An Introduction to Algebraic Multigrid

Wenhao Wu
Chuanbin Li

University of California,Davis

June 3, 2013

Overview

Factors in Multigrid

- Grids and Neighbors;

Factors in Multigrid

- Grids and Neighbors;
- Smoothness;

Factors in Multigrid

- Grids and Neighbors;
- Smoothness;
- Interpolation and Restriction;

Factors in Multigrid

- Grids and Neighbors;
- Smoothness;
- Interpolation and Restriction;
- Coarse Grid Selection.

Introduction

- Apply Multigrid techniques to problems without geometric grids;
- Problems with irregular geometric structures.

Introduction

- AMG is a solver for $Au = f$.
- A is a symmetric M-matrix:
  - Symmetric;
  - Positive definite;
  - Diagonal entries are positive;
  - Off-diagonal entries are non-positive.

Grids

Grids

Identify the grid points with the indices of the unknown quantities.

Grids

Grids

Identify the grid points with the indices of the unknown quantities.

$Au = f$, where

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Then $\{1, \ldots, n\}$ are the fine-grid points, and a coarse grid means some subset of it.

Neighbors, Influence and Dependence

Neighbor

The connection among grid points are defined by the matrix $A = (a_{ij})$. If $a_{ij} \neq 0$ or $a_{ji} \neq 0$, then the ith and jth grid points are neighbors.

Neighbors, Influence and Dependence

Neighbor

The connection among grid points are defined by the matrix $A = (a_{ij})$. If $a_{ij} \neq 0$ or $a_{ji} \neq 0$, then the ith and jth grid points are neighbors.

Strong Dependence

Given a threshold value $0 < \theta \leq 1$, the variable u_i *strongly depends* on u_j if

$$-a_{ij} \geq \theta \max_{k \neq i}\{-a_{ik}\}.$$

Algebraic Smoothness

Algebraic Smoothness

Smooth error is defined to be any error that is not reduced effectively by relaxation.

For a symmetric M-matrix, common relaxation schemes, such as GS, GSRB and weighted Jacobi work well with this definition.

Difference between Algebraic and Geometric Smoothness

An example:

$$-au_{xx} - cu_{yy} + bu_{xy} = 0, \quad (x, y) \in (-1, 1) \times (-1, 1),$$

with homogeneous Dirichlet boundary conditions. We apply the following stencils:

$$D_{xx}^h = \frac{1}{h^2}\begin{pmatrix} 1 & -2 & 1 \end{pmatrix}, D_{yy}^h = \frac{1}{h^2}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

$$D_{xy}^h = \frac{1}{2h^2}\begin{pmatrix} -1 & 1 & \\ 1 & -2 & 1 \\ & 1 & -1 \end{pmatrix}.$$

Difference between Algebraic and Geometric Smoothness

An example:

$$-au_{xx} - cu_{yy} + bu_{xy} = 0, \quad (x, y) \in (-1, 1) \times (-1, 1),$$

Coefficients a, b, and c are locally constant but have different values in different quadrants.

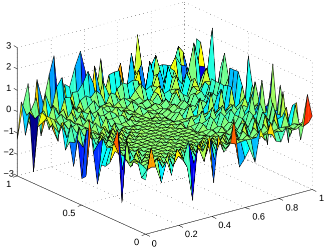
$a = 1$ $c = 1000$ $b = 0$	$a = 1$ $c = 1$ $b = 2$
$a = 1$ $c = 1$ $b = 0$	$a = 1000$ $c = 1$ $b = 0$

Difference between Algebraic and Geometric Smoothness

An example:

$$-au_{xx} - cu_{yy} + bu_{xy} = 0, \quad (x, y) \in (-1, 1) \times (-1, 1),$$

After 10 sweeps of G-S relaxation, the error doesn't change essentially, so we reach algebraic smoothness, but not geometric smoothness.



Interpreting Algebraic Smoothness 1

Algebraic smoothness infers small residual.

If $A = D - L - L^{\mathrm{T}}$ is a symmetric M-matrix and $\mathbf{e}$ is a smooth error, then it can be shown that

$$\|\mathbf{r}\|_{D^{-1}} \ll \|\mathbf{e}\|_A, \quad \text{or} \quad (D^{-1}A\mathbf{e}) \ll (\mathbf{e}, A\mathbf{e}).$$

Interpreting Algebraic Smoothness 1

Algebraic smoothness infers small residual.

