Bootstrap Algebraic Multigrid

Karsten Kahl

Bergische Universität Wuppertal

September 8, 2010
Outline

Computations in QCD
  Properties of Wilson stabilized operators in QED/QCD

Algebraic Multigrid (AMG)
  Motivation of Multigrid
  Ingredients of Algebraic Multigrid
  AMG in QCD Computations

Bootstrap Algebraic Multigrid
  Least Squares Interpolation
  Bootstrap Setup

Numerical Results
  The Gauge Laplace operator

Conclusions and Outlook
Gauge Theory

- Continuum formulation of Dirac operator in presence of background gauge field $A$

$$D = \sum_{\mu=1}^{n} \gamma_\mu \otimes (\partial_\mu + iA_\mu)$$

- $n$-dimensional euclidean space-time
- $\gamma_\mu$ generators of Clifford-Algebra ($\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 0, \gamma_\mu^2 = 1$)
- $A_\mu$ hermitian, traceless
- $\partial_\mu$ partial derivative in spatial direction $\mu$
Lattice Gauge Theory

- Discretization on equidistant lattice with spacing $a$
- Gauge Configuration ($a = 1$)

$$U^x_\mu = e^{-iA_\mu(x+\frac{1}{2}e_\mu)} \approx e^{-i\int_x^{x+e_\mu} A_\mu(s)ds}$$

- Covariant finite differences of spinor $\psi$ at lattice site $x$

$$\hat{\partial}_\mu \psi_x = U^x_\mu \psi_{x+e_\mu} - \psi_x \quad \text{(forward)}$$

$$\hat{\partial}_\mu \psi_x = \psi_x - \left(U^x_\mu e_\mu\right)^H \psi_{x-e_\mu} \quad \text{(backward)}$$

$$\tilde{\partial}_\mu \psi(x) = \frac{1}{2} \left(\hat{\partial}_\mu + \hat{\partial}_\mu\right) \psi_x \quad \text{(central)}$$

$\Rightarrow$ Discretized Dirac operator

$$D = \sum_{\mu=1}^n \gamma_\mu \otimes \left(\hat{\partial}_\mu + \hat{\partial}_\mu\right) = \sum_{\mu=1}^n \gamma_\mu \otimes \tilde{\partial}_\mu \in \mathbb{C}^{N^d n_s n_c}$$
Quantum Chromodynamics
  - Interaction of Quarks and Gluons
  - 4 dimensional space-time
  - $U_\mu \in SU(3)$ (three colors)

Quantum Electrodynamics
  - Interaction of Electrons and Photons
  - 2 dimensional space
  - $U_\mu \in U(1)$ (phase)

Wilson-Stabilization term (Gauge Laplace $\sum_{\mu=1}^{n} \hat{\partial}_\mu \hat{\partial}^\mu$)

$$D = \sum_{\mu=1}^{n} \left( \gamma_\mu \otimes \tilde{\partial}_\mu + \hat{\partial}_\mu \hat{\partial}^\mu \right)$$

Wilson Schwinger operator of QED $S \in \mathbb{C}^{2N^2 \times 2N^2}$

Wilson operator of QCD $D \in \mathbb{C}^{12N^4 \times 12N^4}$
Properties of Wilson formulation

- Discretization with co-variant FD with stabilization term

\[
D\psi = \begin{pmatrix}
A(U) & B(U) \\
-B(U)^H & A(U)
\end{pmatrix}
\begin{pmatrix}
\psi^{(1)} \\
\psi^{(2)}
\end{pmatrix}
\]

- \(A(U)\) Wilson term, \(\begin{pmatrix}
-B(U)^H & B(U)
\end{pmatrix}\) cov. FD of \(D\)

- \(D \in \mathbb{C}^{N^d n_s n_c \times N^d n_s n_c}\) non-hermitian, \(\Gamma_{d+1}\)-hermitian

\[
\Gamma_{d+1} = \gamma_{d+1} \otimes I_{N^d \times n_c}, \quad D^H = \Gamma_{d+1} D \Gamma_{d+1}
\]

