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: $16^3 \times 32$, $a \simeq 0.12$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

- 2.15TFlops, 512MB memory
- 16 Power5+ $\otimes$16 nodes

IBM System Blue Gene Solution

- 57.3TFlops, 5TB memory
- 1024 nodes $\otimes$10 racks
- $8 \times 8 \times 8$ torus network
- 2 PowerPC440 shares 4MB 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: $\sim$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

$$D(m) = \left(m_0 + \frac{m}{2}\right) + \left(m_0 - \frac{m}{2}\right) \gamma_5 \mathrm{sign}(H_W)$$

  $H_W = \gamma_5 D_W$, D_W 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

$$D(m) = \left(m_0 + \frac{m}{2}\right) + \left(m_0 - \frac{m}{2}\right)\gamma_5 \mathrm{sign}(H_W)$$

Zolotarev's partial fractional approximation

J. van den Eshof et al., Comp. Phys. Comm. 146 (2002) 203.

$$\mathrm{sign}(H_W) = \frac{H_W}{\sqrt{H_W^2}} = H_W\left(p_0 + \sum_{l=1}^{N} \frac{p_l}{H_W^2 + q_l}\right)$$

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

Solver algorithm (2)

☐ Nested CG algorithm

- Outer CG for $D(m)$, inner CG for $(H_W^2 + q_l)^{-1}$ (multishift)
  A.Frommer et al., Int. J. Mod. Phys. C 6 (1995) 627.
- Relaxed CG: ϵ_{in} is relaxed as outer iteration proceeds
  N.Cundy et al., Comp. Phys. Comm. 165 (2004) 221.
- Subtraction of low-modes of H_W applicable (safe from $\lambda_{min} \sim 0$)
- Cost is almost unchanged as N

☐ 5-dimensional CG

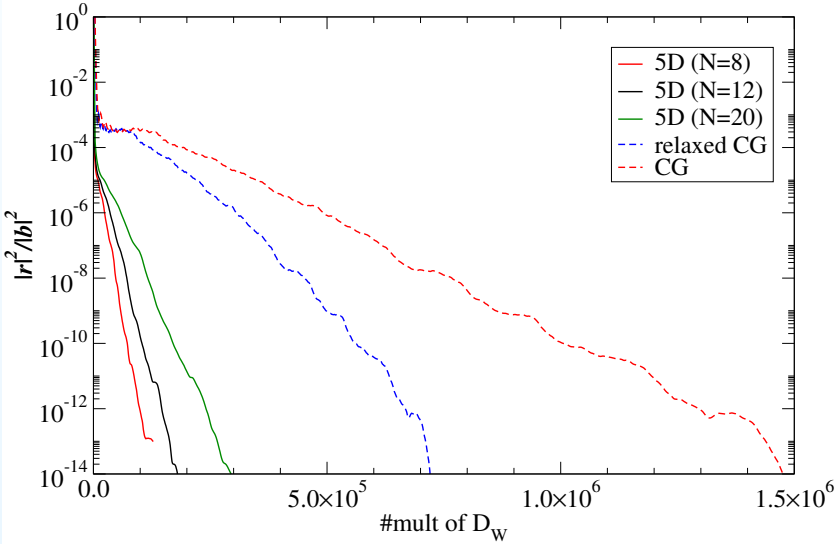
A. Borici, hep-lat/0402035; R.G.Edwards et al., PoS LAT2005 (2006) 146.

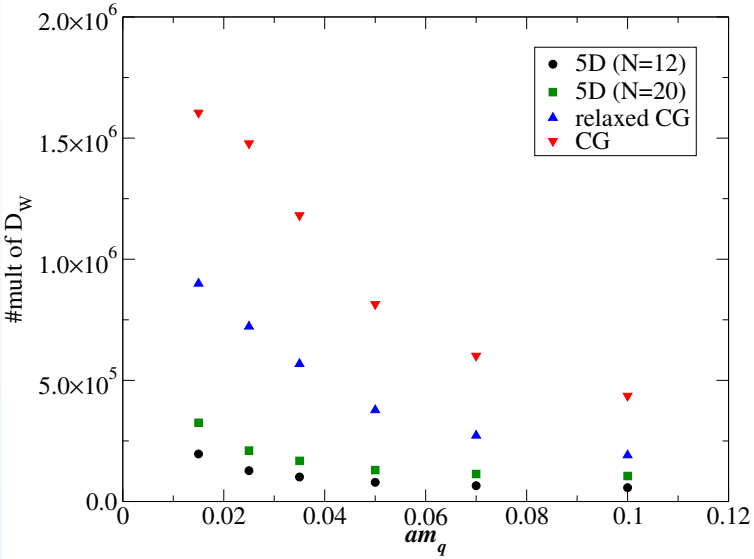
- Making use of Schur decomposition
- Even-odd preconditioning
- Cost increases linearly in N
- Subtraction of low-modes of H_W is not applicable
  $\rightarrow$ difficulty at $\lambda_{min} \sim 0$

