Multigrid in HMC

Meifeng Lin

BROOKHAVEN NATIONAL LABORATORY

Work done with Richard Brower and James Osborn
Partially supported by NSF (BU), SciDAC-3 and ALCF

USQCD SciDAC Software Workshop, Fermilab, Oct. 18-19, 2013

Motivation

- Adaptive Multigrid proved efficient in the Wilson/Clover solver.
  - Babich et al., PRL 2010
  - Osborn et al., PoS 2010.
- HMC involves repeated solving of the Dirac equation
  - In the action (a few)
  - In the force (many)
- Natural to integrate MG solver into HMC

- Project started at Boston University late 2012 under NSF grant

MG solver performance

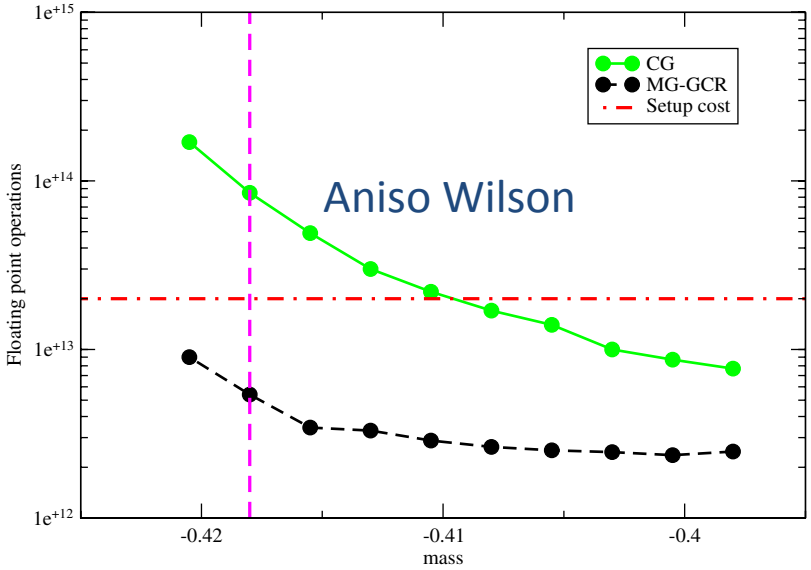


FIG. 2 (color online). Number of floating point operations required to reach convergence for CG and MG-GCR on the $V = 32^3 \times 96$ lattice (parameters given in Fig. 1). The horizontal line indicates the number of floating point operations of the MG setup. Babich et al 2010

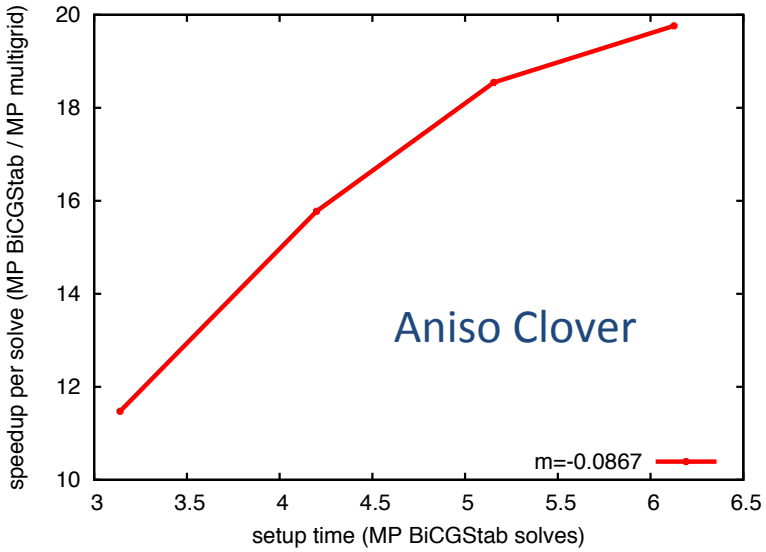


Figure 3: Speedup of multigrid solver relative to BiCGStab versus setup time at the physical quark mass. Osborn et al 2010

- Performance gain depends heavily on the relative setup cost
- Must be able to reuse the setup