If $A = D - L - L^{\mathrm{T}}$ is a symmetric M-matrix and $\mathbf{e}$ is a smooth error, then it can be shown that

$$\|\mathbf{r}\|_{D^{-1}} \ll \|\mathbf{e}\|_A, \quad \text{or} \quad (D^{-1}A\mathbf{e}) \ll (\mathbf{e}, A\mathbf{e}).$$

Writing in components yields $\displaystyle\sum_i \frac{r_i^2}{a_{ii}} \ll a_{ii}|e_i|$.

Interpreting Algebraic Smoothness 1

Algebraic smoothness infers small residual.

If $A = D - L - L^{\mathrm{T}}$ is a symmetric M-matrix and $\mathbf{e}$ is a smooth error, then it can be shown that

$$\|\mathbf{r}\|_{D^{-1}} \ll \|\mathbf{e}\|_A, \quad \text{or} \quad (D^{-1}A\mathbf{e}) \ll (\mathbf{e}, A\mathbf{e}).$$

Writing in components yields $\sum_i \frac{r_i^2}{a_{ii}} \ll a_{ii}|e_i|$.

At least on average $|r_i| \ll a_{ii}|e_i|$.

Interpreting Algebraic Smoothness 1

Algebraic smoothness infers small residual.

If $A = D - L - L^{\mathrm{T}}$ is a symmetric M-matrix and $\mathbf{e}$ is a smooth error, then it can be shown that

$$\|\mathbf{r}\|_{D^{-1}} \ll \|\mathbf{e}\|_A, \quad \text{or} \quad (D^{-1}A\mathbf{e}) \ll (\mathbf{e}, A\mathbf{e}).$$

Writing in components yields $\sum_i \frac{r_i^2}{a_{ii}} \ll a_{ii}|e_i|$.
At least on average $|r_i| \ll a_{ii}|e_i|$.
We write it loosely as

$$A\mathbf{e} \approx 0.$$

Interpreting Algebraic Smoothness 1

$$A\mathbf{e} \approx 0.$$

Interpreting Algebraic Smoothness 1

$$A\mathbf{e} \approx 0.$$

$$a_{ii}e_i \approx -\sum_{j \neq i} a_{ij}e_j.$$

Interpreting Algebraic Smoothness 1

$$A\mathbf{e} \approx 0.$$

$$a_{ii}e_i \approx -\sum_{j \neq i} a_{ij}e_j.$$

This means that e_i can be appoximately expressed as a linear combination of its neighbors.

Interpreting Algebraic Smoothness 2

A smooth error varies slowly among strong dependent grids.

If $\mathbf{e}$ is a smooth error, then for most i,

$$\sum_{j\neq i}\left(\frac{|a_{ij}|}{a_{ii}}\right)\left(\frac{e_i - e_j}{e_i}\right)^2 \ll 1, \quad 1 \leq i \leq n.$$

Interpreting Algebraic Smoothness 2

A smooth error varies slowly among strong dependent grids.

If $\mathbf{e}$ is a smooth error, then for most i,

$$\sum_{j\neq i}\left(\frac{|a_{ij}|}{a_{ii}}\right)\left(\frac{e_i - e_j}{e_i}\right)^2 \ll 1, \quad 1 \leq i \leq n.$$

It justifies the idea that we can interpolate a find grid point from its strong dependent neighbors on the coarse grid.

Tasks before Us

Wenhao Wu
Chuanbin Li

Tasks before Us

- Given a fine grid, select a coarse grid so that smooth error can be represented well;

Tasks before Us

- Given a fine grid, select a coarse grid so that smooth error can be represented well;
- Define an interpolation operator so that smooth error can be transferred accurately from the coarse grid to the fine grid;

Tasks before Us

- Given a fine grid, select a coarse grid so that smooth error can be represented well;
- Define an interpolation operator so that smooth error can be transferred accurately from the coarse grid to the fine grid;
- Define a restriction operator.

Interpolation

Assume that a coarse grid has been selected. On the fine grid:

$$a_{ii} e_i \approx -\sum_{j \neq i} a_{ij} e_j = \sum_{j \in N_i} a_{ij} e_j,$$

where N_i stands for the neighbors of i.

