

Polynomial Filtering for HMC in Lattice QCD

Waseem Kamleh^{*†}

Trinity College Dublin

E-mail: waseem@maths.tcd.ie

Michael J. Peardon

Trinity College Dublin

E-mail: mjp@maths.tcd.ie

TrinLat Collaboration

The use of short order polynomial approximations to $1/x$ as a UV filter for HMC is investigated. Previous work in the Schwinger model (2D QED) showed that the integration step size may be increased within a multiple time scale integration by separating the UV and IR dynamics with a polynomial filter. We test the same method within the framework of Lattice QCD.

XXIIIrd International Symposium on Lattice Field Theory

25-30 July 2005

Trinity College, Dublin, Ireland

^{*}Speaker.

[†]Supported under SFI basic research grant 04/BR/P0266.

The task of performing dynamical simulations at light quark masses remains a challenge to lattice QCD. Available computing power has and continues to increase, yet even at the current stage of multi-Teraflop computing capabilities, algorithmic breakthroughs are needed before the desired quark masses can be realised. Hybrid Monte Carlo[1] is still the most often used algorithm for generating dynamical configurations. There are three main computational bottlenecks that present themselves when attempting to simulate lighter fermions.

The first is that of lattice volume. As the quark mass (and hence pion mass) decreases, larger spatial volumes are required to avoid finite size effects. For standard QCD simulations, this is essentially unavoidable. The second bottleneck is the cost of inverting the fermion matrix. The number of conjugate gradient iterations required to do this increases as the mass decreases. This problem is more severe for Wilson-type fermions as they do not have the protection of chiral symmetry to prevent standard “exceptional” configurations. The use of chiral fermions can ameliorate this problem in the sense that the condition number of the fermion matrix only goes like a single inverse power of the quark mass, but other computational expenses are introduced in the process.

The third bottleneck, and the one which is key to this work, is the ultraviolet slowing down of the molecular dynamics integration. When performing HMC for dynamical fermions, as the quark mass decreases the fluctuations induced by the pseudofermions increases, requiring finer integration step sizes to be used to bring these under control, significantly increasing the computational cost of generating dynamical configurations. One way of dealing with this is to use a Sexton and Weingarten[2] integration scheme with multiple scales, separating the ultraviolet and infrared dynamics. The key point in this case is that one must possess an inexpensive means of evaluating the UV dynamics of the fermions. In this work we propose to do this using a polynomial filter, following earlier work within the Schwinger model[5].

1. Multiple Time Scale HMC

Given a configuration U , the hybrid Monte Carlo algorithm generates the next element in the Markov chain by first performing a Molecular Dynamics (MD) integration to generate a new configuration U' , and then performing a Metropolis accept/reject step on the proposed configuration with a probability $\rho(U \rightarrow U') = e^{-\Delta\mathcal{H}}$. Here,

$$\mathcal{H}(P, U) = \sum_{x, \mu} \frac{1}{2} \text{Tr} P_{\mu}(x)^2 + S[U], \quad (1.1)$$

is a Hamiltonian system in which the four dimensional lattice is embedded through the addition of a fictitious “simulation” time τ , and P is defined as the conjugate momenta to U .

By requiring that the Hamiltonian be conserved along the molecular dynamics trajectory one obtains the discretised equations of motion,

$$U_{\mu}(x, \tau + \Delta\tau) = U_{\mu}(x, \tau) \exp(i\Delta\tau P_{\mu}(x, \tau)), \quad (1.2)$$

$$P_{\mu}(x, \tau + \Delta\tau) = P_{\mu}(x, \tau) - U_{\mu}(x, \tau) \frac{\delta S}{\delta U_{\mu}(x, \tau)}. \quad (1.3)$$

Define the corresponding time evolution operators,

$$V_T(\Delta\tau) : \{U(\tau), P(\tau)\} \rightarrow \{U(\tau + \Delta\tau), P(\tau)\}, \quad (1.4)$$

$$V_S(\Delta\tau) : \{U(\tau), P(\tau)\} \rightarrow \{U(\tau), P(\tau + \Delta\tau)\}. \quad (1.5)$$

For sufficiently small finite step sizes $\Delta\tau$ the Hamiltonian is approximately conserved along the trajectory, giving high acceptance probabilities. Different integration schemes give differing discretisation errors. The simplest (and most used) integration scheme is the leapfrog

$$V(\Delta\tau) = V_S\left(\frac{\Delta\tau}{2}\right)V_T(\Delta\tau)V_S\left(\frac{\Delta\tau}{2}\right). \quad (1.6)$$

