The HMC-algorithm: Schwarz-preconditioning with a one-dimensional domain decomposition

Martin Hasenbusch

Dipartimento di Fisica, Università di Pisa and I.N.F.N.

Lattice 2007, 2 August 2007
Overview

M. Hasenbusch, in preparation

- Definition of the Model
- Improved Pseudo-fermions, Schwarz-preconditioning
- Iterated decomposition, (overlapping blocks)
- Numerical Results
- Conclusions
Lattice QCD:
- 4 dimensional hyper-cubic lattice
- $x$ sites of the lattice, $\mu$ direction
- Gauge field $U_{x,\mu} \in SU(3)$ lives on the link $(x, \mu)$
- quark fields live on the sites

Wilson gauge action

$$S_G[U] = -\frac{\beta}{3} \sum_x \sum_{\mu > \nu} \text{Re} \text{Tr} \left( U_{x,\mu} U_{x+\hat{\mu},\nu} U_{x+\hat{\nu},\mu}^\dagger U_{x,\nu}^\dagger \right)$$

Wilson fermions

$$M[U]_{xy} = 1 - \sum_{\mu} \left\{ (1 - \gamma_\mu) \ U_\mu(x) \ \delta_{x+\hat{\mu},y} + (1 + \gamma_\mu) \ U_\mu^\dagger(x - \hat{\mu}) \ \delta_{x-\hat{\mu},y} \right\}$$
QCD according to the particle data book:

\[ m_u = 1.5 \div 4.5 \text{ MeV} \quad m_d = 4 \div 8.5 \text{ MeV} \quad m_s = 80 \div 155 \text{ MeV} \]
\[ m_c = 1 \div 1.4 \text{ GeV} \quad m_b = 4 \div 4.5 \text{ GeV} \quad m_t = 174.3 \pm 5.1 \text{ GeV} \]

We simulate: Two degenerate flavours of dynamical quarks

\[ m_u = m_d \quad \text{finite} \quad m_s = m_c = m_b = m_t = \infty \]

Until quite recently \( m_\pi / m_\rho > 0.5 \) in large scale simulations (e.g. UKQCD, CPPACCS) of Wilson fermions at \( a \approx 0.1 \text{fm} \).
Physical point \( m_\pi / m_\rho \approx 0.18 \)

\[ Z = \int \mathcal{D}[U] \exp(-S_G[U]) \det M[U]^2 \]
Hybrid Monte Carlo (HMC) Algorithm

determinant is too expensive ( $\propto \text{Volume}^3$) $\rightarrow$ pseudo-fermions

$$\det M^2 = \det MM^\dagger \propto \int D[\phi^\dagger] \int D[\phi] \exp(-|M^{-1}\phi|^2)$$

Problem: $S_F = |M^{-1}\phi|^2$ is non-local

$\implies$ molecular dynamics evolution of all gauge fields.
Introduce conjugate momenta $P$ for the gauge field

$\implies$ Hamiltonian:

$$H(U, \phi, \phi^\dagger, P) = S_G(U) + S_F(U, \phi, \phi^\dagger) + \frac{1}{2} \sum_{x, \mu} \text{Tr} P_{x, \mu}^2$$
What is the problem?

The simulation becomes more expensive as the quark mass becomes smaller, (Lattice 2001, Berlin “Berlin wall”):

\[ \text{cost} = m_{PS}^{-2.8(2)} \quad \text{(for } \beta = 5.6, \text{ Lippert 2001)} \]

- Condition number of \( M \) increases
  \[ \implies \text{solver needs more iterations} \]

- step size must be decreased to get constant acceptance
Improved pseudo-fermions

Introduce $N$ (in practice $N = 2, 3$) matrices $W_i$ such that

$$ M = \prod_{i=1}^{N} W_i $$

The $W_i$ should have a smaller condition number than $M$.