Interpolation

Assume that a coarse grid has been selected. On the fine grid:

$$a_{ii}e_i \approx -\sum_{j \neq i} a_{ij}e_j = \sum_{j \in N_i} a_{ij}e_j,$$

where N_i stands for the neighbors of i.

$$N_i = C_i \cup D_i^s \cup D_i^w,$$

where

Interpolation

Assume that a coarse grid has been selected. On the fine grid:

$$a_{ii}e_i \approx -\sum_{j\neq i} a_{ij}e_j = \sum_{j\in N_i} a_{ij}e_j,$$

where N_i stands for the neighbors of i.

$$N_i = C_i \cup D_i^s \cup D_i^w,$$

where

- C_i stands for the neighboring coarse grid points that strongly influence i, which is the coarse interpolatory set for i;

Interpolation

Assume that a coarse grid has been selected. On the fine grid:

$$a_{ii}e_i \approx -\sum_{j \neq i} a_{ij}e_j = \sum_{j \in N_i} a_{ij}e_j,$$

where N_i stands for the neighbors of i.

$$N_i = C_i \cup D_i^s \cup D_i^w,$$

where

- C_i stands for the neighboring coarse grid points that strongly influence i, which is the coarse interpolatory set for i;
- D_i^s stands for neighboring fine grid points that strongly influence i;

Interpolation

Assume that a coarse grid has been selected. On the fine grid:

$$a_{ii} e_i \approx -\sum_{j \neq i} a_{ij} e_j = \sum_{j \in N_i} a_{ij} e_j,$$

where N_i stands for the neighbors of i.

$$N_i = C_i \cup D_i^s \cup D_i^w,$$

where

- C_i stands for the neighboring coarse grid points that strongly influence i, which is the coarse interpolatory set for i;
- D_i^s stands for neighboring fine grid points that strongly influence i;
- D_i^w stands for neighboring fine grid points that weakly influence i.

Interpolation

Our goal is to find a operator I_{2h}^h s.t.

$$(I_{2h}^h \mathbf{e})_i = \begin{cases} e_i, & \text{if } i \in C; \\ \sum_{j \in C_i} \omega_{ij} e_j, & \text{if } i \in F. \end{cases}$$

Interpolation

Our goal is to find a operator I_{2h}^{h} s.t.

$$(I_{2h}^{h}\mathbf{e})_i = \begin{cases} e_i, & \text{if } i \in C; \\ \sum_{j\in C_i} \omega_{ij} e_j, & \text{if } i \in F. \end{cases}$$

On the fine grid we have

$$a_{ii} e_i \approx -\sum_{j\in C_i} a_{ij} e_j - \sum_{j\in D_i^s} a_{ij} e_j - \sum_{j\in D_i^w} a_{ij} e_j .$$

Interpolation

Our goal is to find a operator I_{2h}^{h} s.t.

$$(I_{2h}^{h}\mathbf{e})_i = \begin{cases} e_i, & \text{if } i \in C; \\ \sum_{j \in C_i} \omega_{ij} e_j, & \text{if } i \in F. \end{cases}$$

On the fine grid we have

$$a_{ii} e_i \approx -\sum_{j \in C_i} a_{ij} e_j - \sum_{j \in D_i^s} a_{ij} e_j - \sum_{j \in D_i^w} a_{ij} e_j .$$

Represent e_j in the last two sums with e_i and/or e_j in the coarse grid.

Interpolation

$$a_{ii}e_i \approx -\sum_{j\in C_i} a_{ij}e_j - \sum_{j\in D_i^s} a_{ij}e_j - \sum_{j\in D_i^w} a_{ij}e_j.$$

First we deal with D_i^w simply by attributing it to diagonal terms, i.e.

$$(a_{ii} + \sum_{j\in D_i^w} a_{ij})e_i \approx -\sum_{j\in C_i} a_{ij}e_j - \sum_{j\in D_i^s} a_{ij}e_j.$$

Interpolation

$$a_{ii} e_i \approx -\sum_{j \in C_i} a_{ij} e_j - \sum_{j \in D_i^s} a_{ij} e_j - \sum_{j \in D_i^w} a_{ij} e_j .$$

First we deal with D_i^w simply by attributing it to diagonal terms, i.e.

$$(a_{ii} + \sum_{j \in D_i^w} a_{ij}) e_i \approx -\sum_{j \in C_i} a_{ij} e_j - \sum_{j \in D_i^s} a_{ij} e_j .$$

- If the dependence is underestimated, then e_i strongly depends on $e_j \in D_i^w$. Since smooth error varies slowly among them, it is justified to replace e_j with e_i;

Interpolation

$$a_{ii}e_i \approx -\sum_{j\in C_i} a_{ij}e_j - \sum_{j\in D_i^s} a_{ij}e_j - \sum_{j\in D_i^w} a_{ij}e_j.$$

First we deal with D_i^w simply by attributing it to diagonal terms, i.e.

$$(a_{ii} + \sum_{j\in D_i^w} a_{ij})e_i \approx -\sum_{j\in C_i} a_{ij}e_j - \sum_{j\in D_i^s} a_{ij}e_j.$$

- If the dependence is underestimated, then e_i strongly depends on $e_j \in D_i^w$. Since smooth error varies slowly among them, it is justified to replace e_j with e_i;
- If the dependence is weak, then $a_{ij} \ll a_{ii}$, making the error insignificant.

