# Multigrid solver for clover fermions

James C. Osborn (ANL)

with

R. Babich, R. C. Brower, S. D. Cohen, C. Rebbi (BU)

J. Brannick (Penn. State)

M. A. Clark (Harvard)

T. Manteuffel, S. McCormick (UC Boulder)

June 17, Lattice 2010, Sardinia, Italy

# Outline

- Multigrid intro
- Adaptive multigrid
- Application to Wilson Clover
- Results
- Setup cost
- Plans

# The problem

- Lattice QCD requires repeated solution of Dirac equation

$$[D(U) + m]\psi = \eta$$

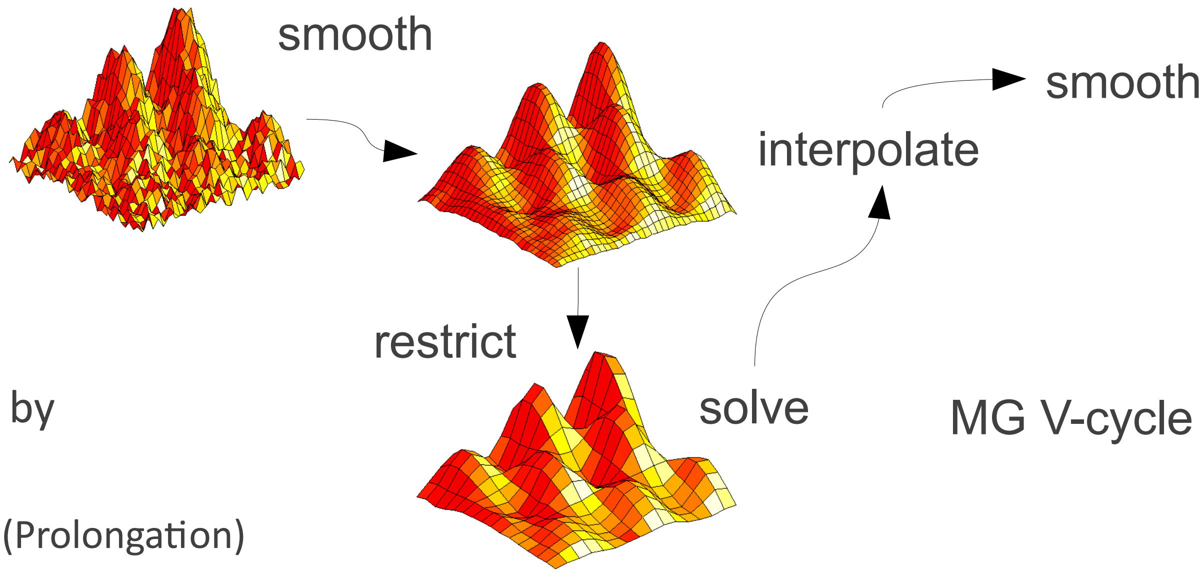
- Much of the work goes into solution
  - Typically over 50% for lattice generation
  - Can be over 90% for analysis

- Exhibits critical slowing down
  - Condition number diverges as mass decreases ($\kappa \sim 1/m$)
  - Standard Krylov solvers (CG, BiCGstab, …) become inefficient as condition number grows (iterations $\sim \sqrt{\kappa}$)

- Multigrid methods have been very successful in beating this in other fields

# Multigrid

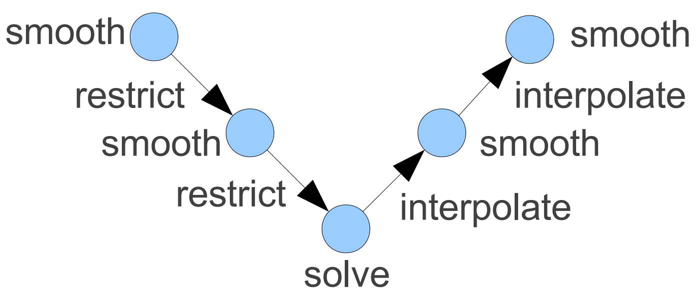
- Standard solvers (stationary, Krylov) good at reducing high frequency error components, not good with low frequency errors
- MG projects error onto coarse grid, solves, then interpolates correction back to fine grid

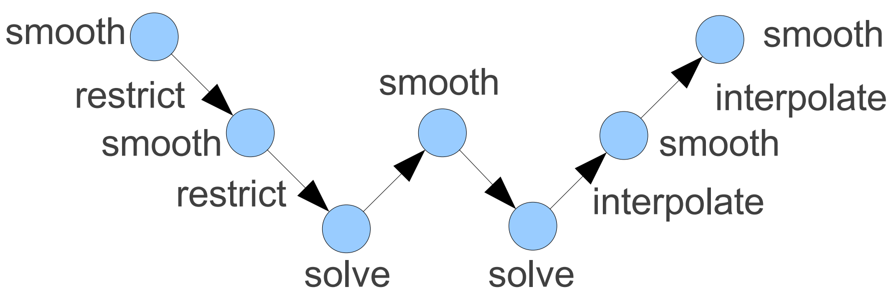



- V-cycle determined by
  - Restriction (R)
  - Interpolation (P) (Prolongation)
  - smoother