- Eigenvalues of \(D\) are real or complex conjugate pairs

- Eigentriplets \((\lambda, \nu^l_\lambda, \nu^r_\lambda), (\bar{\lambda}, \nu^l_{\bar{\lambda}}, \nu^r_{\bar{\lambda}})\)

\[
\nu^l_\lambda = \Gamma_{d+1} \nu^r_{\bar{\lambda}}, \quad \nu^l_{\bar{\lambda}} = \Gamma_{d+1} \nu^r_\lambda
\]
Multigrid in a Nutshell

- Smooth
- Finest Grid
- Fewer Dofs
- First Coarse Grid
- Restriction
- The Multigrid V-cycle
- Prolongation
Ingredients of Algebraic Multigrid (Needed, Given)

- Sparse linear system of equations $Au = f$
- Hierarchy of sparse systems of linear equations $A_l u_l = f_l$
- Appropriate smoothing iterations $S_l = I + B_l (b_l - Au_l)$
- Definition of coarse spaces
  $$C^{m_0} \supset C^{m_1} \supset \ldots \supset C^{m_L}$$
- Definition of restriction and interpolation
  $$R_{l+1}^l : C^{m_l} \rightarrow C^{m_{l+1}} \quad \text{and} \quad P_{l+1}^l : C^{m_{l+1}} \rightarrow C^{m_l}$$
  - Guided by approximation property, complementary to smoother
  - $A$ hermitian positive definite, $R = P^H$
- Definition of operator hierarchy $A = A_0, A_1, \ldots, A_L$
  $$A_{l+1} = R_{l+1}^l A_l P_{l+1}^l$$
Generic Multigrid Algorithm

\[
\text{MG}_l(A_l, b_l)
\]

\[
u_l = 0
\]

\[
\text{for } i = 1, \ldots, \nu_1 \text{ do}
\]

\[
u_l = u_l + B_l(b_l - A_l u_l) \quad \text{“Pre-smoothing”}
\]

\[
\text{end for}
\]

\[
\text{if } l + 1 = L \text{ then}
\]

\[
u_{l+1} = A_{l+1}^{-1} R_{l+1}^l (b_l - A u_l) \quad \text{“Coarse-grid correction”}
\]

\[
\text{else}
\]

\[
u_{l+1} = \text{MG}(A_{l+1}, R_{l+1}^l (b - A u_l)) \quad \text{“Post-smoothing”}
\]

\[
\text{end if}
\]

\[
u_l = u_l + P_{l+1}^l u_{l+1}
\]

\[
\text{for } i = 1, \ldots, \nu_2 \text{ do}
\]

\[
u_l = u_l + B_l(b_l - A_l u_l) \quad \text{“Post-smoothing”}
\]

\[
\text{end for}
\]
Application of AMG to linear system

\[ Du = f, \quad D \in \mathbb{C}^{N^d n_s n_c}, \quad \text{with} \quad D^H = \Gamma_5 D \Gamma_5, \]

where \( D \) is non-hermitian pos. real and \( \Gamma_5 D \) hermitian indefinite.

⇒ Idea: Define \( S_I, A_I \) such that
  
  ▶ The \( \Gamma_5 \)-'ambiguity' is hidden to the AMG algorithm
  ▶ The \( \Gamma_5 \)-hermiticity is preserved in the multigrid hierarchy
Choice of $S_i$ in QCD Computations

**Kaczmarz-Iteration**

(Gauss-Seidel on the normal equations $D^H Du = D^H f$)

<table>
<thead>
<tr>
<th>Forward Sweep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialize $x, r = b - Ax$</td>
</tr>
<tr>
<td><strong>for</strong> $i = 1, \ldots, m$ <strong>do</strong></td>
</tr>
<tr>
<td>$x_i = x_i + \frac{\langle r, De_i \rangle}{| De_i |}^2$</td>
</tr>
<tr>
<td>$r = r - \frac{\langle r, De_i \rangle}{| De_i |}^2 De_i$</td>
</tr>
<tr>
<td><strong>end for</strong></td>
</tr>
</tbody>
</table>