Integration trajectories typically have unit length, and hence as the step size $\Delta\tau$ decreases, the number of integration steps (and the computational cost) increases. Now, our action $S = S_g + S_{\text{pf}}$, consists of the gauge field and pseudo-fermion field components,

$$S_g = \beta \sum_{x, \mu < \nu} \frac{1}{3} \text{Re Tr}(1 - U_{\mu\nu}(x)), \quad (1.7)$$

$$S_{\text{pf}} = \sum_x \phi^\dagger(x) (D_w^\dagger D_w)^{-1} \phi(x). \quad (1.8)$$

Here, D_w is the Wilson fermion matrix. Each time we act with $V_S(\Delta\tau)$ we need to evaluate the pseudofermion force term,

$$F_{\text{pf}} = \frac{\delta S_{\text{pf}}}{\delta U}. \quad (1.9)$$

This involves inverting the fermion matrix, and is by far the major cost in standard HMC simulations. However, for split actions $S = S_1 + S_2$ we can use a multiple time scale integration scheme[2],

$$V(\Delta\tau) = V_2\left(\frac{\Delta\tau}{2}\right) \left[V_1\left(\frac{\Delta\tau}{m}\right) \right]^m V_2\left(\frac{\Delta\tau}{2}\right), \quad (1.10)$$

where

$$V_1(\Delta\tau) = V_{S_1}\left(\frac{\Delta\tau}{2}\right)V_T(\Delta\tau)V_{S_1}\left(\frac{\Delta\tau}{2}\right), \quad (1.11)$$

$$V_2 = V_{S_2}(\Delta\tau). \quad (1.12)$$

In this scheme, V_1 is evaluated m times more often than V_2 . Such an integration scheme is only effective if two conditions are satisfied. First, the force term F_1 due to S_1 must be computationally inexpensive to evaluate compared to that of S_2 . Secondly, as the effective step size for S_2 is m times larger than that of S_1 , the force term due to S_2 must correspondingly be *smaller* than that due to S_1 .

2. Polynomial Filtering

As the gauge force F_g is cheap to evaluate compared to the pseudofermion force F_{pf} , one might choose $S_1 = S_g$ and $S_2 = S_{\text{pf}}$. However it is only at heavy quark masses that we have $F_g > F_{\text{pf}}$. At light quark masses the UV fluctuations in the pseudo fermion force become too large for multiple time scales to be effective. However, we can use a polynomial filter $P = P(D_w)$ to separate the ultraviolet and infrared physics in the pseudofermion force[5],

$$S_{\text{poly}} = \chi^\dagger P^\dagger P \chi, \quad (2.1)$$

$$S_{\text{pf}} = \phi^\dagger (P^\dagger D_w^\dagger D_w P)^{-1} \phi. \quad (2.2)$$

For short polynomials, S_{poly} is fast to evaluate compared to S_{pf} so we split the action in the following way,

$$S_1 = S_g + S_{\text{poly}}, \quad (2.3)$$

$$S_2 = S_{\text{pf}}. \quad (2.4)$$

In order that the force terms satisfy $F_{\text{poly}} > F_{\text{pf}}$ we need to choose a polynomial term which captures the ultraviolet physics and filters it from the pseudofermion term.

An effective UV filter is an n^{th} order polynomial approximation to $1/z$,

$$P_n(z) = a_n \prod_{k=1}^n (z - z_k) \approx \frac{1}{z}. \quad (2.5)$$

In this work we test two different types of approximation[6]. The first, a Hermitian Chebyshev approximation, has the following roots (with $\theta_k = \frac{2\pi k}{n+1}$),

$$z_k = \lambda \left[\frac{1}{2}(1 + \varepsilon)(1 - \cos \theta_k) - i\sqrt{\varepsilon} \sin \theta_k \right]. \quad (2.6)$$

The normalisation is defined by $z_0 = \frac{1}{2}(1 + \varepsilon)$, with $a_n = \frac{1}{z_0 \prod_{k=1}^n (z_0 - z_k)}$. The approximation is good between $[\varepsilon, 1]$, so we rescale with $\lambda = 1 + 8\kappa$. The second, a non-Hermitian Chebyshev approximation has the same normalisation, but slightly different roots,

$$z_k = d(1 - \cos \theta_k) - i\sqrt{d^2 - c^2} \sin \theta_k. \quad (2.7)$$

3. Results

All simulation results are for $8^3 \times 16$ lattices using the Wilson gauge and fermion actions, at $\beta = 5.6$ and $\kappa = 0.1575$. Firstly, equilibrium configurations were generated using a standard HMC algorithm. In order to tune the available parameters the force terms due to the (filtered) pseudofermion term F_{pf} and the polynomial term F_{poly} were measured. To compare the size of the forces we used the norm

$$\|F\| = \left(\sum_{\mu, x, i, j} [F_{\mu, x}]_{ij}^2 \right)^{\frac{1}{2}}. \quad (3.1)$$

Figure 1 shows the results for the Hermitian Chebyshev polynomial. We see that the minimum in $\|F_{\text{pf}}\|$ occurs near $\varepsilon = 0.3$. We also observe that at this choice of ε that $\|F_{\text{poly}}\|$ only increases very slowly with the order of the polynomial n , indicated that even for very small polynomials most of the ultraviolet physics is already being captured.

