Multigrid Methods for Lattice QCD

James Brannick
Center for Computational Mathematics and Applications
The Pennsylvania State University
http://math.psu.edu/∼brannick/
brannick@psu.edu

Fourth International Workshop
on Numerical Analysis and Lattice QCD

Participants in the MG-QCD project

Ball State	I. Livshits	Columbia	A. Bessen
Boston Univ.	R. Brower		D. Keyes
	M. Clark	LLNL	R. Falgout
	J. Osborne	Penn State	J. Brannick
	C. Rebbi		V. Nistor
Boulder	M. Brezina		J. Xu
	C. Ketelesen		L. Zikatanov
	T. Manteuffel	T. U. Delft	S. MacLachlan
	S. McCormick		Kees Oosterlee
	J. Ruge	Weizmann Inst.	A. Brandt

Presentation Plan

* Model problems - solver challenges
  * Dirac Wilson system of QCD
  * Schwinger model of QED
* Smoothed Aggregation (A)MG
* Adaptive (A)MG Approaches
* Numerical Results
* Future Work

4D Dirac Wilson system

$$(M\phi)(x) = \phi(x) - \kappa\Bigg(\sum_{\mu=1}^{4}(I-\gamma^{\mu})\otimes U_{\mu}(x)\phi_{x+e_{\mu}} + (I+\gamma^{\mu})\otimes U_{\mu}^{*}(x-e_{\mu})\phi_{x-e_{\mu}}\Bigg)$$

* For fixed x, $U_{\mu}(x) \in SU(3)$: $\beta \to \infty \Rightarrow U_{\mu} \to I$
* $\gamma^{\mu} \in \mathbb{C}^{4\times 4}$: sparse matrices with 4 nonzero entries, $\pm 1, \pm i$
* $M \in \mathbb{C}^{n\times n}$, $n = 12 \cdot n_1 \cdot n_2 \cdot n_3 \cdot n_4$
* M is positive for $0 \leq \kappa < \kappa_c$ and M becomes very ill-conditioned for $\kappa \approx \kappa_c$
* $\gamma^5 M = M^* \gamma^5$

Simplification: 2D Dirac Wilson system

$$M = \frac{1}{2}\left(\sum_{\mu=1}^{2} \sigma_\mu (\nabla_\mu^+ + \nabla_\mu^-) + \nabla_\mu^- - \nabla_\mu^+ \right) + m$$

* Cov. finite difference operators ($U_\mu \in U(1)$):

$$(\nabla_\mu^+ f)(x,s) = U_\mu(x) f(x + e_\mu, s) - f(x,s)$$

$$(\nabla_\mu^- f)(x,s) = f(x,s) - U_\mu^*(x - e_\mu) f(x - e_\mu, s)$$

* $H = \sigma_3 M = M^* \sigma_3 = H^*$, with $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

2D Hermitian Dirac Wilson System $Hf = \psi$

$$\begin{pmatrix} A & B \\ B^* & -A \end{pmatrix} \begin{pmatrix} f(x,1) \\ f(x,2) \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

Taking $a = \frac{1}{N}$ on a fixed domain and setting $U_\mu \equiv 1$ gives

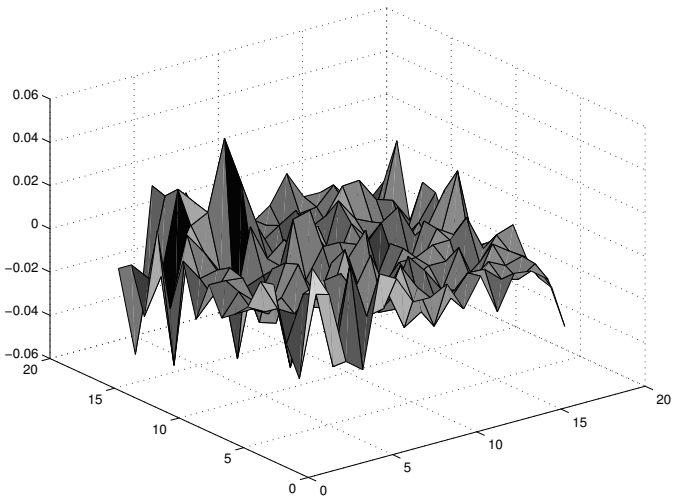
$$\begin{pmatrix} A & B \\ B^* & -A \end{pmatrix}^2 = \begin{pmatrix} (mI - \frac{a}{2}\Delta_a)^2 - \Delta_a & 0 \\ 0 & (mI - \frac{a}{2}\Delta_a)^2 - \Delta_a \end{pmatrix}$$

For $m = 0$, $\mathbf{1} \in \mathrm{Null}(H)$ and $(mI - \frac{a}{2}\Delta_a)^2 - \Delta_a = -\Delta_a(I - \frac{a^2}{4}\Delta_a)$, with

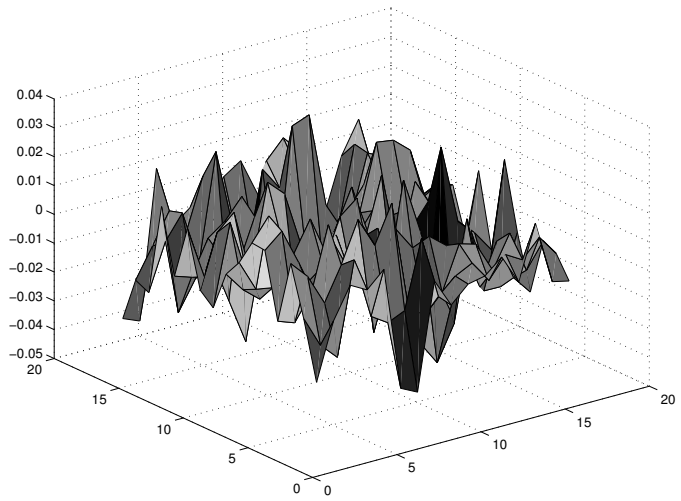
$$\sigma(-\Delta_a) \subseteq [0, \frac{8}{a^2}] \quad \text{and} \quad \sigma(I - \frac{a^2}{4}\Delta_a) \subseteq [1, 3]$$