Introduce pseudo-fermions for each $W_i$.

$$ \det MM^\dagger \propto \int D[\phi_1^\dagger] \int D[\phi_1] \ldots \int D[\phi_N^\dagger] \int D[\phi_N] \exp(- \sum_{i=1}^{N} |W_i^{-1}\phi_i|^2) $$
- **Mass-preconditioning** (Hasenbusch 2001):

\[
\begin{align*}
W_1 & = M + \rho_1 \\
W_i & = (M + \rho_{i-1})^{-1}(M + \rho_i) \quad \text{for } 1 < i < N \\
W_N & = (M + \rho_{N-1})^{-1}M
\end{align*}
\]

Alternative (R. Sommer; Hasenbusch and Jansen 2003): add \( \rho_i \gamma_5 \)

- **Polynomial splitting** (Peardon 2001)

formally the same splitting as in PHMC, however low order polynomial

\[
W_1 = P(M) \approx M^{-1}
\]

Easy to combine with “UV-filtering” or PHMC. (not done yet?)
- **RHMC** (Clark, Kennedy 2004): take the $n^{th}$ root of $M$, introduce a pseudo-fermion field for each of the roots. Technically done with a rational approximation. Requires a multi-mass solve for each of the roots.

- **Schwarz preconditioned HMC** (Lüscher 2004)
  M.H. and K. Jansen (2003) conclusions and outlook:
  ”... For example, one could divide the lattice in sub-lattices and construct the matrix $W$ by eliminating all hopping terms from $\hat{Q}$ that connect different sub-lattices. Such a construction might be particularly useful for a massively parallel computer. ...”
Figure taken from M. Lüscher, Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD, hep-lat/0409106, Comput.Phys.Commun. 165 (2005) 199

Checkerboard (red-black) decomposition of the block-lattice in two subsets $A$ and $B$

$$M = \begin{pmatrix} M_A & M_{A\leftarrow B} \\ M_{B\leftarrow A} & M_B \end{pmatrix}$$

$$\det M = \det M_A \det M_B \det(1 - M_A^{-1}M_{A\leftarrow B}M_B^{-1}M_{B\leftarrow A})$$
The remainder:

\[ \tilde{W}_2 = 1 - M^{-1}_A M_{A \leftarrow B} M^{-1}_B M_{B \leftarrow A} \]

acts only on fields that live on the inner boundary of \( A \)
Hence: \( \det \tilde{W}_2 = \det W_2 \), where

\[ W_2 = 1 - P_{\delta A} M^{-1}_A M_{A \leftarrow B} M^{-1}_B M_{B \leftarrow A} \]

\( P_{\delta A} \): Projection on the interior boundary of \( A \);
Further reduction due to the structure of the \( \gamma \) matrices.

 Needed for the pseudo-fermion action:

\[ W_2^{-1} = 1 - P_{\delta A} M^{-1} M_{B \leftarrow A} \]
Only active links are updated

Both end-points in one block, at maximum one endpoint in the interior boundary

\[ \Rightarrow \text{decoupling with respect to } S_G \text{ and } S_{F1}, \]

avoid the update of links with the largest force with respect to \( S_{F2} \)

<table>
<thead>
<tr>
<th>block</th>
<th># active links / # all links</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 6^4 )</td>
<td>0.247...</td>
</tr>
<tr>
<td>( 8^4 )</td>
<td>0.369...</td>
</tr>
<tr>
<td>( 12^4 )</td>
<td>0.530...</td>
</tr>
<tr>
<td>( 12 \times 8 \times 6^2 )</td>
<td>0.331...</td>
</tr>
</tbody>
</table>
Forces at the end of the trajectory for: $k = 0$ gauge field, $k = 1$ fermion matrix restricted to blocks, $k = 2$ remainder of the fermion determinant.

Figure taken from M. Lüscher, Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD, hep-lat/0409106, Comput. Phys. Commun. 165 (2005) 199.
One-dimensional decomposition (in many Schrödinger functionals)

- sufficient to reduce the condition number of the fermion matrix
- simplifies the implementation
- allows for iteration of the decomposition (and overlapping) of domains
Iteration of the one-dimensional decomposition:
Implementation

- Division of the lattice into 2 blocks
- Sub-division of each block into 2 sub-blocks
- Mass-preconditioning of the fermion matrix restricted to the sub-blocks