Figure 2 shows the results for the non-Hermitian Chebyshev polynomial, with $d = 1$ fixed (corresponding to the centre of the spectrum of D_w), varying c . It is clear that the minimum in $\|F_{\text{pf}}\|$ occurs at $c = 0$, which is a degenerate case, corresponding a Hermitian Chebyshev polynomial with $\varepsilon = 1$. Hence we conclude that for $d = 1$ there is no advantage to the non-Hermitian approximation.

The left-hand plot of Figure 3 shows the relative size of the pseudofermion force terms for the unfiltered and filtered cases for different polynomial orders (choosing $\varepsilon = 0.3$ and $c = 0$). We see that the Hermitian Chebyshev filter achieves a reduction of a factor of 3 for $n = 4$ compared to the standard case, improving to a factor of 6 at $n = 16$.

The right hand plot of Figure 3 shows that the number of iterations required to invert the (filtered) fermion matrix is also reduced by tuning ε , and we see that for $n = 16$ the choice $\varepsilon = 0.3$ is again optimal. Although for smaller n the choice of $\varepsilon < 0.3$ reduces the number of CG iterations slightly, this would be a poor choice due to being a less effective UV filter.

Finally, we performed HMC simulations using the Hermitian Chebyshev filter to gain a feel for the effectiveness of the filter in an actual calculation. Results are shown in Table 1. The results show that with a polynomial filter one can significantly reduce the step-size while maintaining a high acceptance rate.

Action	n_{poly}	ε	$\Delta\tau$	n_{md}	m	ρ_{accept}
Std.	0	0.0	0.02	50	2	0.87
Poly-H.	6	0.3	0.05	20	10	0.87
Poly-H.	6	0.3	0.05	20	5	0.84

Table 1: Simulation results using standard and polynomial filtered HMC. Shown are the order of the polynomial n_{poly} , choice of ε , step size for the pseudofermions $\Delta\tau$, the number of pseudofermion integration steps n_{md} , the number of inner integration steps per pseudofermion step m and finally the acceptance rate ρ_{accept} .

4. Conclusions

Dynamical light quark simulations are the next major hurdle for Lattice QCD. The use of a polynomial approximation to the inverse as a filter successfully separates the UV and IR pseudofermion dynamics. This enables the use of a multiple time scale integration to reduce the cost of dynamical simulations. We tested both Hermitian and non-Hermitian Chebyshev approximations, and saw the Hermitian version is the most effective. Preliminary simulation results are promising, and a detailed investigation will be the subject of future work.

References

- [1] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, *Phys. Lett.* **B195** (1987) 216–222.
- [2] J. C. Sexton and D. H. Weingarten, *Hamiltonian evolution for the hybrid Monte Carlo algorithm*, *Nucl. Phys.* **B380** (1992) 665–678.
- [3] M. Luscher, *A new approach to the problem of dynamical quarks in numerical simulations of lattice QCD*, *Nucl. Phys.* **B418** (1994) 637–648, [hep-lat/9311007].
- [4] R. Frezzotti and K. Jansen, *A polynomial hybrid monte carlo algorithm*, *Phys. Lett.* **B402** (1997) 328–334, [hep-lat/9702016].
- [5] **TrinLat** Collaboration, M. J. Peardon and J. Sexton, *Multiple molecular dynamics time-scales in hybrid Monte Carlo fermion simulations*, *Nucl. Phys. Proc. Suppl.* **119** (2003) 985–987, [hep-lat/0209037].
- [6] A. Borici and P. de Forcrand, *Systematic errors of Luscher’s fermion method and its extensions*, *Nucl. Phys.* **B454** (1995) 645–662, [hep-lat/9505021].