General case: Lowest eigenmode of H^2

Real



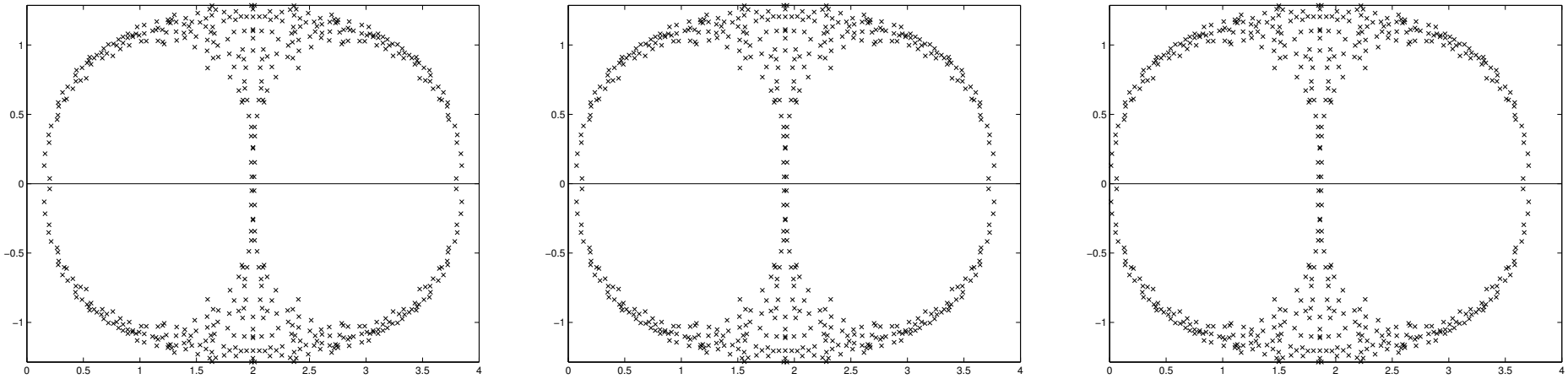
Imag



$m \approx m_{cr}, \beta = 3$, $n = 2 \cdot 16 \cdot 16$ lattice

Spectrum of M

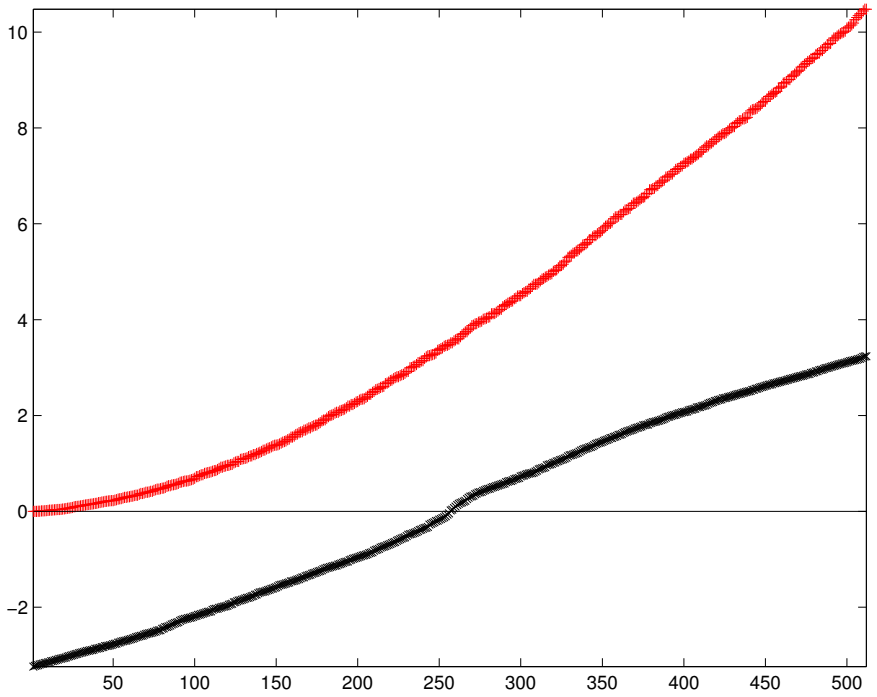
$m = 0$ $m = -.1$ $m = -.2$



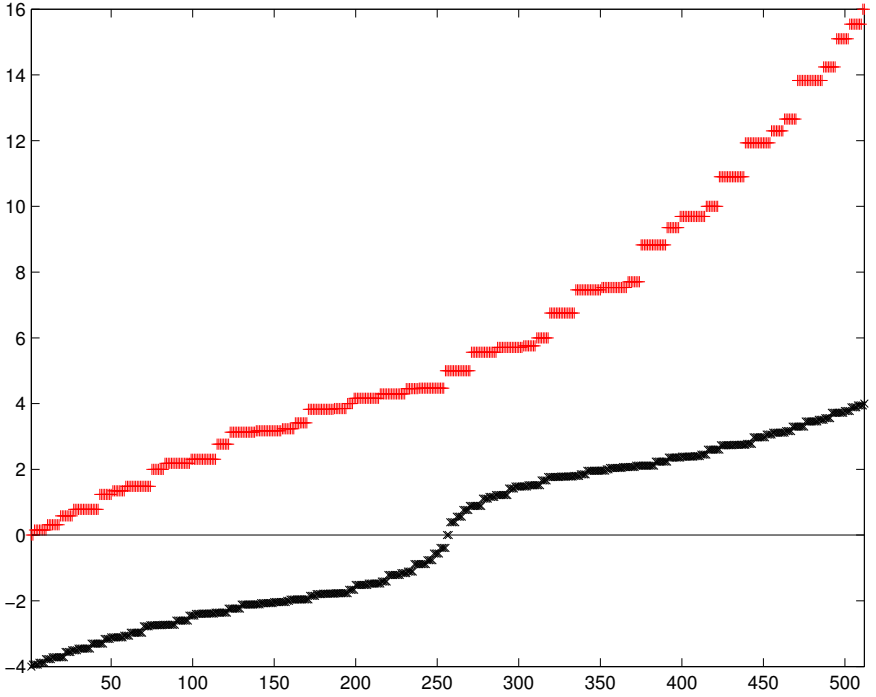
$\beta = 3, m_{cr} \approx -.21, n = 2 \cdot 16 \cdot 16$