Solver algorithm (3)

Comparison:

$(a \simeq 0.12\mathrm{fm}, m \simeq 0.4m_s$, *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: $S_F = S_{PF1} + S_{PF2}$

  M.Hasenbusch, Phys. Lett. B 519 (2001) 177.

$$
\begin{aligned}
S_{PF1} &= \phi_1^\dagger [D(m')^\dagger D(m')]^{-1} \phi_1 \quad \text{(preconditioner)} \\
S_{PF2} &= \phi_2^\dagger \left\{ D(m') [D(m)^\dagger D(m)]^{-1} D(m')^\dagger \right\} \phi_2
\end{aligned}
$$

- Multi-time step: $\Delta\tau_{(PF2)} > \Delta\tau_{(PF1)} > \Delta\tau_{(G)} = \Delta\tau_{(E)}$

  J.C.Sexton and D.H.Weingarten, Nucl. Phys. B 380 (1992) 665.

- Overlap solver: relaxed CG/5D CG

- Reflection/refraction at $\lambda_{min} = 0$

  Z.Fodor, S.D.Katz and K.K.Szabo, JHEP0408 (2004) 003.

  – Needs monitoring of λ_{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 λ_{thrs}, $N \simeq 10 \rightarrow$ adopting 5D solver
- no eigenvalue determination

Error in MD is corrected by Noisy Metropolis:

A.D.Kennedy and J.Kuti, Phys. Rev. Lett. 54 (1985) 2473.

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

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

where ξ is Gaussian noise vector, $W = D(m)/D'(m)$,

- D': relaxed overlap operator used in MD
- D: accurate overlap operator

Performance of N_f=2 simulations

Performance on Blue Gene (512-node)

$a \sim 0.12$*fm*, $\mu = 0.2$, *trajectory length:* $\tau = 0.5$

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

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

- HMC-2: less precise 5D solver in MD + noisy Metropolis
  $\longrightarrow$ factor $\sim$2 accelerated

m_{ud}	$N_{\tau(PF2)}$	$\frac{\Delta\tau_{(PF2)}}{\Delta\tau_{(PF1)}}$	$\frac{\Delta\tau_{(PF1)}}{\Delta\tau_{(G,E)}}$	N_{PF1}	$N_{PF2}^{(MD)}$	$N_{PF2}^{(NM)}$	P_{acc}	time[min]
0.015	13	6	8	10	16	10	0.68	52
0.025	10	6	8	10	16	10	0.82	43
0.035	10	6	8	10	16	10	0.87	36

$N_f = 2 + 1$ algorithm (1)

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

T. DeGrand and S. Schaefer, JHEP 0607 (2006) 020

$H^2 = D^\dagger(m)D(m)$ commutes with γ_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):

$$S_{PF1} = \phi_{1\sigma}^\dagger Q_\sigma^{-1}(m')\phi_{1\sigma}, \qquad S_{PF2} = \phi_{2\sigma}^\dagger \left(\frac{Q_\sigma(m')}{Q_\sigma(m)}\right) \phi_{2\sigma}$$

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

$$\phi_{1\sigma} = \sqrt{Q_\sigma(m')} \cdot \xi_{1\sigma}, \qquad \phi_{2\sigma} = \sqrt{\frac{Q_\sigma(m)}{Q_\sigma(m')}} \cdot \xi_{2\sigma}.$$

— Polynomial or partial fractional approx.