# Multigrid

- MG V-cycle typically used as preconditioner for outer solver
  - Here using GCR
- Used recursively: MG cycle used to solve on coarse grid, …
- Choice of cycle:
  - V-cycle, W-cycle, …
  - Here using GCR solver for coarse system with MG preconditioner



# Choosing R & P

- Coarse grid solve:

$$PA_c^{-1}Rr$$

$$A_c = RAP$$

- Algebraic MG: P & R formed from elements of A (or approximation to)
- Adaptive MG: P & R formed from slow-to-converge modes of A
  - Want P to preserve low modes of A
  - Form P from representative low modes chopped into blocks
  - R from low modes of A$^\dagger$

$$P = \begin{pmatrix} v1 & v2 & & & \\ v1 & v2 & & & \\ v1 & v2 & & & \\ & & v1 & v2 & \\ & & v1 & v2 & \\ & & v1 & v2 & \\ & & & & \ddots \end{pmatrix}$$

# Fine and coarse operators

- MG normally done on Hermitian positive definite systems ($D^\dagger D$)
  - Coarse operator constructed from Galerkin prescription $R = P^\dagger$, $A_c = P^\dagger A P$
  - Increases complexity of coarse operator (has 2-hop corner terms)
- Instead using just D
  - Want R to be rich in low left-modes
  - For $\gamma_5$-Hermitian operator can set $R = P^\dagger \gamma_5$
- Solving Wilson-clover operator
  - Using even-odd preconditioning on fine system
  - $D\,x = b \rightarrow D_p\,x_p = b \rightarrow D_r\,x_r = b_r$
  - Construct coarse operator from $D_p$ then construct reduced operator
  - $D_p$ no longer $\gamma_5$-Hermitian, but use same R anyway

$$D = \begin{pmatrix} D_{ee} & D_{eo} \\ D_{oe} & D_{oo} \end{pmatrix}$$

$$D_p = \begin{pmatrix} 1 & D_{eo} D_{oo}^{-1} \\ D_{oe} D_{ee}^{-1} & 1 \end{pmatrix}$$