Interpolation

$$(a_{ii} + \sum_{j \in D_i^w} a_{ij}) e_i \approx - \sum_{j \in C_i} a_{ij} e_j - \sum_{j \in D_i^s} a_{ij} e_j .$$

Now we want to treat D_i^s. Experience shows that it is better to attribute them to points in C_i .

Interpolation

$$(a_{ii} + \sum_{j \in D_i^w} a_{ij}) e_i \approx - \sum_{j \in C_i} a_{ij} e_j - \sum_{j \in D_i^s} a_{ij} e_j .$$

Now we want to treat D_i^s. Experience shows that it is better to attribute them to points in C_i .

$$e_j \approx \frac{\sum_{k \in C_i} a_{jk} e_k}{\sum_{k \in C_i} a_{jk}}, \quad j \in D_i^s .$$

Interpolation

$$(a_{ii} + \sum_{j \in D_i^w} a_{ij}) e_i \approx - \sum_{j \in C_i} a_{ij} e_j - \sum_{j \in D_i^s} a_{ij} e_j .$$

Now we want to treat D_i^s. Experience shows that it is better to attribute them to points in C_i .

$$e_j \approx \frac{\sum_{k \in C_i} a_{jk} e_k}{\sum_{k \in C_i} a_{jk}}, \quad j \in D_i^s .$$

This approximation requires that if e_i and e_j are strongly connected, then $C_i \cap C_j \neq \varnothing$.

Interpolation

After Combining the above and calculation, we get

$$(I_{2h}^{h}\mathbf{e})_i = \begin{cases} e_i, & \text{if } i \in C; \\ \sum_{j\in C_i} \omega_{ij} e_j, & \text{if } i \in F; \end{cases}$$

$$\text{where} \quad \omega_{ij} = -\frac{a_{ij} + \sum_{m\in D_i^s} \frac{a_{im}a_{mj}}{\sum_{k\in C_i} a_{mk}}}{a_{ii} + \sum_{n\in D_i^w} a_{in}}.$$

Fundamental Premise of Coarse Grid

Requirement 1

Smooth error can be approximated accurately on the coarse grid.

Requirement 2

Smooth functions can be interpolated accurately from the coarse grid.

Requirement 3

Coarse grid must have substantially fewer points than the fine grid.

Heuristic Criteria for Selection of C-points

H-1: Interpolation Accuracy

For each F-point i, every point $j \in S_i$ that strongly influences i either should be in the coarse interpolatory set C_i or should strongly depend on at least one point in C_i

H-2: Fewest Points in Coarse Grid

The set of coarse points C should be a maximal subset of all points with the property that no C-point strongly depends on another C-point.

In selection of coarse grid points, H-1 is going to be enforced rigorously while H-2 is going to be used as a guide.

The Coloring Scheme - Pass 1/2

Select a set of C-points that have good approximation property and tend to satisfy H-2.

- Define λ_i: a measure of potential quality of point i as a C-point.
- Initialize $\lambda_i = |S_i^T|$
- While there are points unassigned to C or F
  - Assign $\hat{i} = \arg \max_{i \text{ unassigned}} \{\lambda_i\}$ to C.
  - Assign all unassigned points $j \in S_i^T$ to F.
  - For each $j \in S_i^T$ in the above step, $\lambda_k = \lambda_k + 1$ for all unassigned points $k \in S_j$.

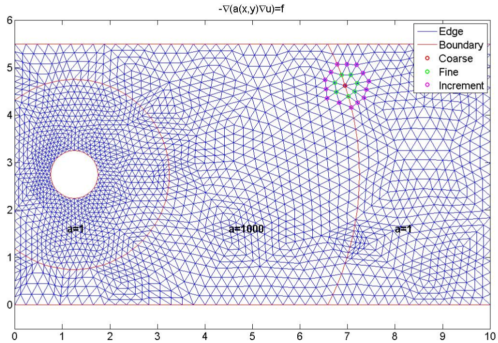
The Coloring Scheme - Pass 2/2

Ensure that H-1 is satisfied: no 2 F-F dependency without depending on a common C-point.