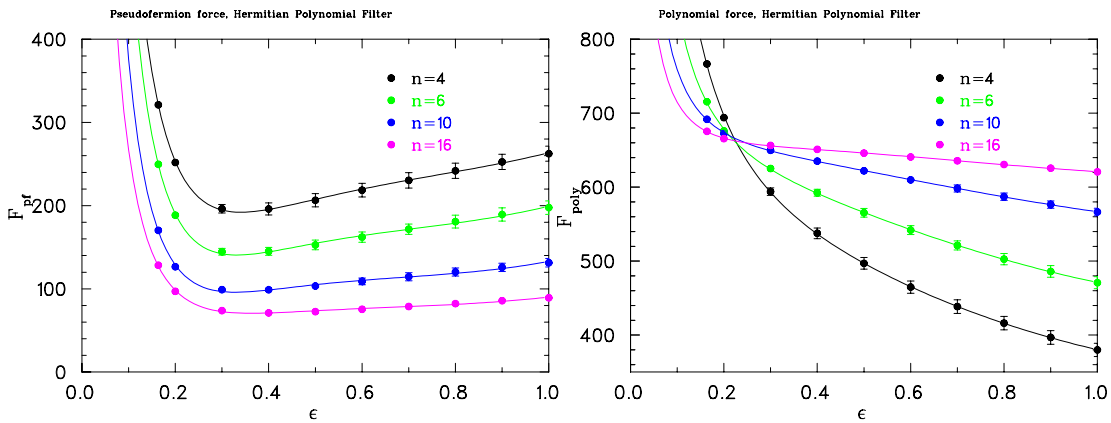


Figure 1: Size of the pseudofermion (left) and polynomial (right) force terms as a function of ϵ for a Hermitian Chebyshev filter.

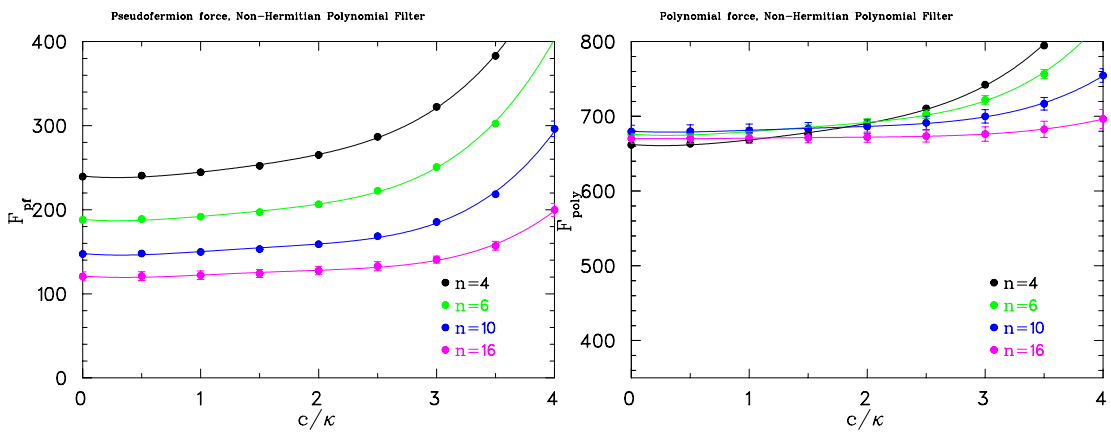


Figure 2: Size of the pseudofermion (left) and polynomial (right) force terms as a function of c for a non-Hermitian Chebyshev filter.

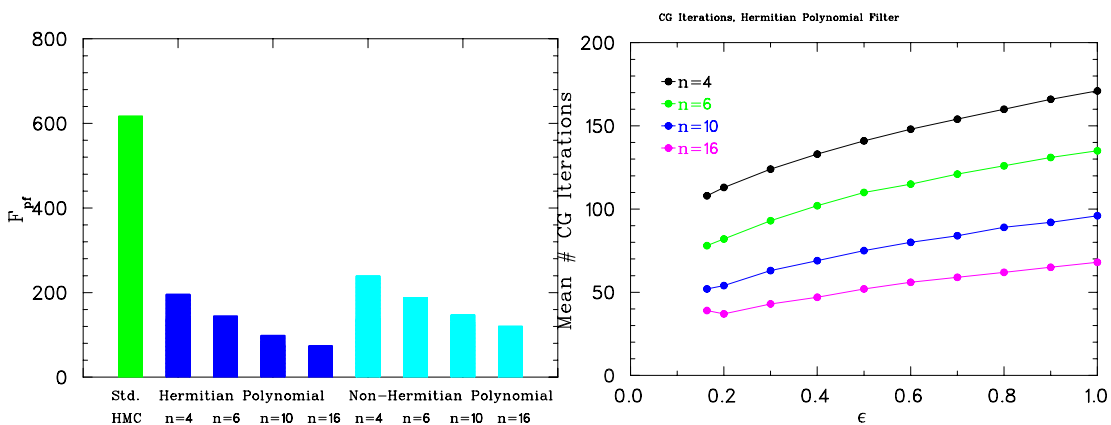


Figure 3: (Left) Comparative size of the pseudofermion force term for standard HMC and for different orders n of Hermitian and non-Hermitian Chebyshev filter (for $\epsilon = 0.3$ and $c = 0$) (Right) Mean number of CG iterations as a function of ϵ for a Hermitian Chebyshev filter.

POS(LAT2005)106