Integration scheme: (trajectory of length $t = 1$)
- **Multiple time scale** integration (Sexton Weingarten 1992)
- **Leap-frog** for $S_{F2}$, $S_{F3}$ and $S_{F4}$
- **Sexton Weingarten improved scheme**, second order minimal norm (2MN) scheme (de Forcrand Takaishi 2005) for $S_G$ and $S_{F1}$

Solver:
- **even-odd preconditioning**
- $level = 4$ BiCGstab solver
- $level < 4$ geometric series (special case of Chebyshev iteration)
step size depending on $x_0$

$$dt(x_0) = \lambda(x_0) \Delta t$$

Scheme

<table>
<thead>
<tr>
<th>Scheme</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$0.5$</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1.5</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Both cases: Fraction of active links $0.59375$
Numerical test

- Wilson gauge action $\beta = 5.6 \rightarrow a \approx 0.08 \text{ fm}$

- Wilson fermions with $\kappa = 0.1575, 0.1580, 0.15825$ corresponds to a pion mass $am_\pi \approx 0.28, 0.20, 0.15$ or ($\approx 690 \text{ MeV}, 490 \text{ MeV}, 370 \text{ MeV}$)
  Real world: $m_\pi \approx 135 \text{ MeV}$

- Lattice size: $32 \times 16^3$ and $32 \times 24^3$

Forces

$24^3 \times 32 \beta=5.6 \ \kappa=0.15825$

average on spacial links

$24^3 \times 32 \beta=5.6 \ \kappa=0.15825$

average and maximal force; $n=4$

The HMC-algorithm: Schwarz-preconditioning with a one-dimensional domain decomposition

Martin Hasenbusch
stepsizes, acceptance rates ... 

<table>
<thead>
<tr>
<th>$L$</th>
<th>$\kappa$</th>
<th>$\rho$</th>
<th>$\Delta t$</th>
<th>$n_4$</th>
<th>$n_3$</th>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n_0$</th>
<th>$P_{acc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>A 0.1575</td>
<td>0.15</td>
<td>0.16..</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.892(2)</td>
</tr>
<tr>
<td>12</td>
<td>A 0.1580</td>
<td>0.15</td>
<td>0.16..</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.916(2)</td>
</tr>
<tr>
<td>16</td>
<td>A 0.1575</td>
<td>0.20</td>
<td>0.25</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.704(5)</td>
</tr>
<tr>
<td>16</td>
<td>A 0.1580</td>
<td>0.15</td>
<td>0.2</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.826(4)</td>
</tr>
<tr>
<td>16</td>
<td>A 0.15825</td>
<td>0.15</td>
<td>0.2</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.826(4)</td>
</tr>
<tr>
<td>24</td>
<td>A 0.15825</td>
<td>0.15</td>
<td>0.14..</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.83(2)</td>
</tr>
<tr>
<td>24</td>
<td>B 0.15825</td>
<td>0.15</td>
<td>0.2</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.875(4)</td>
</tr>
</tbody>
</table>

For Comparison:
GRAL $\kappa = 0.1580$, $32 \times 16^3$: $\Delta t = 0.006$, $P_{acc} = 0.66$
M. Lüsher, $\kappa = 0.15825$, $32 \times 24^3$: $\Delta t = 0.05$. $P_{acc} = 0.86$
Urbach et al, $\kappa = 0.15825$, $32 \times 24^3$: $\Delta t = 0.1$. $P_{acc} = 0.8$
(Clark-Kennedy similar in performance)
Autocorrelation times?

![Graphs showing autocorrelation times](image)

Red line: value of the plaquette Del Debbio et al.; M. Lüscher private communication.
\[ \tau_{plaq} = 8(2), \quad \tau_{solv} = 16(5) \]

Comparison:
SESAM: \( \tau_{plaq} = 7(4), \quad \tau_{solv} = 18(6) \).
Lüscher: \( \tau_{plaq} = 68(25), \quad \tau_{solv} = 168(42) \).
Conclusions and Outlook

- Problem of decreasing step-size with decreasing mass is eliminated
- Performance of mass-preconditioning, (also polynomial preconditioning), RHMC and Schwarz-preconditioning seems similar (does this depend on the lattice spacing $a$?)
- mass-preconditioning, Polynomial preconditioning, RHMC simpler to implement, no restrictions on the geometry, easy to adopt to more complicated actions