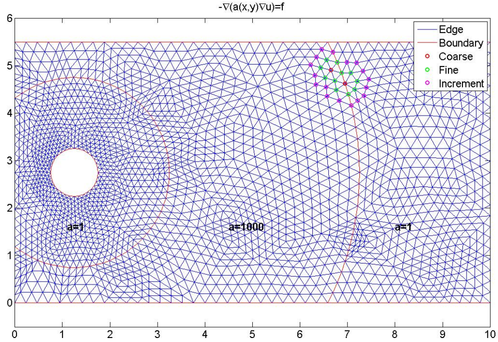
- (Following Pass 1) While there are F-F pairs unchecked
  - Check the next pair of F-points
  - If 1 of the 2 F-points strongly depends on the other, but they don't strongly depend on a common C-point, change one of them to be C-point.

Remark: The pass 2 ensures H-1 to be satisfied, which in turn guarantees that the interpolation operation is possible.

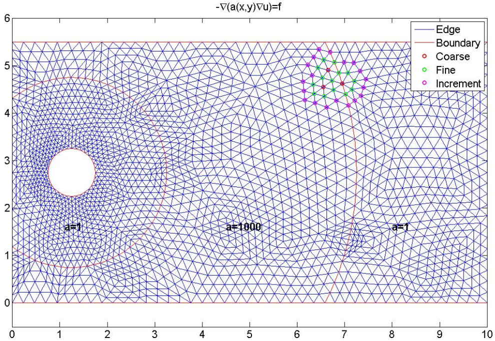
The Coloring Scheme - Example: Pass 1/2



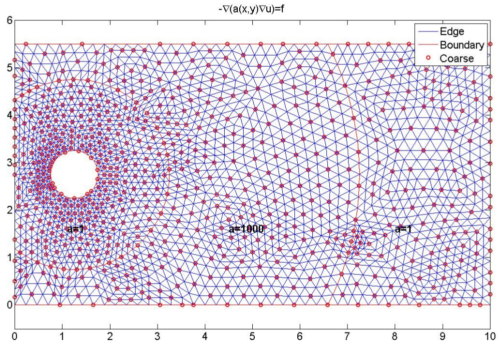
The Coloring Scheme - Example: Pass 1/2



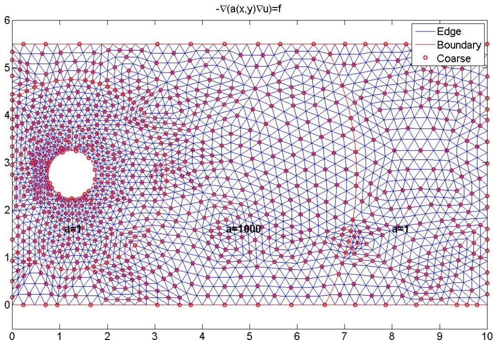
The Coloring Scheme - Example: Pass 1/2



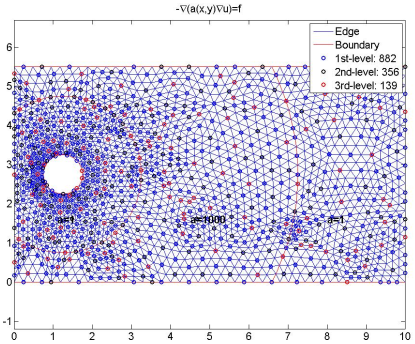
The Coloring Scheme - Example: Pass 1/2



The Coloring Scheme - Example: Pass 2/2



The Coloring Scheme - Example: AMG Grid Hierarchy



Coarse-Grid Operators

- Restriction operator: $I_h^{2h} = (I_{2h}^h)^T$
- Coarse-grid operator: $A^{2h} = I_h^{2h} A^h I_{2h}^h$

V-Cycle AMG

Recursive description: $\mathbf{v}^h \leftarrow \mathrm{AMG}(A^h, \mathbf{v}^h, \mathbf{f}^h)$.