- **Converging Iterative method**
- **Error propagator**

\[
\epsilon_i = \epsilon_i - \frac{\langle D\epsilon, De_i \rangle}{\| De_i \|}^2
\]

\[
= \epsilon_i - \frac{\langle \Gamma_5 D\epsilon, \Gamma_5 De_i \rangle}{\| \Gamma_5 De_i \|}^2
\]

$\Rightarrow$ Blind to $\Gamma_5$-'ambiguity'}
Choice of $A_i$ in QCD Computations

Coarse-grid correction error propagator (two-grid context)

\[ e = e - PD_c^{-1}RDe, \text{ with } D_c = RDP \]

Define $P, R$ such that

\[ D_c^H = \Gamma_5^c D_c \Gamma_5^c \]

\[ e - P (RDP)^{-1} RDe = e - P (R\Gamma_5 DP)^{-1} R\Gamma_5 De \]

▶ With $R = P^H$ and $\Gamma_5 P = P\Gamma_5^c$ \(\Rightarrow\) \(P = \begin{pmatrix} P_{s_1} & \vdots \\ \vdots & P_{s_2} \end{pmatrix}\) we have

\[ D_c^H = \left( P^H DP \right)^H = P^H \Gamma_5 D \Gamma_5 P = \Gamma_5^c D_c \Gamma_5^c \]

\[ e - P \left( P^H DP \right)^{-1} P^H De = e - P \left( P^H \Gamma_5 DP \right)^{-1} P^H \Gamma_5 De \]
Bootstrap AMG

Adaptive Algebraic Multigrid Framework

- Compatible Relaxation
  - Adaptive computation of nested subspaces
  - Quality measurement of grid hierarchies
- Least Squares interpolation
  - Adaptive interpolation
  - Local relaxation
- Bootstrap setup techniques
  - Multigrid discovery of algebraically smooth error
  - Multigrid quality control mechanisms
Least Squares Interpolation

Computation of interpolation weights \((p_i)_j, \quad i \in \mathcal{F}, j \in \mathcal{C}_i\)

\[
\mathcal{L}_{\mathcal{C}_i}(p_i) = \sum_{s=1}^{k} \omega_s (u_i^{(s)} - \sum_{j \in \mathcal{C}_i} (p_i)_j u_j^{(s)})^2 \to \min_{p_i}
\]

- Test vectors \(u^{(1)}, \ldots, u^{(k)} \in V_c \subset \mathbb{C}^n\),
- Set of interpolatory points \(\mathcal{C}_i \subset \mathcal{C}\) for \(i \in \mathcal{F}\)
- Weights \(\omega_s = f(\|Au^{(s)}\|^{-1}) \in \mathbb{R}^+\)
- Splitting of variables \(\Omega = \mathcal{F} \cup \mathcal{C}\), Interpolation \(P\) from \(\mathcal{C}\) to \(\Omega\),

\[
P = \begin{pmatrix} P_{fc} \\ I \end{pmatrix}, \quad p_{ij} \neq 0, i \in \mathcal{F}, j \in \mathcal{C}_i
\]
Least Squares Interpolation – Implementation

- Caliber of interpolation \( c \)
- \( \mathcal{L} \) measure for local accuracy
- Limit \( \mathcal{N}_i \) to graph neighborhood
- Best choice in each iteration
- \( QR \)-decomposition update scheme

Greedy setup

Given TVs \( u^{(l)}, l = 1, \ldots, k \)
Weights \( \omega_l = f(\|Au^{(l)}\|^{-1}) \)

for \( i \) in \( \mathcal{F} \) do

\[ \mathcal{N}_i \subset \mathcal{C}, C_i = \emptyset \]

while \( |C_i| < c \) and \( \mathcal{N}_i \neq \emptyset \) do

\[ g^* = \arg\min_{g \in \mathcal{N}_i} (\mathcal{L}_{C_i \cup g(p_i)}) \]

\[ \mathcal{N}_i = \mathcal{N}_i \setminus g^* \] and \( C_i = C_i \cup g^* \)

end while

end for

Complexity and accuracy control
Bootstrap Setup

No a priori information on algebraically smooth error available!