Spectrum of H and H^2

Free Field



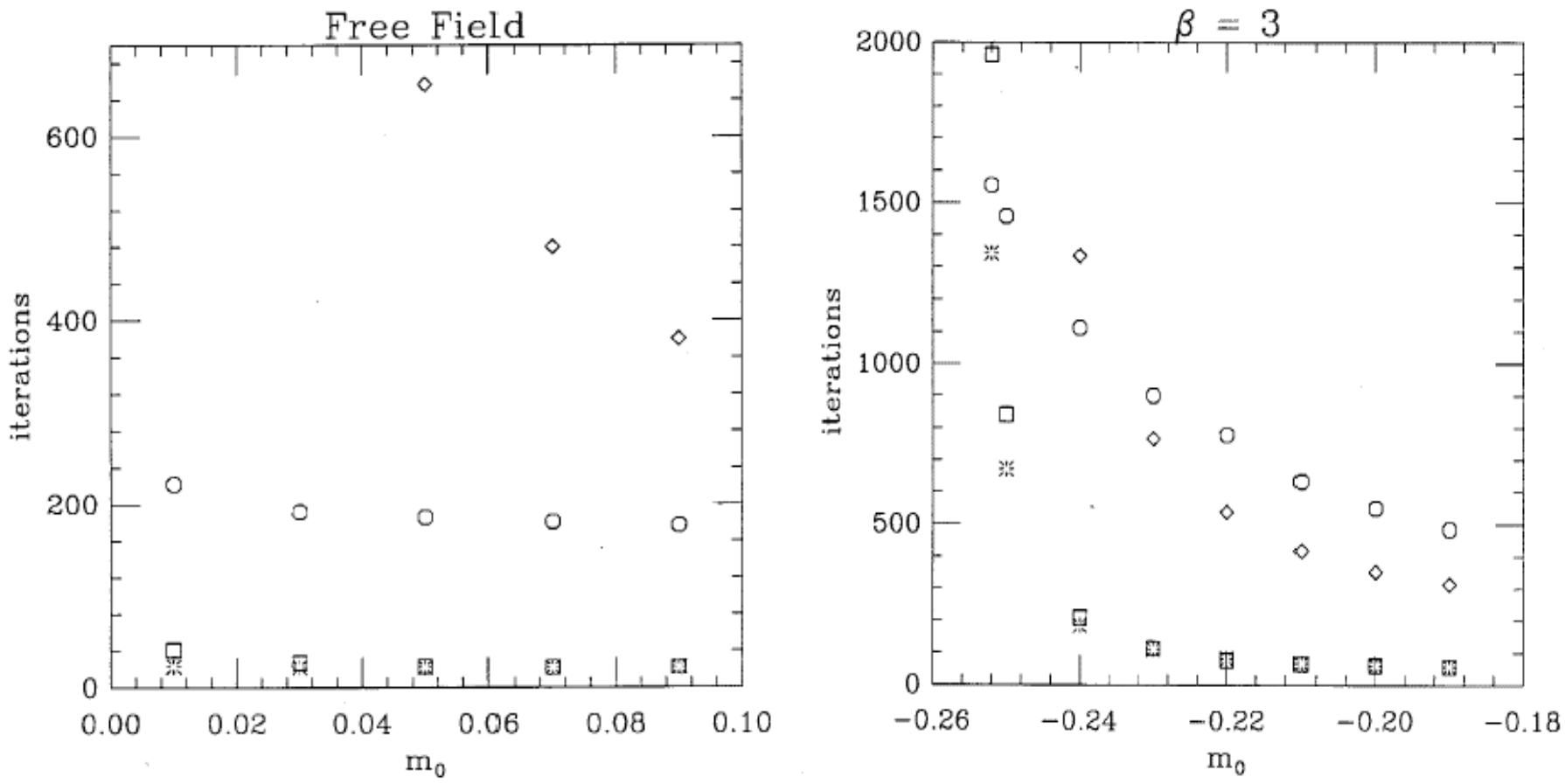
$\beta = 3$



$m \approx m_{cr}$, $n = 2 \cdot 16 \cdot 16$

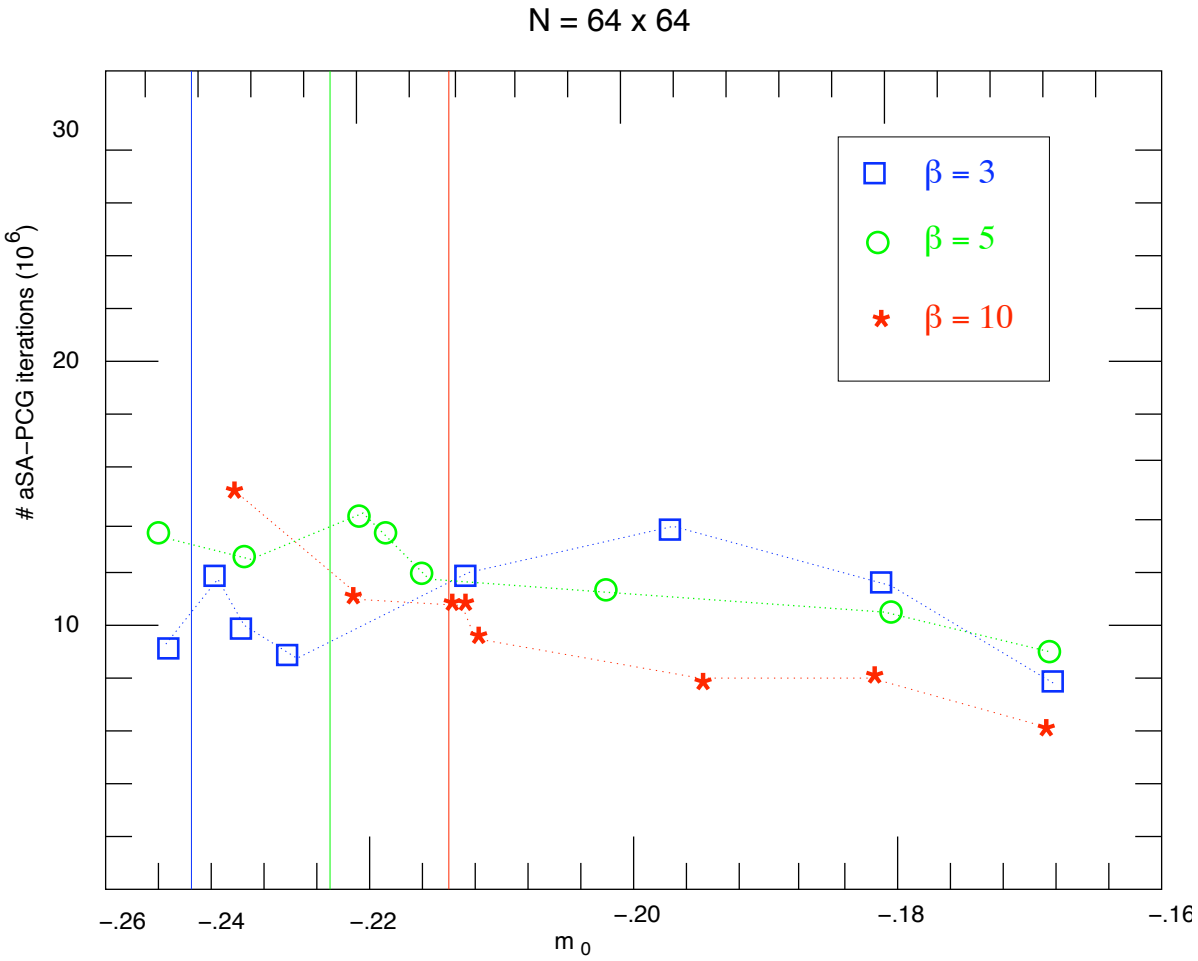
MG for lattice fermions circa 2000

* R. Ben-Av, et al, *Fermion simulations using parallel transported multigrid*, Phys. Lett. **B**253 (1991), pp. 185–192
* R. Ben-Av, M. Harmatz, S. Solomon, and P. G. Lauwers, *Parallel* transported multigrid for inverting the dirac operator: Variants of the method and their efficiency, Nucl. Phys. **B**405 (1993), pp. 623–666
* A. Brandt, *Multigrid methods in lattice field computations*, Nucl Phys. Proc. Suppl. **2**6 (1992), pp. 137–180
* R. C. Brower, R. G. Edwards, C. Rebbi, and E. Vicari, *Projective multigrid for Wilson fermions*, Nucl. Phys. **B**366 (1991), pp. 689–705
* P. Hasenfratz, *Prospects for perfect actions*, Nucl. Phys. Proc. Suppl. **6**3 (1998), pp. 53–58
* Many others ...



$N = 64$, Jacobi (Diamond), CG (circle), MG V-cycle (square), W-cycle (star)

Brower, et. al., *Projective multigrid for Wilson fermions* (1991)

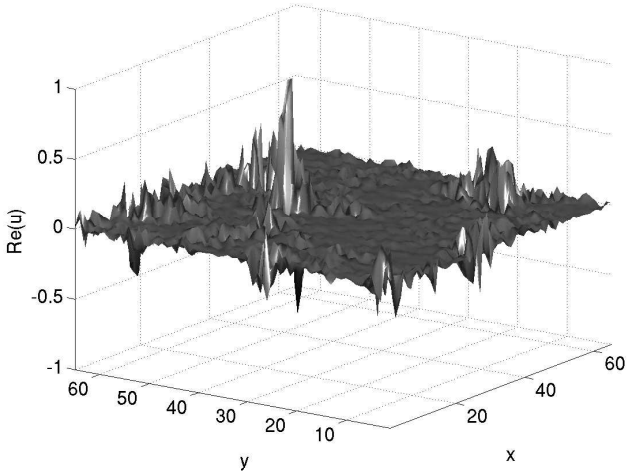


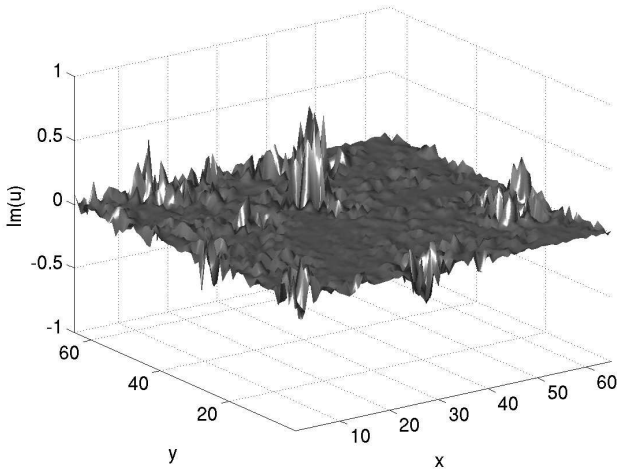
Brannick, et. al., *Adaptive smoothed aggregation in lattice QCD*, Lecture Notes Comp. Sci. Eng., Springer-Verlag, 2006