$$D_r = 1 - D_{eo} D_{oo}^{-1} D_{oe} D_{ee}^{-1}$$

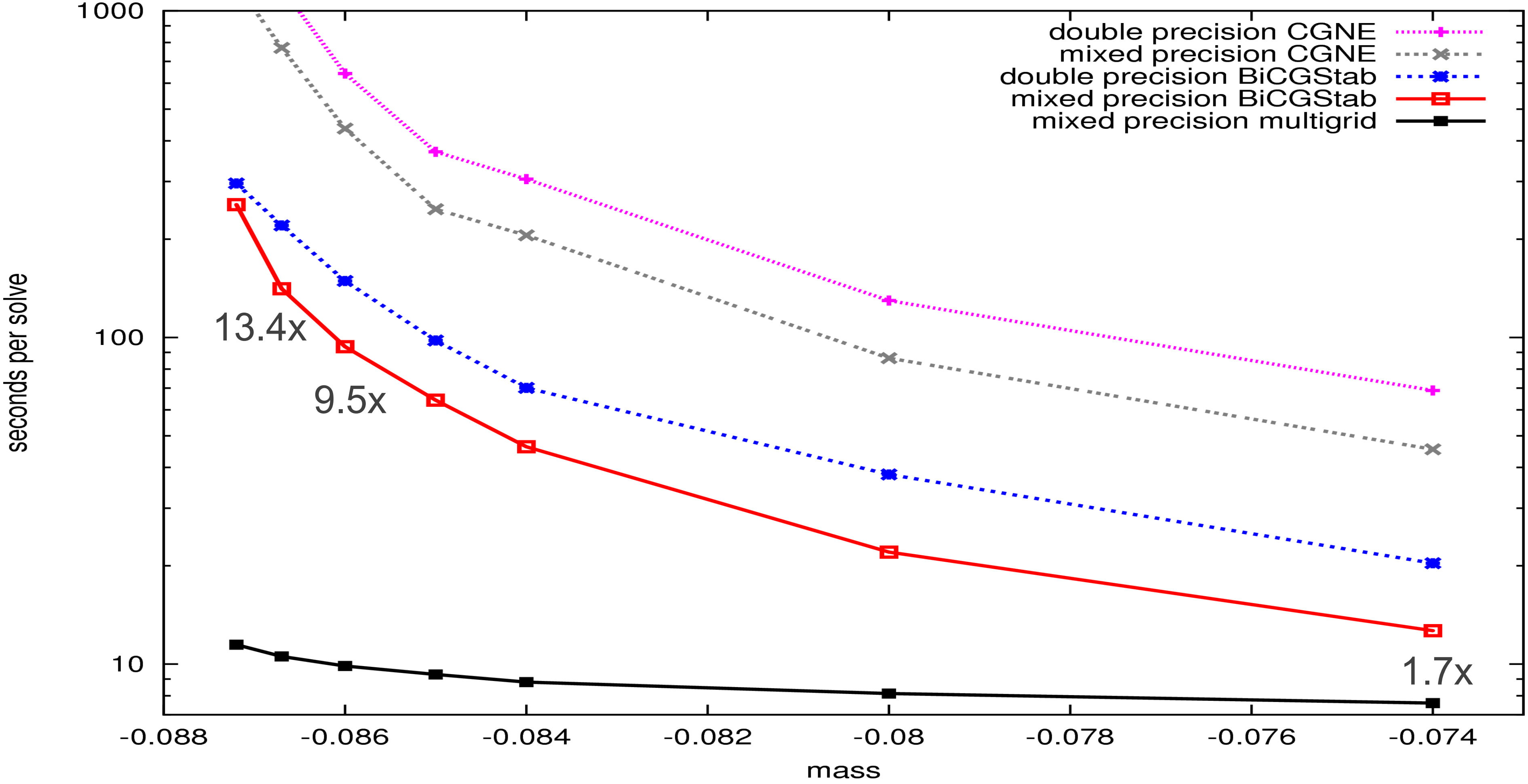
# Implementation Details

- Mixed precision
  - Outer GCR solver on fine level in double precision
  - MG preconditioner and all levels below in single
  - Comparison to mixed precision Krylov methods (iterative refinement)
- Implemented in US SciDAC QDP/C
  - Multi-lattice support and improved arbitrary Nc support