- Other parts are straightforward

  e.g., force:

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

  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.
- $M_{MD}^{(pf2)} = 4$, $R_{MD}^{(pf1)} = 5$, $R_{MD}^{(GE)} = 6$, $l_{trj} = 0.5$, $m' = 0.4$

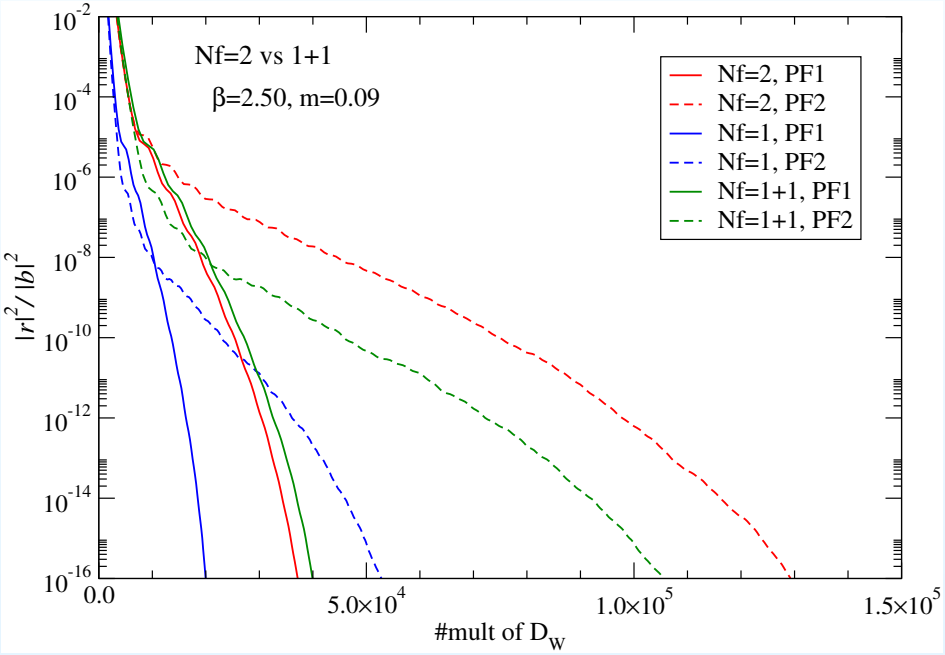
	trj	plaq	P_{acc}	min/trj(BG 512 node)
Nf=1+1	1500	0.651219(16)	$\sim$0.8	23
Nf=2	1000	0.651173(21)	0.81	13

- Consisitent with $N_f = 2$.
- Increased cost: largely due to refreshment of ϕ's
  (Now Zolotarev approx. is used)

N_f=2 vs N_f=1+1 (2): solver

For Q_σ, number of H_W mult is effectively half of H^2.

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



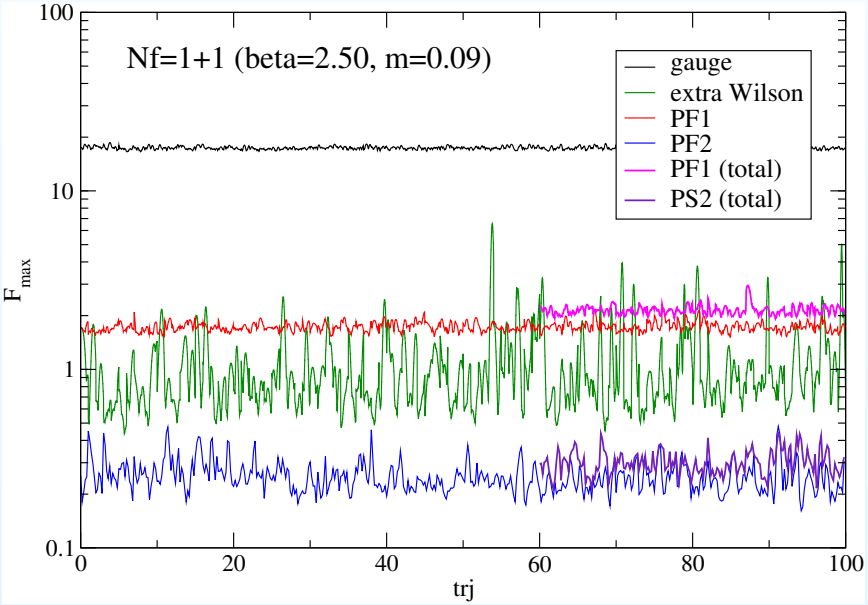
(on different config.)