Extreme simplification: Gauge Laplacian

$$L = \sum_{\mu} \nabla_{\mu}^{-} - \nabla_{\mu}^{+} + m$$

* Scaling by $\frac{1}{2(2+m)}$ gives $L = I - \kappa D$, $D = \begin{pmatrix} 0 & D_{eo} \\ D_{oe} & 0 \end{pmatrix}$ with $L > 0$ for $0 \leq \kappa < \kappa_{cr} := \frac{1}{\lambda_{\max}(D)}$





Presentation Plan

* Model problems - solver challenges
  * Dirac Wilson system of QCD
  * Schwinger model of QED
* Smoothed Aggregation (A)MG
* Adaptive (A)MG Approaches
* Numerical Results
* Future Work

Method of subspace corrections

* Abstract problem: Find $u \in V$ such that $Au = f$, $f \in V^*$, $A > 0$
* Space decomposition: $V = \sum_i V_i$

  Idea: Correct the error in each subspace by solving $A_i e_i = r_i$ for $i = 1 : J$

  * exactly, if each of the subspaces is small (e.g., multiplicative Schwarz)
  * approximately, if the subspaces are of large dimension (e.g., multigrid)

  Algorithm MSC With $u \leftarrow u^0$, for $i = 1 : J$ repeat the subspace correction $u \leftarrow u + e_i$ where $e_i \in V_i$ is given by

$$e_i = M_i^{-1} \Pi_i (f - Au)$$

* **Examples**: (1) $V = \mathbb{R}^n$ with $V_i := \text{span}\{\delta_i\}$ and $M_i = A_i = a_{ii}$ gives Gauss-Seidel; (2) $V_J \supset V_{J-1} \supset \ldots \supset V_1$ and $M_i \approx A_i$ results is MG

Let $T_i = M_i^{-1} A_i \pi_i$, then $u - u^\ell = E_J(u - u^{\ell-1})$,

$$E_J = (I - T_J)(I - T_{J-1}) \dots (I - T_1)$$

Theorem.[XU AND ZIKATANOV, J. of AMS (2003)] The convergence factor is obtained via $\|E_J\|_A^2 = 1 - \frac{1}{c_0+1}$,

$$c_0 = \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} (\bar{T}_i^{-1} T_i^* w_i, T_i^* w_i)_A \text{ with } w_i = \sum_{j=i}^{J} v_j - T_i^{-1} v_i$$

where $\bar{T}_i \equiv T_i^* + T_i - T_i^* T_i$ SPD $\Rightarrow c_0 < \infty$

Application to MG

* For finite element spaces and regular refinement $V_J \supset V_{J-1} \supset ... \supset V_1$. In such settings, under appropriate assumptions, it is straightforward to show that the convergence depends logarithmically on the number of levels
* If nested finite element meshes are not available, then the standard arguments do not apply!
* Motivated by success of MG methods algebraic versions were proposed to handle general problems. Originally for second-order elliptic equations with discontinuous coefficients on unstructured grids
* "Algebraic" stands for the fact that all the tools of the method are constructed solely on the basis of the original matrix A, in a setup phase
* In AMG, the subspaces are built "on the fly", making multilevel theory for the convergence of such algorithms very difficult

AMG estimates ($J = 2$)

Two-grid error propagator:

$$E_{tg} = (I - M^{-t}A)(I - \pi_A)(I - M^{-1}A)$$

with $\pi_A := P(P^tAP)^{-1}P^tA$. Two-level convergence factor is obtained via

$$\|E_{tg}\|_A^2 = 1 - \frac{1}{K}; \quad K(P) = \sup_v \frac{\|(I - \pi_{\bar{M}})v\|_{\bar{M}}}{\|v\|_A}$$

where $\bar{M} := M^t(M^t + M - A)^{-1}M$. Analysis of recursive multi–level method depends in addition on stability:

$$\|PR\|_A < \eta; \quad RP = I_c$$

What does it mean to choose $V_c = \mathrm{Range}(P) \subset V$?

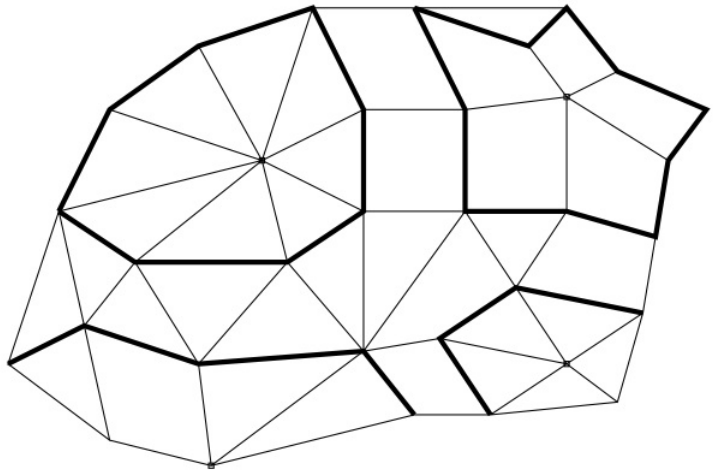
Coarse space is constructed automatically within the algorithm, level by level, in a (hopefully) computationally optimal setup procedure which involves

1. Picking a set of coarse variables, i.e., set of indices
   $\mathcal{N}_c = \{i_1, \ldots, i_{n_c}\} \rightarrow$ graph theoretic approaches

2. Definiing $V_c = \mathrm{span}\{\psi_k\}_{k=1}^{n_c}$ such that each ψ_k is supported in Ω_k, for a vector:
   $\Omega_k \subset \{1, \ldots, n\} \rightarrow$ null space of the system matrix

$*$ Each of the V_c (or V_i obtained recursively) must satisfy certain properties, related to the convergence of the overall algorithm.

Smoothed Aggregation



* Define *strong* nbhd. of dof $i \in \Omega = \{1, .., n\}$ as

$$N_i(\epsilon) := \{j \neq i : \epsilon a_{ij}^2 > a_{ii} a_{jj}\}$$

* Using N_i and ind. set algorithm, Ω partitioned into set of disjoint aggregates $\{\mathcal{A}_i\}_{i=1}^{n_c}$ such that $\cup_{i=1}^{n_c} \mathcal{A}_i = \Omega$