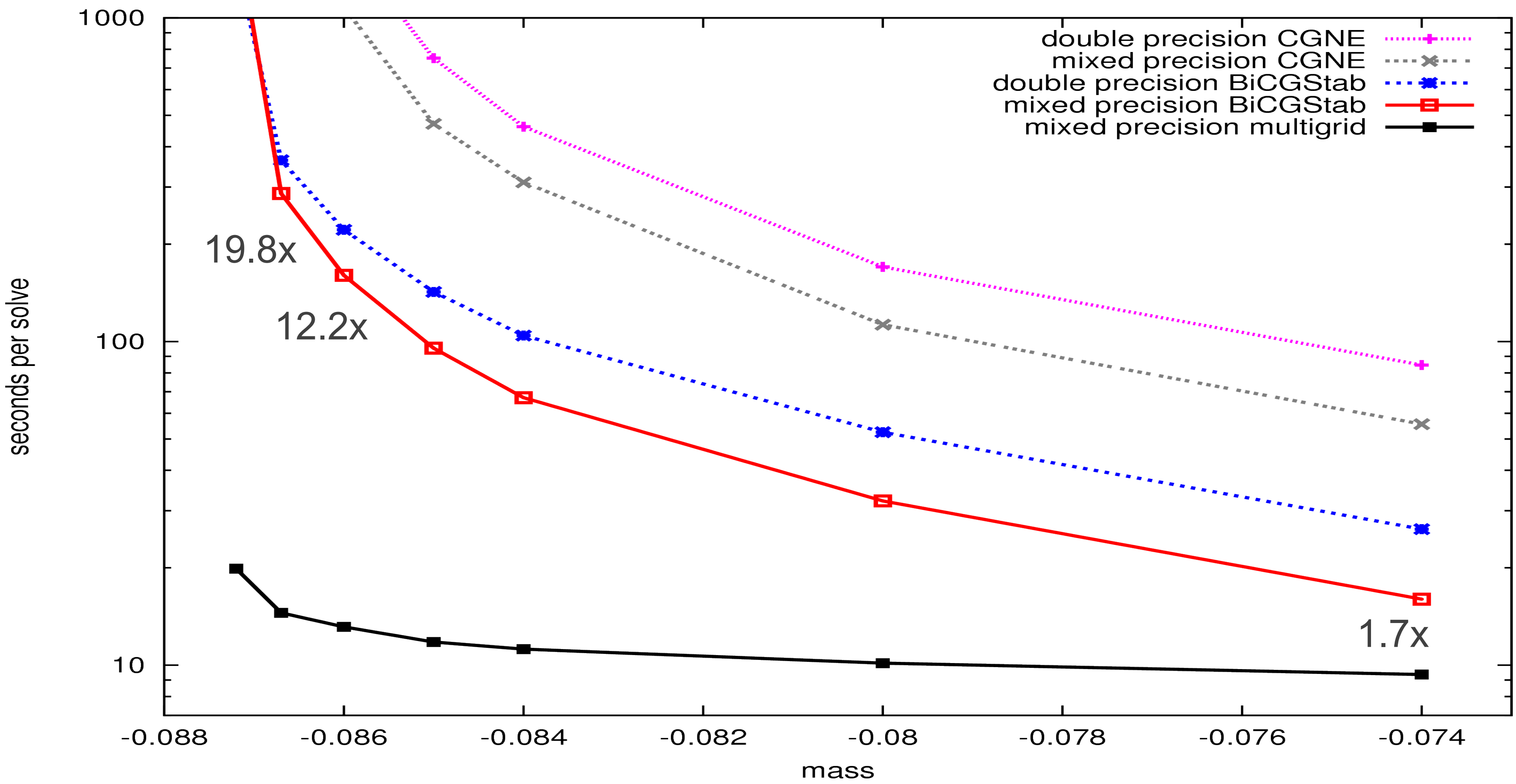
# Numerical results

- Using LHPC Hadron Spectrum Collaboration anisotropic clover lattices
  - $a_s \approx 3.5\ a_t$, $a_t \approx 0.1$ fm
  - $24^3$x128 and $32^3$x256
  - Dynamical $m_\pi \approx 220$ MeV (m = -0.086)
- Results obtained on BG/P
  - 256 cores for $24^3$x128
    - 1st coarse lattice: $8^3$x16 with 24 vectors
    - 2nd coarse lattice: $4^3$x4 with 32 vectors
  - 1024 cores for $32^3$x256
    - 1st coarse lattice: 16x8x8x32 with 24 vectors
    - 2nd coarse lattice: 4x4x4x16 with 32 vectors