— Costs of $N_f = 1 + 1$ and $N_f = 2$ are comparable

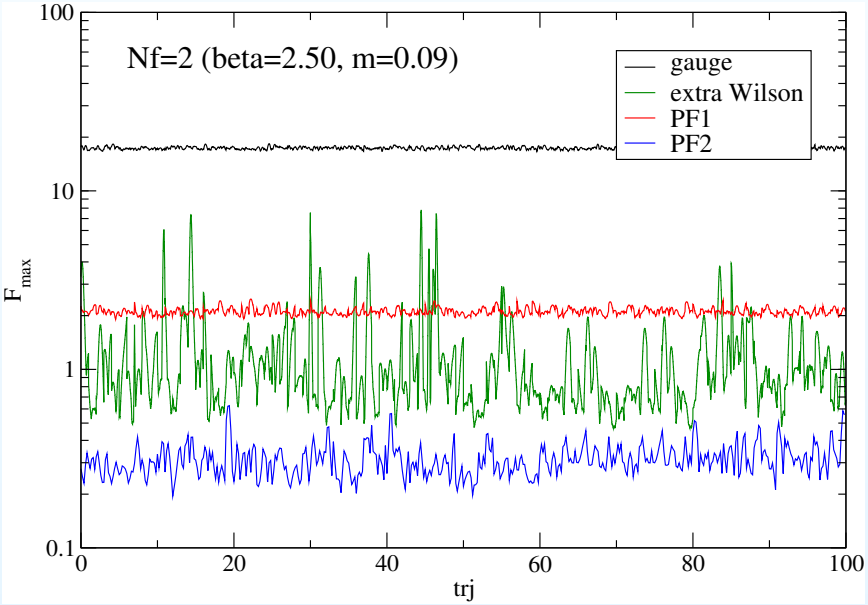
N_f=2 vs N_f=1+1 (2): force

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)
- $M_{MD}^{(pf2)} = 5,\ R_{MD}^{(pf1)} = 5,\ R_{MD}^{(GE)} = 6,\ l_{trj} = 0.5,\ m' = 0.4$
- Thermalization: 300 trjs (very preliminary)

	trj	plaq	P_{acc}	time/trj(BG 512 node)
Nf=2+1	150	0.609724(50)	$\sim$0.76	70 min
Nf=2	4600	0.614685(12)	0.85	40 min

N_f=2+1 (2): solver/force

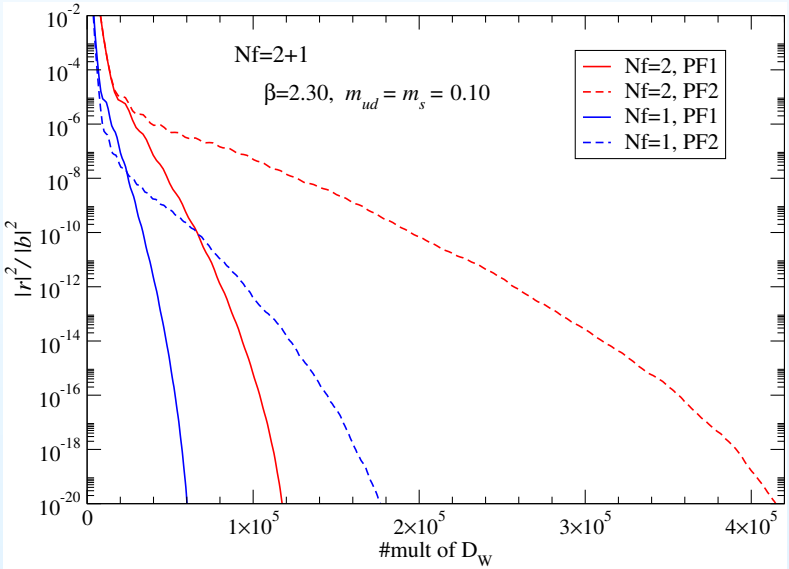
Solver convergence:

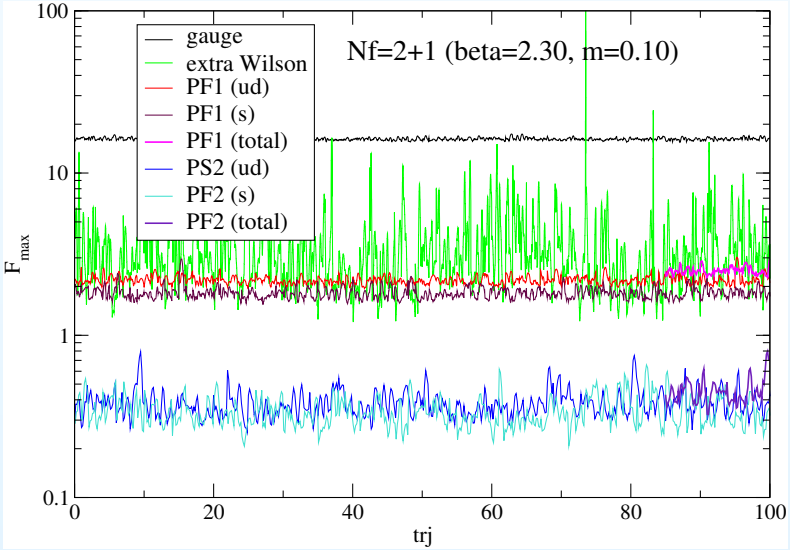
One flavor part is twice faster than $N_f = 2$