Implementation

- Wilson/Clover MG solver available in qopqdp (version >= 0.19.1)
- Integration to HMC done in FUEL
  - Only naïve Wilson HMC is available
  - Clover HMC is next
  - Anisotropy is also implemented
- Gauge field gets updated after every solve in HMC, but is highly correlated over a long MD time.
  - Setup is done at light dynamical mass at beginning of trajectory.
  - Reused in subsequent integration steps and/or MD trajectories until gain is lost
  - Refresh the setup when (trajectory time > setup time + 1st trajectory time)

Challenges for MG-HMC

- Need to compete with modern HMC algorithms
- Hasenbusch mass preconditioning

$$
\begin{aligned}
S_F[U,\{\phi_i^\dagger,\phi_i\}] \;=\; & \phi_0^\dagger \left([W_1^{-1}\hat{M}][W_1^{-1}\hat{M}]^\dagger\right)^{-1}\phi_0 \\
& + \sum_{i=1}^{n-1} \phi_i^\dagger \left([W_{i+1}^{-1}W_i][W_{i+1}^{-1}W_i]^\dagger\right)^{-1}\phi_i \\
& + \phi_n^\dagger (W_n W_n^\dagger)^{-1}\phi_n
\end{aligned}
$$

- Fewer light quark solves, more heavy Hasenbusch-mass solves. MG gains more in light solves.

MG-HMC Tests

- Starting from existing thermalized anisotropic 2-flavor Wilson lattices. (Bulava et al. 2009)
- Apples-to-apples comparison: use the same HMC setup. Simply replace the original solver with MG solver
- Pion mass ~ 420 MeV. Tested on two lattice volumes.
- Run on 32 BG/Q nodes with 32 MPI processes/node at ALCF.

Volume	ξ_0	ν	ξ_{MD}	$\tau^{[\star]}$	n_l	n_H	n_G	stop. cond.
$24^3 \times 64$	2.38	1	2.4	0.707	10	40	240	1e-8

Volume	ξ_0	ν	ξ_{MD}	$\tau^{[\star]}$	n_l	n_H	n_G	stop. cond.
$32^3 \times 96$	2.38	1	2.4	0.707	10	60	360	1e-8

MG Parameter Tuning

MG parameters	Run 1	Run 2	Run 3	Run 4	Run 5 [nvecs=16]
setup_res.	0.4	0.1	0.4	0.1	0.5
cres	0.3	0.3	0.5	0.3	0.3
setup_change_fac	0.4	0.1	0.2	0.4	0.4
npre	5	5	4	0	5
npost	9	9	9	5	9
scale	1	0	0.2	1	1
Setup Time [secs]	49	61	43	65	29
Traj. 1 Time [secs]	160	575	308	152	162
Traj. 2 Time [secs]	182	672	376	214	176
Traj. 3 Time [secs]	201	686	409	282	192
Traj. 4 Time [secs]	222	681	427	344	208

- Many parameters to tune.
- Fixed nvecs = 24 in Run 1-4, and 16 in Run 5.
- Scanned other parameters to find the best set.
- Manual and painful. Needs a better (preferably automatic) tuning strategy.