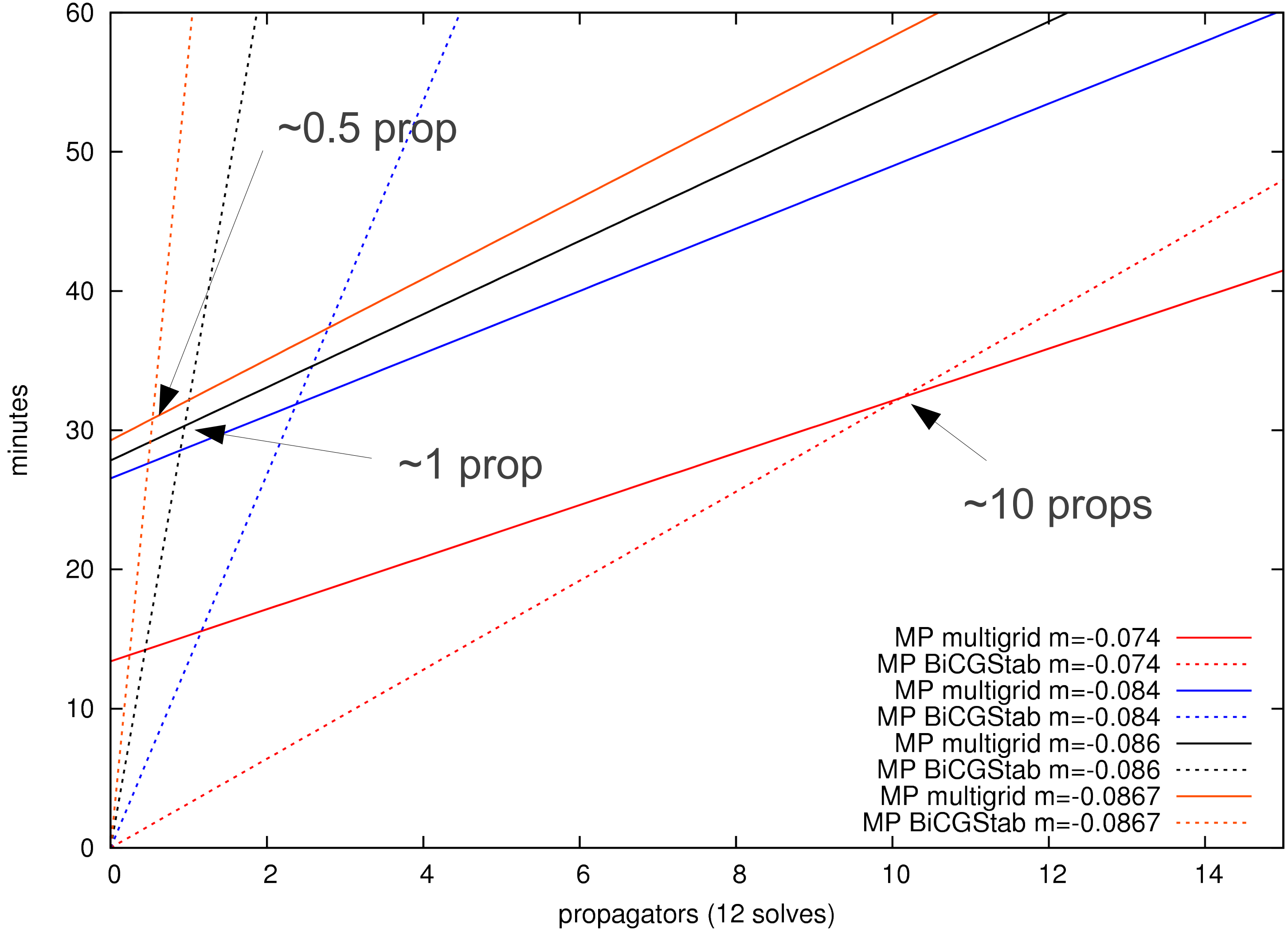
# Results $24^3$x128



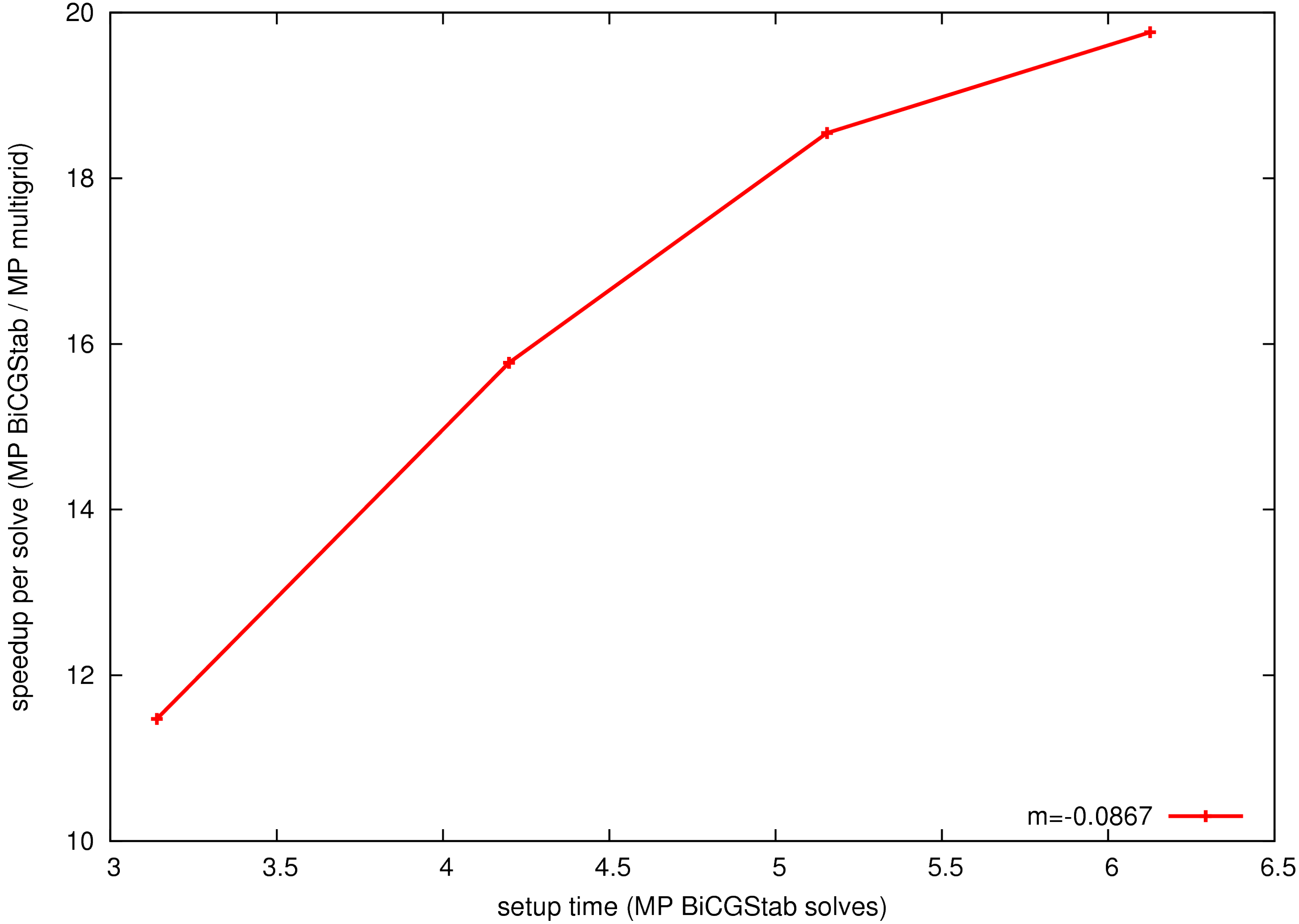
# Results $32^3$x256



