Lattice QCD simulation with 2+1 flavors of dynamical overlap fermions

Hideo Matsufuru for JLQCD Collaboration
(with H.Fukaya, S.Hashimoto, K.Kanaya, T.Kaneko, M.Okamoto, T.Onogi, and N.Yamada)

Joint Meeting of Pacific Region Particle Physics Communities
29 October – 3 November 2006, Hawaii, USA
JLQCD’s overlap project

Dynamical simulation with overlap fermions

- Main run: $16^3 \times 32$, $a \simeq 0.12\text{fm}$ (larger size is planned)
- Lightest quark mass $\simeq m_s/6$
- Fixed topology by extra Wilson fermion
  - need to examine the effect of fixing topology
- $N_f = 2$ is now in productive run
- $N_f = 2 + 1$ is in progress

Overview/results at $N_f = 2 \rightarrow$ T.Kaneko’s talk

In this talk:

- Algorithms of solver and HMC
- $N_f = 2 + 1$ simulation
New machines at KEK

Working since March 2006

Hitachi SR11000
- 2.15 TFlops, 512 MB memory
- 16 Power5+ × 16 nodes

IBM System Blue Gene Solution
- 57.3 TFlops, 5 TB memory
- 1024 nodes × 10 racks
- 8 × 8 × 8 torus network
- 2 PowerPC440 shares 4 MB cache

Wilson kernel for BG:
Tuned by IBM Japan (J. Doi and H. Samukawa)
- double FPU instructions for complex arithmetics
- low level communication API

Wilson solver: ~29% of peak performance (on cache)
**Action**

\[ S = S_G + S_F + S_E \]

- **Gauge field** \( S_G \): Iwasaki (renormalization group improved)
- **Overlap fermion** \((N_f = 2)\): \( S_F = \phi^\dagger [D(m)^\dagger D(m)]^{-1} \phi \)

  **overlap Dirac operator**

  \[ D(m) = \left( m_0 + \frac{m}{2} \right) + \left( m_0 - \frac{m}{2} \right) \gamma_5 \text{sign}(H_W) \]

  \[ H_W = \gamma_5 D_W, \ D_W \text{ is Wilson-Dirac operator with } -M_0 \]

- **Extra Wilson fermion:**

  \[ \det \left( \frac{H_W^2}{H_W^2 + \mu^2} \right) = \int \mathcal{D} \chi^\dagger \mathcal{D} \chi \exp[-S_E] \]

  — suppresses near-zero modes of \( H_W \)

  *Vranas (2000); Fukaya (2006); S.Hashimoto et al., hep-lat/0610011*
Solver algorithm (1)

Overlap Dirac operator

\[ D(m) = \left( M_0 + \frac{m}{2} \right) + \left( M_0 - \frac{m}{2} \right) \gamma_5 \text{sign}(H_W) \]

Zolotarev’s partial fractional approximation

\[ \text{sign}(H_W) = \frac{H_W}{\sqrt{H^2_W}} = H_W \left( p_0 + \sum_{l=1}^{N} \frac{p_l}{H^2_W + q_l} \right) \]

- \((H^2_W + q_l)^{-1}\): determined by Multishift CG simultaneously
- For smaller \(\lambda_{min}\), larger \(N\) is needed for accuracy
  - e.g. for \(N=10\), \(O(10^{-7})\) accuracy for \(\lambda_{min}=0.05\) and \(O(10^{-5})\) for 0.01.
- Subtraction of low modes of \(H_W\)
  - \(\text{sign}(\lambda)\) (\(\lambda < \lambda_{thr}\)) is explicitly determined
Solver algorithm (2)

- Nested CG algorithm
  - Outer CG for $D(m)$, inner CG for $(H_W^2 + q_l)^{-1}$ (multishift)
  - Relaxed CG: $\varepsilon_{in}$ is relaxed as outer iteration proceeds
  - Subtraction of low-modes of $H_W$ applicable (safe from $\lambda_{min} \sim 0$)
  - Cost is almost unchanged as $N$

- 5-dimensional CG
  - Making use of Schur decomposition
  - Even-odd preconditioning
  - Cost increases linearly in $N$
  - Subtraction of low-modes of $H_W$ is not applicable
    $\rightarrow$ difficulty at $\lambda_{min} \sim 0$
Solver algorithm (3)

Comparison:

\((a \simeq 0.12\text{fm}, m \simeq 0.4m_s, \text{single conf.})\)

- Relaxed CG is factor 2 faster than standard CG
- 5D solver is 2-3 times faster than relaxed CG for \(N = 20\)
- If \(\lambda \simeq 0\) does not appear, 5D solver has advantage
HMC algorithm (1)

Building blocks of accelerating HMC:

• Hasenbusch preconditioning: \( S_F = S_{PF_1} + S_{PF_2} \)


  \[
  S_{PF_1} = \phi_1^\dagger [D(m')^\dagger D(m')]^{-1} \phi_1 \quad \text{(preconditioner)}
  \]

  \[
  S_{PF_2} = \phi_2^\dagger \{ D(m') [D(m)^\dagger D(m)]^{-1} D(m')^\dagger \} \phi_2
  \]