- TVs initially random, slightly smoothed (small residual locally)
  
  \[ S^n u^{(1)}, \ldots, S^n u^{(k)} \]

- Multigrid hierarchy with interpolation \( P_{l+1}^l \)

  \[
  A_{l+1} = \left( P_{l+1}^l \right)^H A_l P_{l+1}^l, \quad A_0 = A
  \]

  \[
  T_{l+1} = \left( P_{l+1}^l \right)^H T_l P_{l+1}^l, \quad T_0 = I
  \]

- Observation with \( P_l = P_1^0 \cdot \ldots \cdot P_l^{l-1} \),

  \[
  \frac{\langle v_l, v_l \rangle_{A_l}}{\langle v_l, v_l \rangle_{T_l}} = \frac{\langle P_l v_l, P_l v_l \rangle_A}{\langle P_l v_l, P_l v_l \rangle_2}
  \]
Interpolation of coarse-grid vectors \( v_l \) with small gen. energy yields vectors with same small gen. energy on successive finer grids.

Relaxation on gen. EW equation improves approximation quality:

\[
A_l v_l^{(s)} = \lambda_l^{(s)} T_l v_l^{(s)}
\]

Enhancement of TV set and quality control of MG hierarchy.
Bootstrap Setup – Cycling Strategies

- Relax on $Au = 0$, $u \in \mathcal{U}$
- Compute $\mathcal{V}$, s.t., $Av = \lambda Tv$, $v \in \mathcal{V}$
- Relax on $Av = \lambda Tv$, $v \in \mathcal{V}$
- Relax on $Au = 0$, $u \in \mathcal{U}$ and $Av = \lambda Tv$, $v \in \mathcal{V}$
Development of adaptive AMG for LGT

I. Gauge Laplace operator for $U(1)$ gauge configurations
   - Random couplings due to background gauge field
   - Complex valued
   - Locally supported algebraically smooth error

II. Wilson-Schwinger operator of QED
    - System of PDEs
    - $\sigma_3$-hermitian operator
    - Non-hermitian positive-real

III. Wilson-Dirac operator of QCD
     - $\gamma_5$-hermitian operator
     - 3 variables per spin per lattice site
     - $SU(3)$ lattice couplings
The Gauge Laplace operator – $U(1)$ gauge configurations

- Gauge Laplace $A(U) = \hat{\partial}_x \hat{\partial}^x + \hat{\partial}_y \hat{\partial}^y$
- Hermitian part of Wilson-Schwinger operator
- Hermitian positive definite
- Simple example for an operator with background gauge field
- Eigenvectors to small eigenvalues locally supported

- Interested in systems with mass-shift s.t. $\lambda_{\text{min}} = m_0$
  
  \[
  (A(U) + ml) \psi = \varphi
  \]
Gauge Laplace – Numerical Results

- Shifted to $\lambda_{\text{min}} = \frac{1}{N^2}$
- Gauss-Seidel
- Full Coarsening
- Coarsest grid $8 \times 8$
- Operator comp. $\approx 1.6$

<table>
<thead>
<tr>
<th>$\beta \ \backslash \ N$</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.242_W 8</td>
<td>.286_W 8</td>
<td>.478_W 9</td>
<td>.493_V^3 9</td>
</tr>
<tr>
<td>5</td>
<td>.284_W 8</td>
<td>.225_W 8</td>
<td>.389_W 10</td>
<td>.469_V^3 8</td>
</tr>
<tr>
<td>10</td>
<td>.120_W 8</td>
<td>.223_W 7</td>
<td>.349_W 9</td>
<td>.423_V^3 8</td>
</tr>
</tbody>
</table>

- Asymptotic convergence of $V(2,2)$-cycle
- Number of pCG iterations for relative acc. $10^{-8}$
- Setup-cycle with $|\mathcal{U}| = 8$, $|\mathcal{V}| = 16$, 4 relaxations
Algebraic Multigrid for the Wilson-Schwinger operator