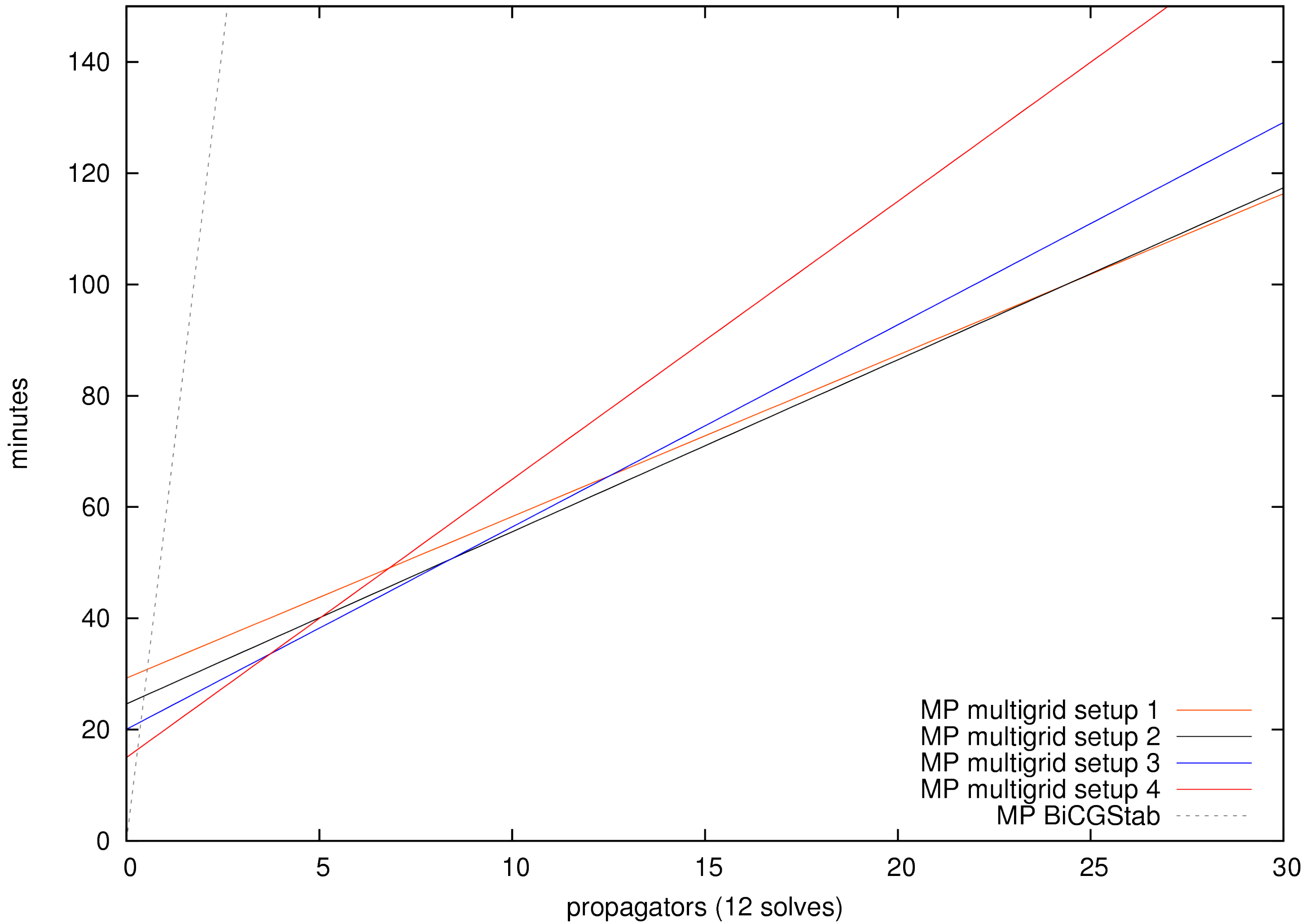
# Setup cost $32^3$x256



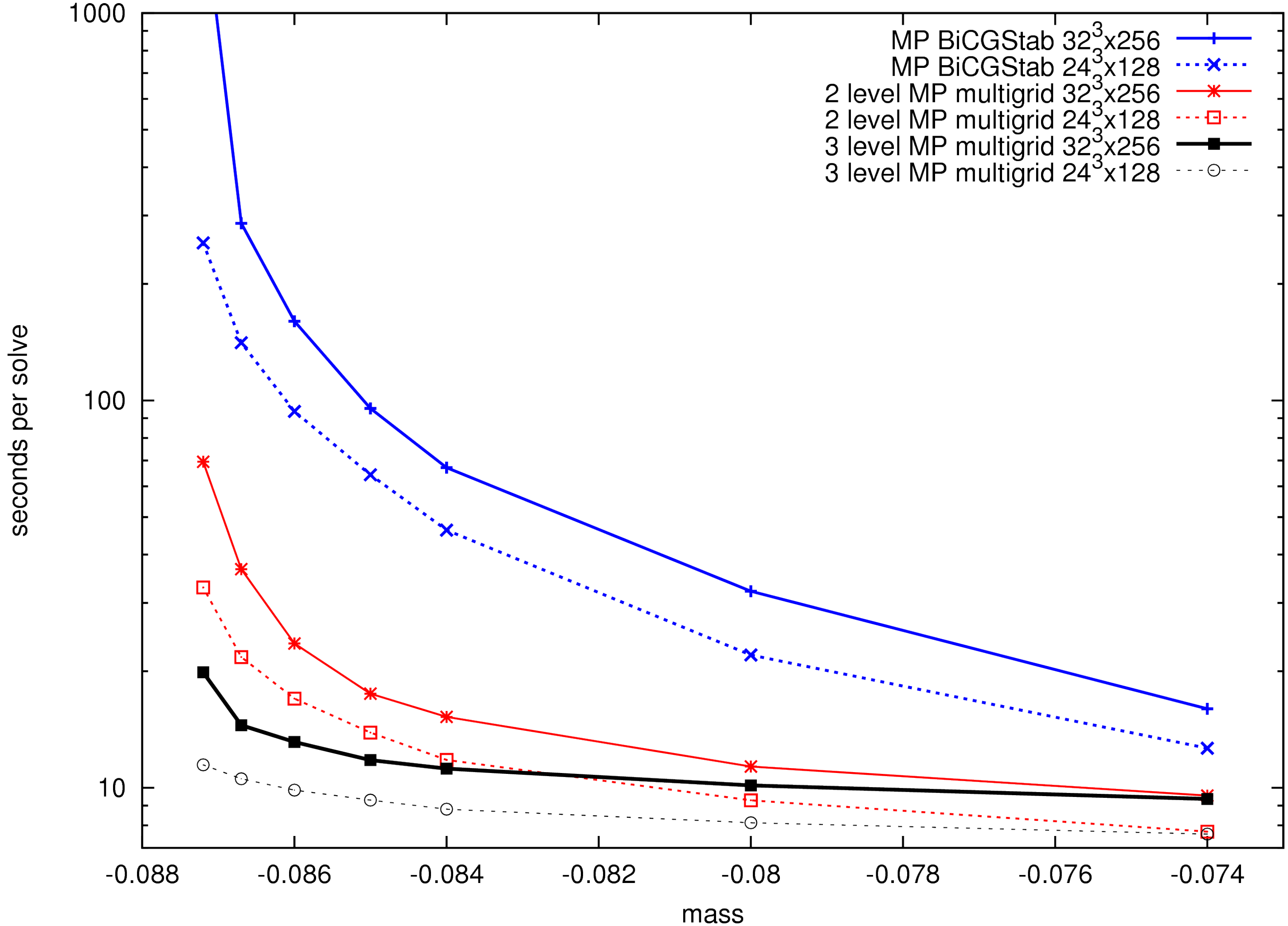
# Setup cost vs speedup $32^3$x256 (physical mass)