• Multi-time step: \( \Delta \tau_{(PF_2)} > \Delta \tau_{(PF_1)} > \Delta \tau_{(G)} = \Delta \tau_{(E)} \)


• Overlap solver: relaxed CG/5D CG

• Reflection/refraction at \( \lambda_{min} = 0 \)


  – Needs monitoring of \( \lambda_{min} \) and inverting \( D^\dagger D \) twice

  \( \Rightarrow \) skipped: \( \lambda_{min} = 0 \) is avoided by \( S_E \)
HMC algorithm (2): Noisy Metropolis

Most time consuming part: solvers in molecular dynamics
Cost in MD is reduced by

- assuming no near-zero mode
- fixed $\lambda_{\text{thrs}}$, $N \simeq 10 \rightarrow$ adopting 5D solver
- no eigenvalue determination

Error in MD is corrected by Noisy Metropolis:


After usual Metropolis, accept $U_{\text{new}}$ with $P = \min\{1, e^{-dS}\}$,

$$dS = | W^{-1}[U_{\text{new}}]W[U_{\text{old}}] \xi |^2 - |\xi|^2$$

where $W = D(m)/D'(m)$,

- $D'$: relaxed overlap operator used in MD
- $D$: accurate overlap operator
Performance of $N_f=2$ simulations

Performance on Blue Gene (512-node)

$a \sim 0.12 \text{fm}, \mu = 0.2, \text{trajectory length: } \tau = 0.5$

- **HMC-1**: With 4D (relaxed CG) solver

<table>
<thead>
<tr>
<th>$m_{ud}$</th>
<th>$N_{\tau(PF2)}$</th>
<th>$\frac{\Delta \tau_{(PF2)}}{\Delta \tau_{(PF1)}}$</th>
<th>$\frac{\Delta \tau_{(PF1)}}{\Delta \tau_{(G,E)}}$</th>
<th>$N_{PF1,2}$</th>
<th>$P_{acc}$</th>
<th>time[min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.015</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>0.87</td>
<td>112</td>
</tr>
<tr>
<td>0.025</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>0.90</td>
<td>94</td>
</tr>
<tr>
<td>0.035</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>0.74</td>
<td>63</td>
</tr>
</tbody>
</table>

- **HMC-2**: less precise 5D solver in MD + noisy Metropolis

  $\rightarrow$ factor $\sim 2$ accelerated

<table>
<thead>
<tr>
<th>$m_{ud}$</th>
<th>$N_{\tau(PF2)}$</th>
<th>$\frac{\Delta \tau_{(PF2)}}{\Delta \tau_{(PF1)}}$</th>
<th>$\frac{\Delta \tau_{(PF1)}}{\Delta \tau_{(G,E)}}$</th>
<th>$N_{PF1}$</th>
<th>$N_{PF2}^{(MD)}$</th>
<th>$N_{PF2}^{(NM)}$</th>
<th>$P_{acc}$</th>
<th>time[min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.015</td>
<td>13</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>10</td>
<td>0.68</td>
<td>52</td>
</tr>
<tr>
<td>0.025</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>10</td>
<td>0.82</td>
<td>43</td>
</tr>
<tr>
<td>0.035</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>10</td>
<td>0.87</td>
<td>36</td>
</tr>
</tbody>
</table>
\[ N_f = 2 + 1 \text{ algorithm (1)} \]

A. Bode et al., hep-lat/9912043

\[ H^2 = D\dagger(m)D(m) \text{ commutes with } \gamma_5 \]

\[ H^2 = P_+ H^2 P_+ + P_- H^2 P_- \equiv Q_+ + Q_- \]

\[ \det H^2 = \det Q_+ \cdot \det Q_- \]

Eigenvalues of \( Q_+ \) and \( Q_- \) are the same except for zero modes

\[ \Downarrow \]

One of chirality sector realizes odd number of flavor
(zero modes give const. contribution)

- Topology change can be implemented
  — Not necessary in our case
\( N_f = 2 + 1 \) algorithm (2)

Pseudofermion action (\( \sigma = 1 \) or \(-1\)):

\[
S_{PF1} = \phi_{1\sigma} \left( Q_{\sigma}^{-1}(m') \right) \phi_{1\sigma}, \quad S_{PF2} = \phi_{2\sigma} \left( \frac{Q_{\sigma}(m')}{Q_{\sigma}(m)} \right) \phi_{2\sigma}
\]

- Refreshing \( \phi_{1\sigma} \) and \( \phi_{2\sigma} \) (with Gaussian \( \xi_{\sigma} \))

\[
\phi_{1\sigma} = \sqrt{Q_{\sigma}(m')} \cdot \xi_{1\sigma}, \quad \phi_{2\sigma} = \sqrt{\frac{Q_{\sigma}(m)}{Q_{\sigma}(m')}} \cdot \xi_{2\sigma}.
\]

--- Polynomial or partial fractional approx.