$\text{agg}(\Omega, A)$

do $i \in \bar{\Omega}$

if $N_i(\epsilon) \geq 1$

$\mathcal{A}_j = N_i(\epsilon) \cup \{i\}$

$\bar{\Omega} \leftarrow \bar{\Omega} \setminus \mathcal{A}_j$

$j \leftarrow j + 1$

end if

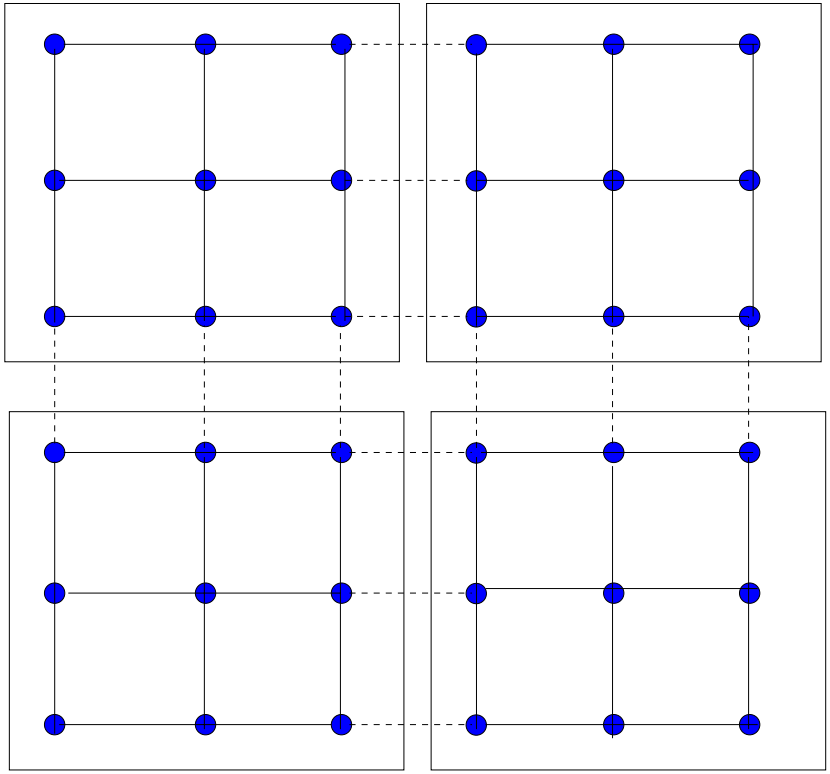
end do

Smoothed Aggregation MG on a lattice

* On structured lattices we define $\{\mathcal{A}_i\}_{i=1}^{n_c}$ using simple geometric blocking, again such that

$$\cup_{i=1}^{n_c} \mathcal{A}_i = \Omega$$

* This *fixes* the sparsity pattern of P and hence also of $A_c = P^t A P$



Smoothed Aggregation (A)MG

Given $B = [x],\ x \in \mathbb{R}^n$, and the set of aggregates, $\{\mathcal{A}_i\}_{i=1}^{n_c}$

$$\tilde{P} = \left(\begin{array}{c|c|c} x_1 & & \\ \vdots & & \\ x_{l_1} & & \\ \hline & \vdots & \\ \hline & & x_{n_{l_c}} \\ & & \vdots \\ & & x_n \end{array}\right) \begin{array}{l} \}\mathcal{A}_1 \\ \\ \vdots \\ \\ \}\mathcal{A}_{n_c} \end{array}$$

* $n_{l_c} = n - l_{n_c} + 1,\ l_i = |\mathcal{A}_i|$
* $\tilde{P}$ scaled: $\tilde{P}^t\tilde{P} = I,\ \tilde{P}x_c = x$
* $P = (I - \tau A)\tilde{P}$

Multiple vector preserving interpolation

For $B := [x^{(1)}, ..., x^{(r)}]$, $x^{(i)} \in \mathbb{R}^n$,

$$\tilde{P} = \left(\begin{array}{c|c|c} X_1 & & \\ \hline & \vdots & \\ \hline & & X_{n_c} \end{array}\right) \begin{array}{l} \} \mathcal{A}_1 \rightarrow Q_1 R_1 \\ \\ \vdots \\ \\ \} \mathcal{A}_{n_c} \rightarrow Q_{n_c} R_{n_c} \end{array}$$

* $X_i \in \mathbb{R}^{l_i \times r}$
* $\tilde{P}^t \tilde{P} = Q^t Q = I$
* $\tilde{P} B_c = \tilde{P} R = B$

Smoothed Aggregation (A)MG convergence theory

Theorem: (P. Vanek, M. Brezina, J. Mandel, 2001) Assume that there exists a constant, $C_a > 0$, such that for every $u \in \mathbb{R}^n$ and every $l \in \{1, , , L\}$, the following approximation property holds:

$$\sum_i \min_{w \in \mathbb{R}^r} \|u - B^1 w\|^2_{\mathcal{A}_i^l} \leq C_a \|u\|^2_A$$

Then

$$\|\tilde{x} - \mathbf{MG}_1(x, b_1)\|_A \leq \left(1 - \frac{1}{c(L)}\right) \|\tilde{x} - x\|_A$$

Presentation Plan

* Model problems - solver challenges
  * Dirac Wilson system of QCD
  * Schwinger model of QED
* Smoothed Aggregation (A)MG
* Adaptive (A)MG Approaches
* Numerical Results
* Future Work

The need for adaptive (A)MG method(s)

For $G_c \approx A_c^{-1}$ we have the standard error transfer iteration

$$x \leftarrow (I - M^{-t}A)(I - PG_cP^tA)(I - M^{-1}A)x$$

The smoother is typically fixed, e.g. (block) Gauss Seidel. Thus, the sources of potential failure of the above V–cycle are:

* either the interpolation error $x - Px_c$ is large, or
* the coarse solver G_c cannot handle the coarse component x_c of the current level error x

Hence, the need for adapting the MG cycle and thus the coarse-level hierarchy.

To guide the adaptation on the basis of a current algebraically smooth error vector x, represents the main feature of the "adaptive (algebraic) multigrid methods"

The idea of incorporating "smooth vector(s)" in the construction of a method was first introduced by Brandt, McCormick, and Ruge (1983)

Some recent references:

A. Brandt, "Bootstrap AMG", 2001

M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge, "Adaptive smoothed aggregation (αSA)," SIAM J. Sci. Comp., **25** (2004), pp 1896–1920

M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge, "Adaptive algebraic multigrid methods," SIAM J. Sci. Comp., **27** (2006), pp 1261–1286

J. Brannick and L. Zikatanov, "A compatible relaxation and trace minimimization based AMG setup algorithm," DD16 Proceedings (2006)

A basic adaptive (A)MG algorithm

[(T)] For a target $\rho \in (0,1)$ test the method G by applying it to $Ax = 0$:

$$x \leftarrow (I - M^{-t}A)(I - PG_cP^tA)(I - M^{-1}A)x \qquad (1)$$

If $x^tAx \leq \varrho^{2\nu}\, x_0^tAx_0$ **break**,

else set $P = [P,\ P_{\text{new}}]$, $A_c = P^tAP$

for $k = 2 : L-1$

(i) set $x = x_c$, $A = A_k$, $A_c = A_{k+1}$, $M = M_k$, $P = P_k$, $G_c = G_{k+1}$, and update x_c using γ applications of (1)

(ii) $P = [P,\ P_{\text{new}}]$, $A_c = P^tAP$

end for

goto step [(T)] to test the new solver

end if

A basic adaptive (A)MG algorithm ...