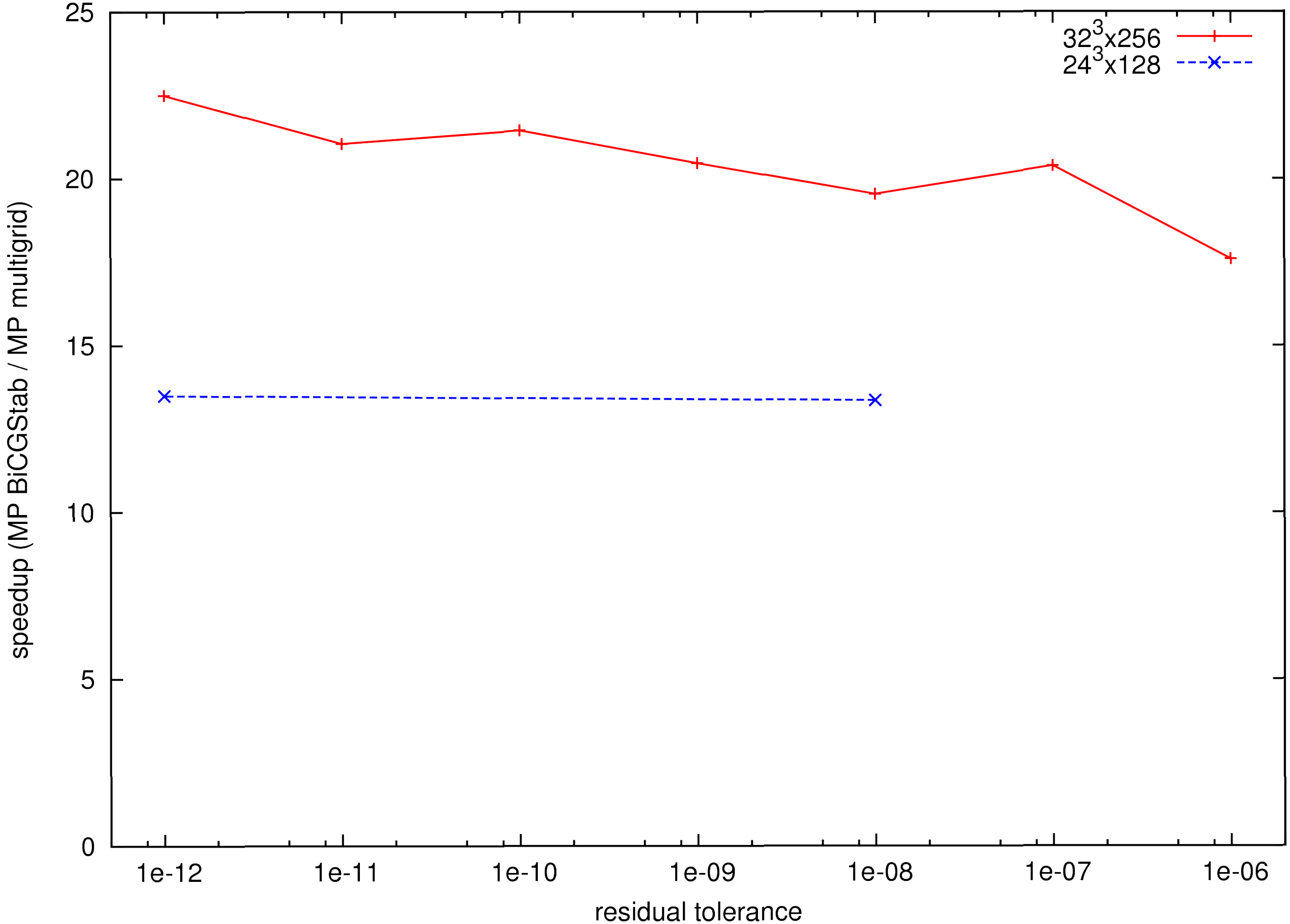
# $32^3$x256 setup cost



# 2 level vs 3 level



# Speedup vs residual



# Error vs residual

- Error:

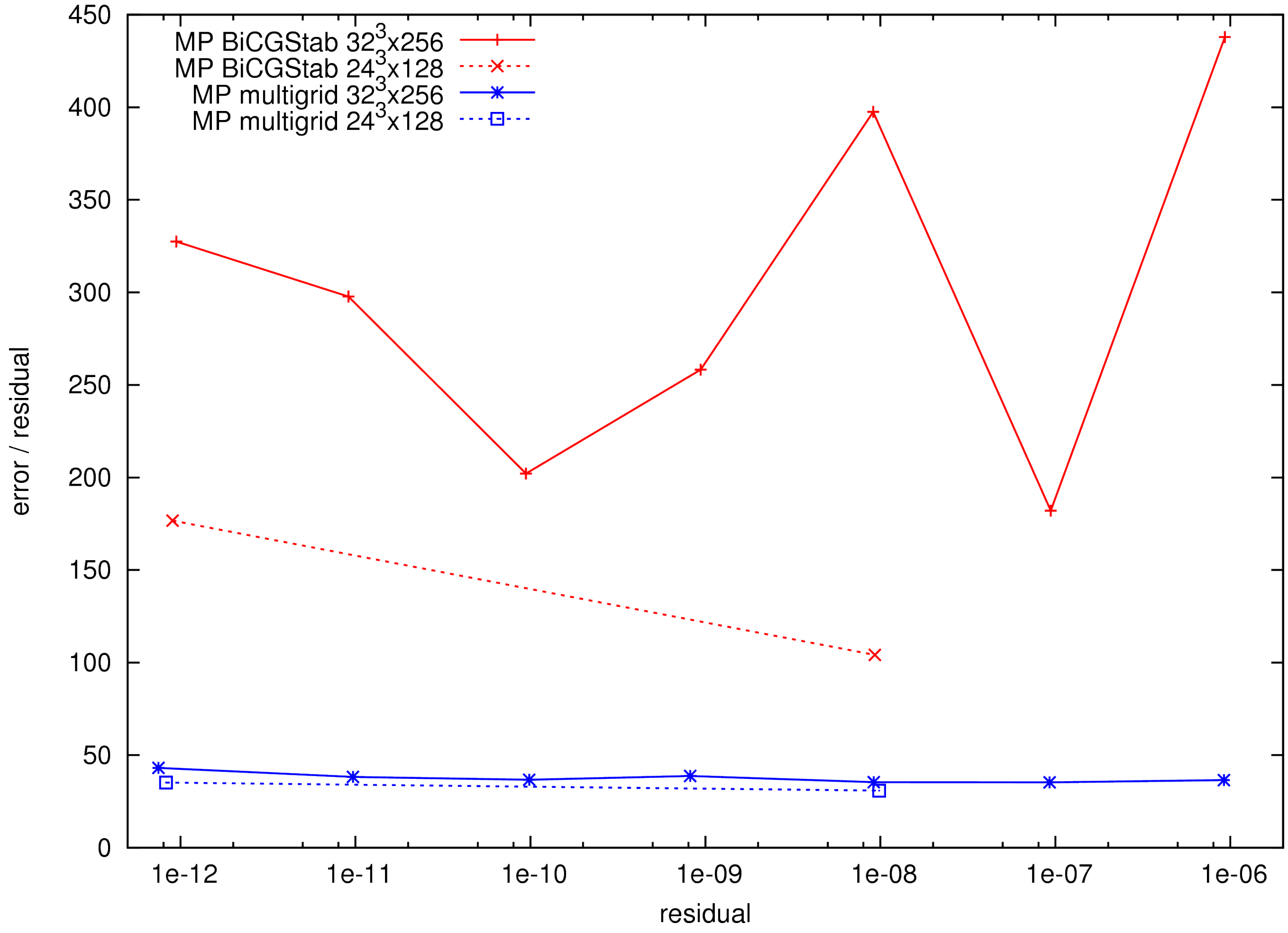
  $e = x^* - x$

- Residual:

  $r = b - A x$

  $= A e$

- Residual not as sensitive to low modes



# Setup methods

- Repeated Relaxation (inverse iteration) on random vectors
  - Simple (don't need to construct coarse operator)
  - Can vary number of iterations/cycles
  - Vectors may be locally redundant
- Adaptive smooth aggregation (αSA)
  - Construct new MG cycle with current vectors, use to find new vector
  - Requires construction of coarse operator
  - New vectors should be new important

# Summary and outlook

- Efficient implementation of Clover MG solver currently being used in production (disconnected diagrams)
- Gives 10-20x improvement for light quarks
- Error very stable and relatively small
- Speedup (and relative error) improves for larger lattices

- Test on larger lattices & extend to more levels
- Improve setup
  - Currently great for medium to large analysis projects
  - Extend to smaller projects and HMC
- Staggered & chiral quarks
- Port to GPUs