- Other parts are straightforward
  
  e.g., force:

\[
\frac{dS_{PF1}}{d\tau} = \phi_{1\sigma} P_{\sigma} \left( \frac{dH^2(m')^{-1}}{d\tau} \right) P_{\sigma} \phi_{1\sigma}
\]

etc.
Check: $N_f=2$ vs $N_f=1+1$ (1)

$16^3 \times 32$ lattice, $\beta = 2.5$, $m_q = 0.09$

- Two positive chirality PS-fermions
- HMC-1 (4D solver, w/o noisy Metropolis)
  — compared with $N_f = 2$, HMC-1
- Initial: $N_f = 2$ thermalized config.
- $M_{MD}^{(pf2)} = 4$, $R_{MD}^{(pf1)} = 5$, $R_{MD}^{(GE)} = 6$, $l_{trj} = 0.5$, $m' = 0.4$

<table>
<thead>
<tr>
<th></th>
<th>trj</th>
<th>plaq</th>
<th>$P_{acc}$</th>
<th>min/trj(BG 512 node)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nf=1+1</td>
<td>1500</td>
<td>0.651219(16)</td>
<td>$\sim 0.8$</td>
<td>23</td>
</tr>
<tr>
<td>Nf=2</td>
<td>1000</td>
<td>0.651173(21)</td>
<td>0.81</td>
<td>13</td>
</tr>
</tbody>
</table>

- Consistent with $N_f = 2$.
- Increased cost: largely due to refreshment of $\phi$’s
  (Now Zolotarev approx. is used)
\( N_f=2 \) vs \( N_f=1+1 \) (2): solver

For \( Q_\sigma \), number of \( H_W \) mult is effectively half of \( H^2 \).

\[
P_\sigma H^2 P_\sigma = P_\sigma \left[ a + \frac{b}{2} \{\gamma_5, \text{sign}(H_W)\} \right] P_\sigma = P_\sigma \left[ a + \sigma b \cdot \text{sign}(H_W) \right] P_\sigma
\]

— Costs of \( N_f = 1 + 1 \) and \( N_f = 2 \) are comparable
$N_f=2$ vs $N_f=1+1$ (2): force

Total forces are similar to $N_f = 2$
— Same HMC parameters are applicable

$N_f = 1 + 1$

$N_f = 2$
Test run: $N_f=2+1$ (1)

$\beta = 2.30, \ m_{ud} = 0.10, \ m_s = 0.10, \ Q = 0$

- $N_f = 2 \oplus$ positive chirality sector
- Other parameters are same as $N_f = 2$
- HMC-1 (4D solver, w/o noisy Metropolis)
- $M_{MD}^{(p^f_2)} = 5, \ R_{MD}^{(p^f_1)} = 5, \ R_{MD}^{(GE)} = 6, \ l_{trj} = 0.5, \ m' = 0.4$
- Thermalization: 300 trjs (very preliminary)

<table>
<thead>
<tr>
<th></th>
<th>trj</th>
<th>plaq</th>
<th>$P_{acc}$</th>
<th>time/trj (BG 512 node)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_f=2+1$</td>
<td>150</td>
<td>0.609724(50)</td>
<td>~0.76</td>
<td>70 min</td>
</tr>
<tr>
<td>$N_f=2$</td>
<td>4600</td>
<td>0.614685(12)</td>
<td>0.85</td>
<td>40 min</td>
</tr>
</tbody>
</table>
$N_f=2+1$ (2): solver/force

Solver convergence:
One flavor part is twice faster than $N_f = 2$
→ total cost is $\sim 1.5$ times

Force hierarchy:
Total forces of 2+1 flavors are similar to $N_f = 2, 1+1$
$N_f=2+1$ (3): $\beta$ shift

$\beta = 2.30$, $m_{ud} = m_s = 0.10$, $Q = 0$

Very preliminary result
- 300 thermalization trjs.
- 30 configs (5 trj separated)

$\alpha$ is determined by hadronic radius (Sommer scale)
— tendency consistent with $N_f = 2$

\[
\begin{array}{c|c|c|c|c|c|c|c|c}
\beta & 1.6 & 1.8 & 2.0 & 2.2 & 2.4 & 2.6 & 2.8 \\
\hline
m_{ud} (= m_s) & 0.04 & 0.06 & 0.08 & 0.10 & 0.12 & 0.14 & 0.16 \\
\end{array}
\]

- $N_f=0$, Iwasaki (CP-PACS, 2004)
- $N_f=0$, Iwasaki + extra Wilson
- $N_f=2$, Iwasaki + ex Wilson + overlap
- $N_f=2$, Iwasaki + clover (CP-PACS, 2002)
Summary/Outlook

JLQCD’s dynamical overlap project

$N_f = 2$ is now in productive run at $16^3 \times 32$, $a \simeq 0.12\text{fm}$, $\simeq m_s/6$

- Best solution: less precise 5D solver $\oplus$ Noisy Metropolis

We are preparing for $N_f = 2 + 1$ simulations

Improvement and parameter tuning are in progress

- 5D CG solver/Noisy Metropolis
- PS-fermion refreshment
- Tuning of HMC parameters (trajectory length, etc)