In summary, based on a existing (A)MG code, using adaptation, one improves the solver quality by changing its components one level at a time

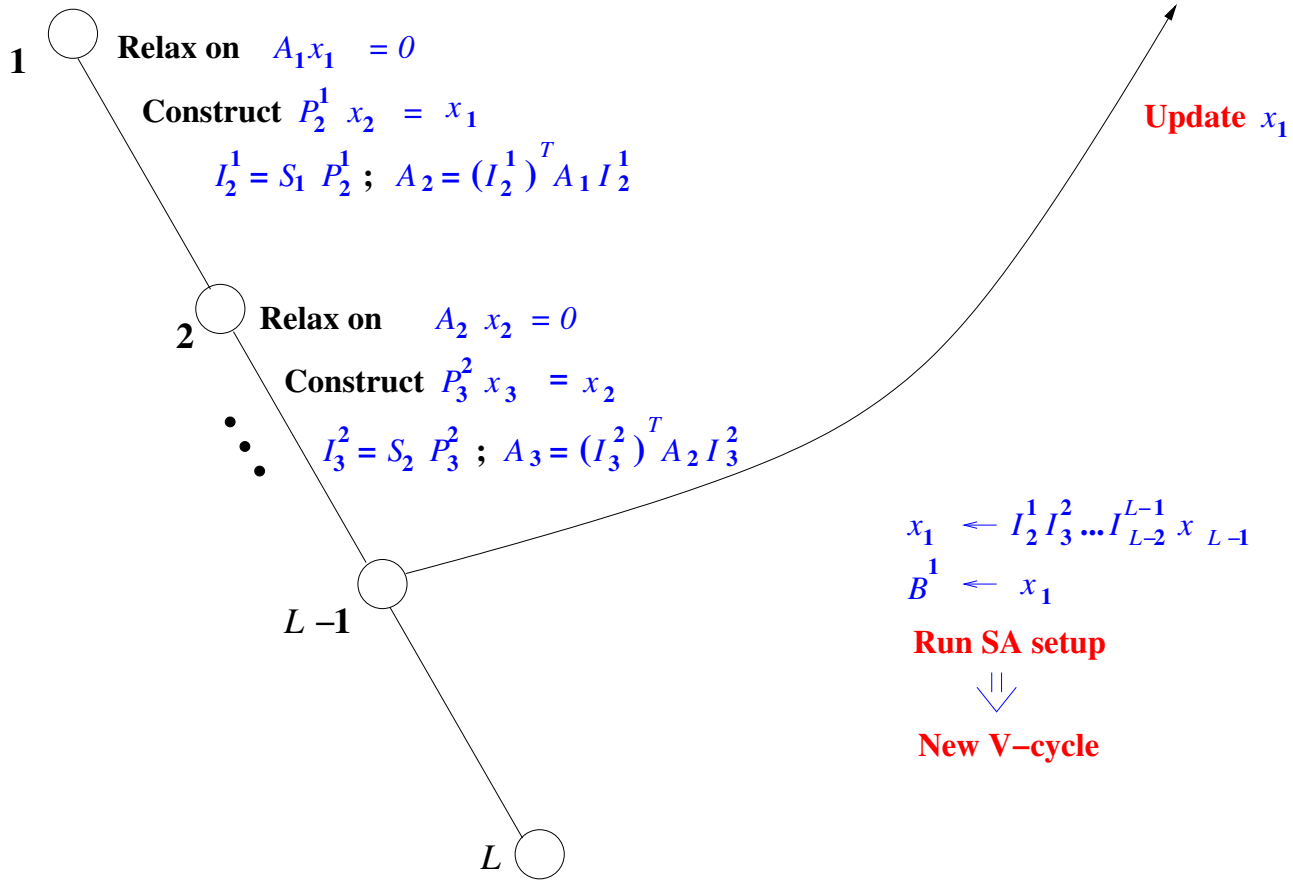
To guarantee improved solver quality after every adaptation step, one needs efficient algorithms for building high–quality updates P_{new} of P

Today we consider the performance of a complex valued SA-based adaptive MG algorithm

* Brezina, et. al., "Adaptive smoothed aggregation", 2004

* Brannick, et. al., "Adaptive smoothed aggregation in lattice QCD", 2006

Adaptive SA Setup



Relation to PCG

* A simple version of the adpative procedure mimics PCG
* Given e_0 relax on $Ae = 0$ such that $e_1 = Se_0$
* Define $P = [e_1]$ and perform two-level MG correction:

$$e_1 \leftarrow (I - \pi_1)Se_0, \quad \pi_1 = P_1(P_1^t A P_1)^{-1} P_1^t A$$

* At kth step define $P_k = [P_{k-1}, e_k]$: $e_k = (I - \pi_k)Se_0$. Then, π_k is the A-orthogonal projector onto:

$$\mathcal{K}_k = \{Se_0, S^2 e_0, ..., S^k e_0\}$$

Presentation Plan

* Model problems - solver challenges
  * Dirac Wilson system of QCD
  * Schwinger model of QED
* Smoothed Aggregation (A)MG
* Adaptive (A)MG Approaches
* Numerical Results
* Future Work

Complex valued adaptive SA solver (M. Clark)

* Variational MG: $A_c = (P_c^f)^t A P_c^f$, $P_c^f = S_f \tilde{P}_c^f$
* Partition fine–level dofs into aggregates
* Given B_f, construct $\tilde{P}_c^f$ and B_c such that $\tilde{P}_c^f B_c = B_f$
* B_f is computed using a multilevel power method based on error propagation operator
* Rely on SA framework to use B_f: in contrast to CG, representative vectors are localized over the aggregates to approximate much larger subspace

Solver components / performance measures

* For all tests we use a point source as rhs and report iteration counts needed to reduce relative residual by 10^8

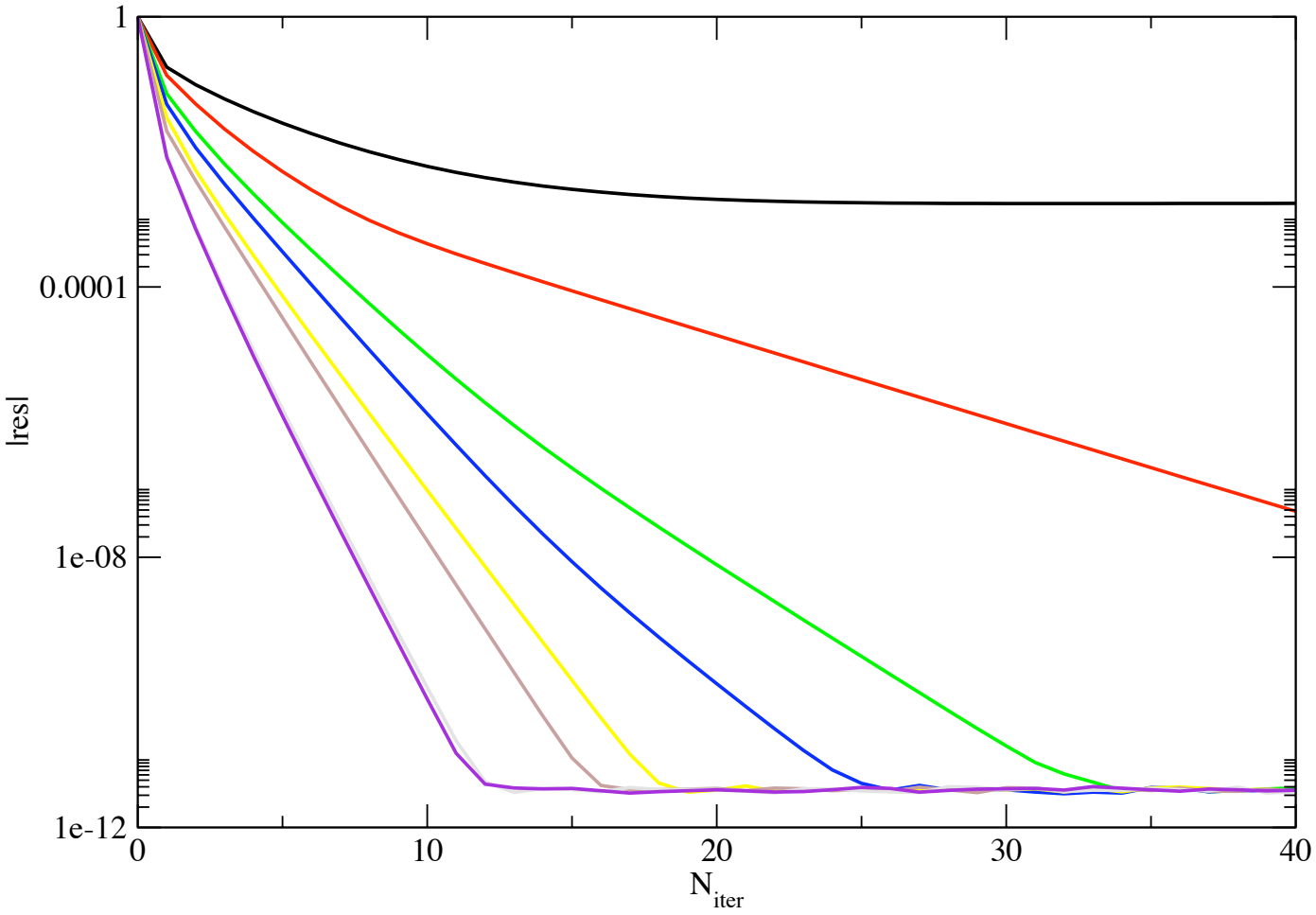
* Smoothers: SOR with $\omega = 1.05$ and ω-Jacobi with $\omega = .6$

