duane, A. Kennedy, B. Pendleton and D. Roweth, Phys. Lett. B 195 (1987) 216.
2. M. Hasenbusch, Phys. Lett. B 519 (2001) 177
3. M. Hasenbusch and K. Jansen, hep-lat/0211042.
4. M. Hasenbusch, these proceedings.
5. A. Ali Khan, T. Bakeyev, M. Göckeler, R. Horsley, D. Pleiter, P. Rakow, A. Schäfer, G. Schierholz and H. Stüben, Phys. Lett. B 564 (2003) 235.
6. J. Sexton and D. Weingarten, Nucl. Phys. B 380 (1992) 665.
7. M. Peardon and J. Sexton, hep-lat/0209037.
8. R. Brower, T. Ivanenko, A. Levi and K. Orginos, Nucl. Phys. B 484 (1997) 353.