Non-hermitian Wilson-Schwinger operator $S_W \in \mathbb{C}^{2N^2 \times 2N^2}$

- **Smoothen**: Kaczmarz
- **Coarsening**: Full coarsening of spin-lattices
- **Interpolation**: Spin-decoupled interpolation $P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}$
\[ F(A) = \{ z \in \mathbb{C} | z = \langle Ax, x \rangle_2, \| x \|_2 = 1 \} \]

\[
\lambda_{\min}\left(\frac{1}{2} (A + A^H)\right) \leq \Re(z) \leq \lambda_{\max}\left(\frac{1}{2} (A + A^H)\right), \quad \text{for all } z \in F(A)
\]
Solve \((S_W + ml) \psi = \varphi\) with mass-shift \(m\)

Moderate shift, \(\lambda_{\text{min}}\left(\frac{1}{2} (S_W + S_W^H) + ml\right) = \frac{1}{N^2}\)

- \(F(S_W + ml)\) positive real
- Gauge Laplace \(A(U) + ml\) pos. def.

“Physical” shift \(m\), \(\min(\Re(\sigma(S_W + ml))) = \frac{1}{N^2}\)

- \(S_W + ml\) positive real
- Gauge Laplace \(A(U) + ml\) indefinite
Moderate mass-shift s.t. $\lambda_{\text{min}}\left(\frac{1}{2} (S_W + S_W^H) + mI\right) = \frac{1}{N^2}$

<table>
<thead>
<tr>
<th>$\beta \backslash N$</th>
<th>32</th>
<th>64</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>.405</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>.259</td>
<td>.387</td>
<td>.459</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>.338</td>
<td>11</td>
</tr>
</tbody>
</table>

- Asymptotic convergence of the stand-alone AMG method
- Number of GMRES iterations for relative acc. $10^{-8}$
- $V^2$-Bootstrap setup-cycle with $k = 8$, $|V| = 16$, 8 relaxations
“Physical” mass-shift s.t. min$(\Re(\sigma(S_W + ml))) = \frac{1}{N^2}$

<table>
<thead>
<tr>
<th>$\beta \setminus N$</th>
<th>32</th>
<th>64</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>♠</td>
<td>♠</td>
<td>♠</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>.259</td>
<td>.434</td>
<td>♠</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>.252</td>
<td>.327</td>
<td>♠</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

- Asymptotic convergence of the $V(2, 2)$-cycle method (Divergence marked by ♠) using Kaczmarz relaxation
- Number of GMRES iterations for relative acc. $10^{-8}$
- $V^2$-Bootstrap setup-cycle with $k = 8$, $|V| = 16$, $(4, 4)$ relaxations
Analysis of stand-alone divergence

- $|\lambda^c| << |\lambda|$
- "Over"-correction
  $\Rightarrow$ Divergence, but (still) good preconditioner

- $\lambda^c < 0 < \lambda$
- Wrong sign correction
Wilson-Dirac operator with “physical” mass-shift s.t.
\[
\min(\Re(\sigma(D_W + ml))) = \frac{1}{N^4}
\]

<table>
<thead>
<tr>
<th>$\beta \setminus N$</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>22/21/23</td>
</tr>
<tr>
<td>11</td>
<td>17/16/15</td>
</tr>
</tbody>
</table>

- Number of pGMRES iterations for relative acc. $10^{-8}$ with $V(4,4)$-cycle preconditioner
- $V^2$-Bootstrap setup-cycle with $k = 8$, $|\mathcal{V}| = 8$, $(4,4)$ relaxations
Conclusions and Outlook

Conclusions

- Bootstrap AMG, adaptive AMG framework
- Adaptive techniques to improve and control AMG methods
- First promising results for LGT applications

Outlook

- Further analysis of the LGT formulation, discretization (?)
- Test bigger problems
- Understand behavior of AMG for $J$-hermitian operators
Thank you for your attention!