* The cost of a single $V(\nu, \mu)$-cycle is approximately $\nu \cdot \mu \cdot GC \cdot OC$ CG iterations, where

$$GC := \frac{\sum_{i=1}^{J} n_{c,i}}{n} \quad \text{and} \quad OC := \frac{\sum_{i=1}^{J} nnz(A_i)}{nnz(A)}$$

3 level, exact coarse solve

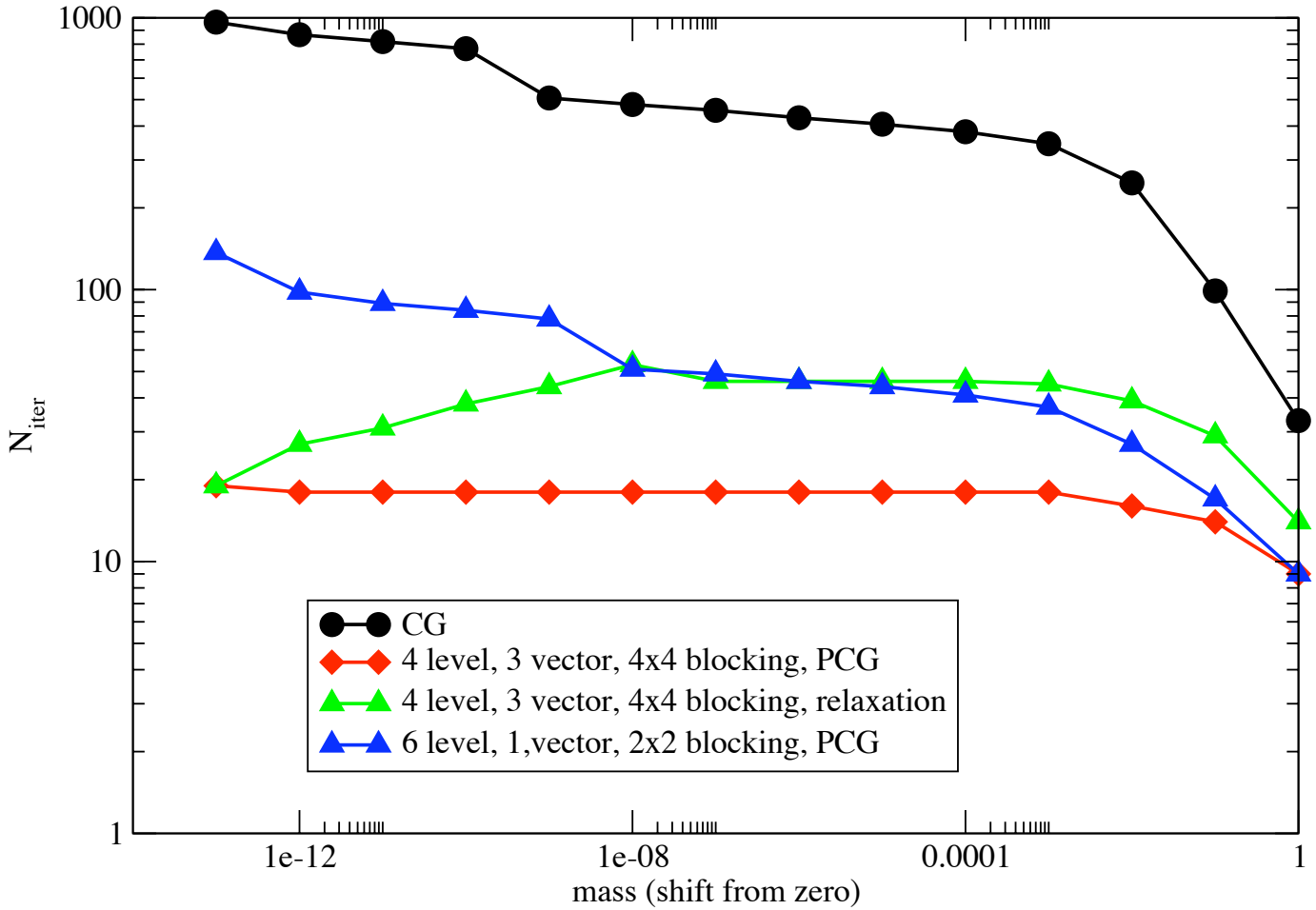
128^2, m=10^{-7}, β=1, 4x4 blocking



Gauge Laplacian: ω-Jacobi ($\omega = .75$), $GC = 1.1$ $OC = 1.1, 1.5, 2.1$

Mass scaling of Gauged Laplacian

V=128^2, β=1.0



ω-Jacobi ($\omega = .75$), $GC = 1.2(1.3)$, $OC = 2.1(2.5)$

Gauge Laplacian

* Ritz values for $N = 128$, $\beta = 1$, and $m = 0$

$$\lambda_1 = 2.1645416870664e-01$$
$$\lambda_2 = 2.2349797163528e-01$$
$$\lambda_3 = 2.2360679843404e-01$$
$$\lambda_4 = 2.3001814768025e-01$$

* Iteration counts needed to reduce relative residual by 10^8 for aSA-PCG, $V(2,2)$-cycles with SOR, $nvec = 1$, 2×2 blocking, $GC = 1.3$, $OC = 2.4$

N/m_0	$1.0e-08$	$1.0e-04$	$1.0e-02$	$1.0e-01$
32	8	7	6	5
64	13	13	10	7
128	14 (425)	13 (343)	11 (232)	8 (107)
256	15	15	13	8
512	15	15	13	9

Gauge Laplacian

* Ritz values for $N = 128$, $\beta = 1$, and $m = 0$

$$\lambda_1 = 2.1645416870664e-01$$
$$\lambda_2 = 2.2349797163528e-01$$
$$\lambda_3 = 2.2360679843404e-01$$
$$\lambda_4 = 2.3001814768025e-01$$

* Iteration counts needed to reduce relative residual by 10^8 for aSA-PCG, $V(2,2)$-cycles with SOR, $nvec = 1$, 4×4 blocking, $GC = 1.1$, $OC = 1.1$

N/m_0	$1.0e-08$	$1.0e-04$	$1.0e-02$	$1.0e-01$
32	8	7	6	5
64	13	13	10	7
128	22 (425)	20 (343)	18 (232)	12 (107)
256	15	15	13	8
512	15	15	13	9

Gauge Laplacian

* Ritz values for $N = 128$, $\beta = 1$, and $m = 0$

$$\lambda_1 = 2.1645416870664e-01$$
$$\lambda_2 = 2.2349797163528e-01$$
$$\lambda_3 = 2.2360679843404e-01$$
$$\lambda_4 = 2.3001814768025e-01$$

* Iteration counts needed to reduce relative residual by 10^8 for aSA-PCG, $V(1,1)$-cycles with SOR, $nvec = 1$, 4×4 blocking, $GC = 1.1$, $OC = 1.1$