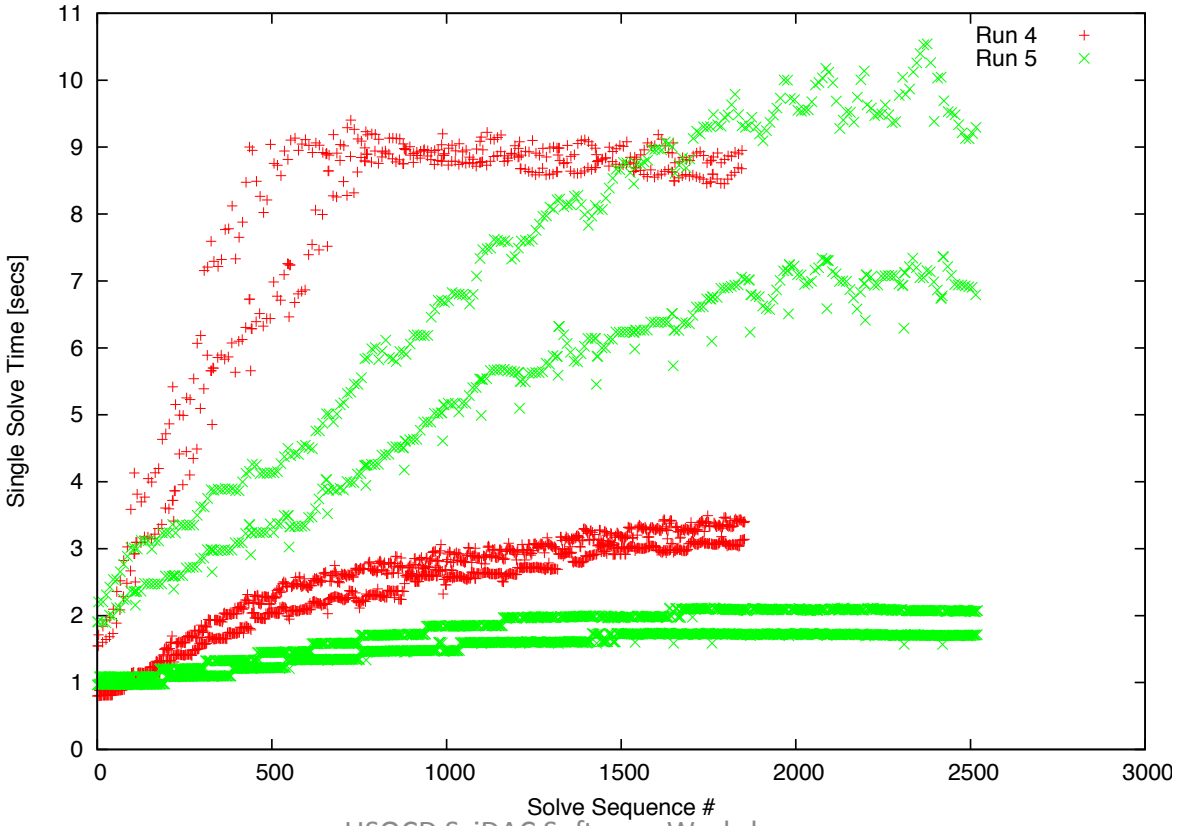
Tuning for MG-HMC

MG parameters	Run 1	Run 2	Run 3	Run 4	Run 5 [nvecs=16]
setup_res.	0.4	0.1	0.4	0.1	0.5
cres	0.3	0.3	0.5	0.3	0.3
setup_change_fac	0.4	0.1	0.2	0.4	0.4
npre	5	5	4	0	5
npost	9	9	9	5	9
scale	1	0	0.2	1	1
Setup Time [secs]	49	61	43	65	29
Traj. 1 Time [secs]	160	575	308	152	162
Traj. 2 Time [secs]	182	672	376	214	176
Traj. 3 Time [secs]	201	686	409	282	192
Traj. 4 Time [secs]	222	681	427	344	208

- Run 4 has the best time for first trajectory, but deteriorates quickly
- Run 5 has the best overall performance.
- Same setup can be used for 3 trajectories.

An Optimization Problem

- If the setup is tuned too well for the first solve, subsequent solves get worse quickly. ➔ physical reasons?
- If it is not tuned well, overall gain is small.
- It is tricky to find the sweet spot.