$\rightarrow$ total cost is $\sim$ 1.5 times

Force hierarchy:

Total forces of 2+1 flavors are similar to $N_f = 2$, 1+1





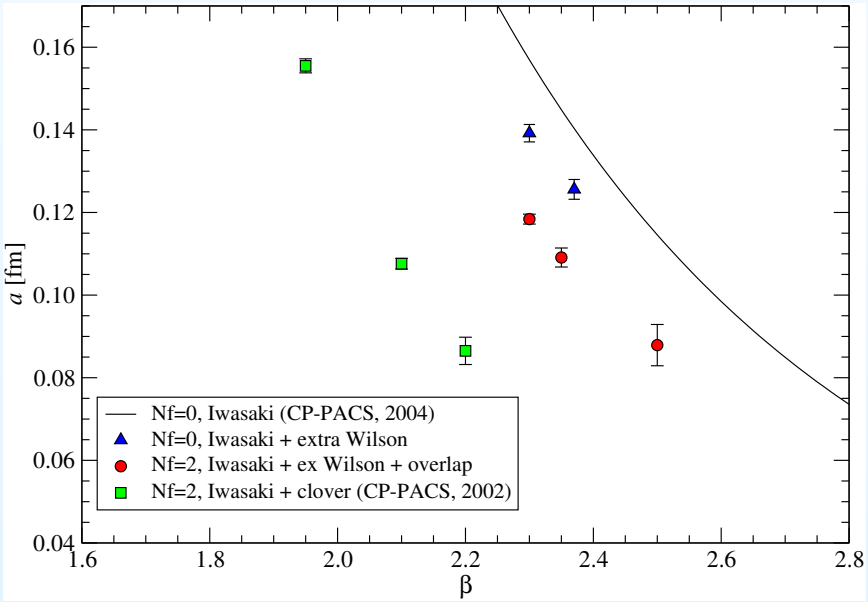
N_f=2+1 (3): β shift

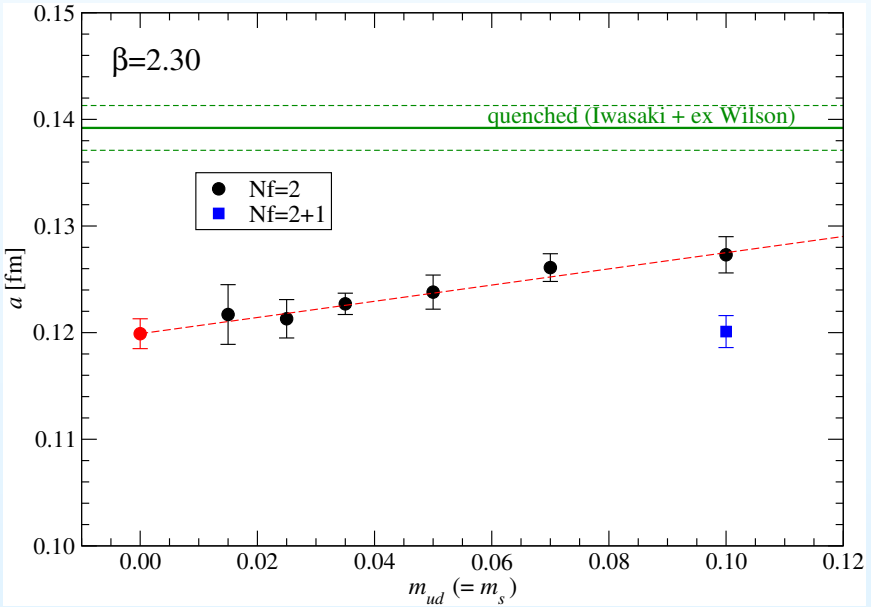
$\beta = 2.30$, $m_{ud} = m_s = 0.10$, $Q = 0$

Very preliminary result

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

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





Summary/Outlook

JLQCD's dynamical overlap project

$N_f = 2$ is now in productive run at $16^3 \times 32$, $a \simeq 0.12$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)