N/m_0	$1.0e-08$	$1.0e-04$	$1.0e-02$	$1.0e-01$
32	8	7	6	5
64	13	13	10	7
128	35 (425)	27 (343)	23 (232)	17 (107)
256	15	15	13	8
512	15	15	13	9

Gauge Laplacian

* Iteration counts needed to reduce relative residual by 10^8 for aSA-PCG, $V(1,1)$-cycles with SOR, $nvec = 1$, 4×4 blocking, $GC = 1.1$, $OC = 1.1$

m_0	$1.0e-08$	$1.0e-04$	$1.0e-02$	$1.0e-01$
cg its	435	343	232	107
setup - MVs	277	186	92	15
aSA-PCG its	32	24	18	10

* Approximate setup costs for $m_0 = 1.0e-08$

setup - MVs	411	277	167	123	84
aSA-PCG its	26	32	36	49	51
aSA its	31	38	59	65	78

Schwinger model

* Ritz values for $N = 64$, $\beta = 1$, $m = 0$, top. charge = 1

λ_1	$2.4072960956544e-01$	$7.4457724811804e-17$
λ_2	$2.7543643782929e-01$	$-1.1102230246252e-16$
λ_3	$2.7644821995514e-01$	$6.0828034281699e-03$
λ_4	$2.7644821995513e-01$	$-6.0828034281690e-03$

* Iteration counts needed to reduce relative residual by 10^8 for aSA-PCG applied to H^2, $V(2,2)$-cycles, $nvec = 1$, 2×2 blocking, $GC = 1.2$, $OC = 2.7$

setup / m_0	$1.0e-04$	$1.0e-03$	$1.0e-02$	$1.0e-01$
917	34	32	20	12
1361	28	27	16	7
1654	23 (2113)	20 (1058)	17 (336)	5 (106)

Schwinger model

* Ritz values for $N = 64$, $\beta = 10$, $m = 0$, top. charge = 0

λ_1	$5.2306384031557e-02$	$-2.3671180635050e-02$
λ_2	$5.2306384031563e-02$	$2.3671180635057e-02$
λ_3	$5.2574243587467e-02$	$-7.9255143077699e-03$
λ_4	$5.2574243587460e-02$	$7.9255143077641e-03$

* Iteration counts needed to reduce relative residual by 10^8 for aSA-PCG applied to H^2, $V(2,2)$-cycles, $nvec = 1$, 2×2 blocking, $GC = 1.2$, $OC = 2.7$

setup / m_0	$1.0e-04$	$1.0e-03$	$1.0e-02$	$1.0e-01$
316	34	23	14	10
504	32	23	15	10
1114	22 (1305)	22 (968)	15 (329)	10 (112)

Concluding Remarks

* Several difficulties must be addressed in designing an AMG method for QCD systems
* Adaptive MG methodology naturally handles these difficulties
  * The adaptive setup is designed to expose the *oscillatory* low energy modes
  * Construction of the sequence of coarse problems is based on these computed low energy modes
  * Key issue now is complexity

Current and future works

* Various complex-valued adaptive AMG test codes
  * In the Classical AMG setting by MacLachlan and Osterlee
  * In SA setting by M. Clark in serial and by A. Bessen in parallel
* Systems MG: block smoothers
* MG for $H f = \phi$
* Preconditioners based on Gauge Laplacian
* Parallel implementation

Current and future works

* Various complex-valued adaptive AMG test codes
  * In the Classical AMG setting by MacLachlan and Osterlee
  * In SA setting by M. Clark in serial and by A. Bessen in parallel
* Systems MG: block smoothers
* MG for $H f = \phi$
* Preconditioners based on Gauge Laplacian
* Parallel implementation

"Ideal" interpolation operator

Given R, let

$$K_* = \inf_P \sup_v \frac{\|(I - PR)v\|_{\bar{M}}^2}{\|v\|_A^2}$$

The minimum and minimizer are given by

$$K_* = \frac{1}{\lambda_{\min}(\bar{M}_{ff}^{-1} A_{ff})} \quad \text{and} \quad P_* = \begin{bmatrix} -A_{ff}^{-1} A_{fc} \\ I_c \end{bmatrix}$$

Thus, if K_* is uniformly bounded, then there exists a P, namely P_*, such that the two-level method is uniformly convergent

Trace minimization

* Assume $\mathcal{N}_c$ and $\{\Omega_i\}_{i=1}^{n_c}$ are given. Then

  Each Ω_i contains exactly one index from $\mathcal{N}_c$

* Consider the following affine subspaces of $\mathbb{R}^{n \times n_c}$:

$$\mathcal{X} = \{Q : Q = \begin{bmatrix} W \\ I_c \end{bmatrix}, W \in \mathbb{R}^{n_f \times n_c}\},$$

$$\mathcal{X}_H = \{Q : Q \in \mathcal{X}, \ Q_{ji} = 0 \text{ for all } j \notin \Omega_i, Q\mathbf{1}_c = e\}$$

Here, e is an arbitrary nonzero element of $\mathbb{R}^n : e = \begin{bmatrix} * \\ \mathbf{1}_c \end{bmatrix}$

Definition of P

* Let $I_i \in \mathbb{R}^{n \times n_i}$ be the characteristic function over Ω_i and define $A_i = I_i^T A I_i$.

* Consider

$$P = \arg\min J(Q) := \arg\min \text{trace}(Q^T A Q), \quad Q \in \mathcal{X}_H$$

* The i-th column of the **unique** solution to the minimization problem is given by

$$[P]_i = I_i A_i^{-1} I_i^T M_a e, \quad M_a^{-1} = \sum_{i=1}^{n_c} I_i A_i^{-1} I_i^T$$

On the best approximation to P_* in the trace norm

The matrix M_a^{-1} is the standard additive Schwarz preconditioner for A where the n_c blocks are defined in terms of $\{\Omega_i\}_{i=1}^{n_c}$. Letting each $\Omega_i = \{1 : n_f\}$ and $e = P_* \mathbf{1}_c$ we have that

$J(P_*) = \text{trace}(\mathcal{S}(A))$ and

$$\begin{aligned} \text{trace}(\mathcal{S}(A)) = \sum_{i=1}^{n_c} (\mathcal{S}(A)e_i, e_i) &= \sum_{i=1}^{n_c} \inf_{v: v_c = e_i} (Av, v) \\ \leq \sum_{i=1}^{n_c} (APe_i, Pe_i) &= \text{trace}(P^T AP) \end{aligned}$$

The best approximation property

Theorem. Let $J(\cdot) = |||\cdot|||_A$ and P be the unique solution of the trace-minimization problem. Then

$$|||P_* - P|||_A = \min_{Q \in \mathcal{X}_H} |||P_* - Q|||_A$$

Proof. Let $Q \in \mathcal{X}_H$ be arbitrary. Since, $P_*^T A(Q - P_*) = 0$, we have

$$J(Q) = J(P_* + (Q - P_*)) = \text{trace}(\mathcal{S}(A)) + |||P_* - Q|||_A^2$$

The optimal choice of e

* By direct computation we have that $J(P) = (M_a e, e)$, implying

$$|||P_* - P|||_A^2 = (M_a e, e) - \text{trace}(\mathcal{S}(A))$$

* We can now take the minimum with respect to e on both sides and arrive at

$$|||P_* - P|||_A^2 = \text{trace}[\mathcal{S}(M_a) - \mathcal{S}(A)]$$

where the minimizer is $e = \begin{bmatrix} -M_{a,ff}^{-1} M_{a,fc} \mathbf{1}_c \\ \mathbf{1}_c \end{bmatrix}$