MG-HMC Performance

24^3x64, time averaged over 20 trajectories

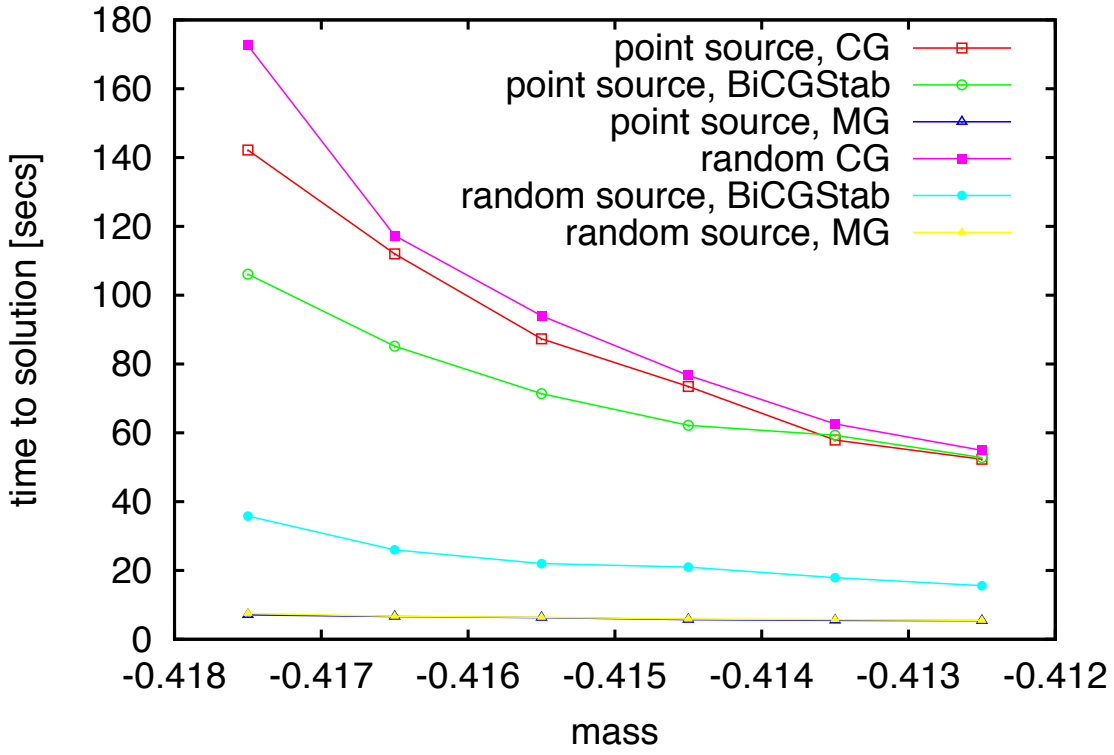
Solver	Light Solve [secs]	Heavy Solve [secs]	Trajectory Time [secs]
CG	176	121	326
BiCGStab	111	99	239
MG	61	91	187

32^3x96, time averaged over 3-8 trajectories

Solver	Light Solve [secs]	Heavy Solve [secs]	Trajectory Time [secs]
CG	830	628	1596
BiCGStab	445	502	1086
MG	209	451	822

- Light solve: MG is 2x faster than BiCGStab, 3-4x faster than CG
- Speedup per trajectory not as big
- Bottleneck is heavy solves ➔ Can rebalance HMC

Dependence on Source Vector



- Source vectors have little effect on CG or MG.
- BiCGStab converges much faster for a random source vector.

Reversibility

- Will reusing the setup affect reversibility?
- No sign so far, but more tests are needed.

```
==With MG==
Sold: 22721701.88   Srev: 22721701.88   dS: 1.329928637e-06
Sold: 22725067.11   Srev: 22725067.11   dS: -0.001061491668
Sold: 22713290.68   Srev: 22713290.68   dS: 0.0005583688617
Sold: 22721697.35   Srev: 22721697.35   dS: -0.0001310259104
Sold: 22724432.14   Srev: 22724432.14   dS: -0.0001665465534

==Without MG (BiCGStab)==
Sold: 22721701.88   Srev: 22721701.88   dS: 0.0003642588854
Sold: 22725067.1    Srev: 22725067.1    dS: -0.0002857670188
Sold: 22713290.69   Srev: 22713290.69   dS: -0.0004257671535
Sold: 22721697.35   Srev: 22721697.35   dS: -0.0006039328873
Sold: 22724432.15   Srev: 22724432.15   dS: -0.0003919377923
```

TODO

- Clover MG-HMC
- Tests on lighter masses and larger volumes.
- Retuning of HMC to see if further speedup is possible.
- Reuse previous near-null vectors to reduce subsequent setup cost.
- Automatic tuning of MG parameters?

Conclusions

- MG-HMC for Wilson has been implemented in FUEL.
- Performance at a pion mass of 420 MeV is already promising.
- Gain should be better with lighter masses and larger volumes.
- More work needs to be done on optimization strategies.