- Relax ν_1 times on $A^h\mathbf{v}^h = \mathbf{f}^h$ with a given initial guess $\mathbf{v}^h$
- If Ω^h =coarsest grid, then solve $A^h\mathbf{v}^h = \mathbf{f}^h$ directly; Else

$$\mathbf{r}^{2h} \leftarrow I_h^{2h}(\mathbf{f}^h - A^h\mathbf{v}^h),$$

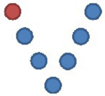
$$A^{2h} = I_h^{2h} A^h I_{2h}^h,$$

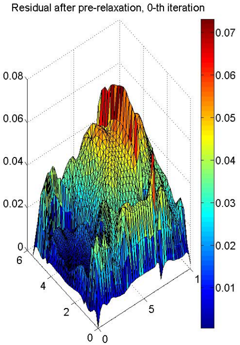
$$\mathbf{e}^{2h} \leftarrow \mathrm{AMG}(A^{2h}, \mathbf{0}, \mathbf{r}^{2h})$$

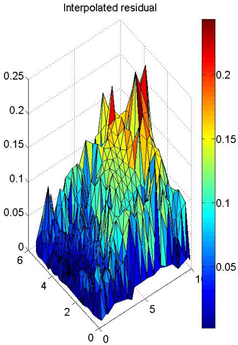
- Prolong $\mathbf{e}^h = I_{2h}^h \mathbf{e}^{2h}$
- Correct $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h \mathbf{e}^{2h}$
- Relax ν_2 times on $A^h\mathbf{v}^h = \mathbf{f}^h$ with a given initial guess $\mathbf{v}^h$

V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

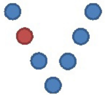


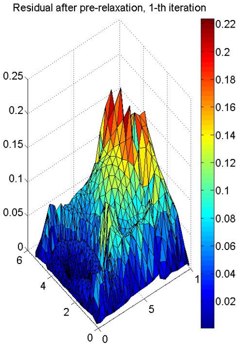


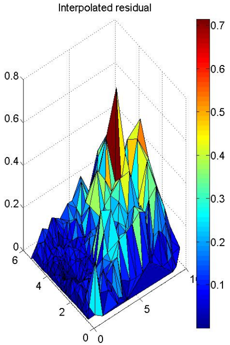


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

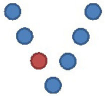


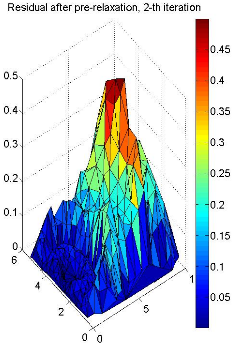


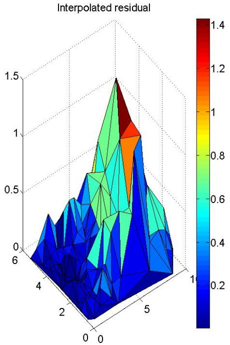


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

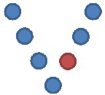


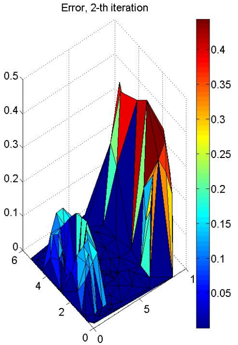


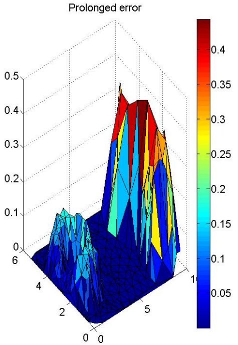


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

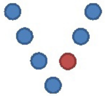


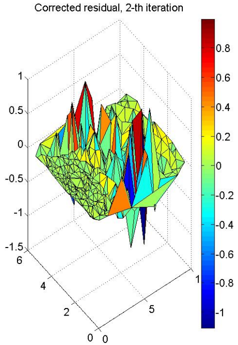


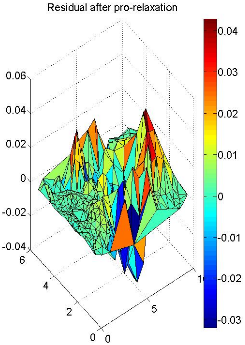


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

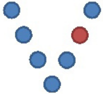


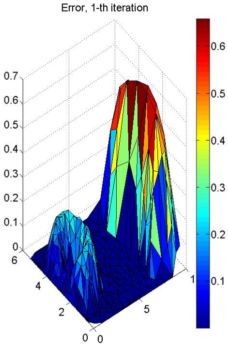


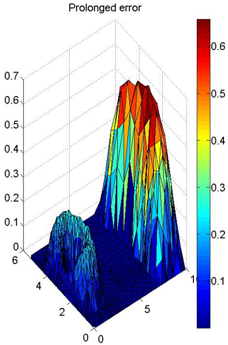


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

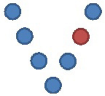


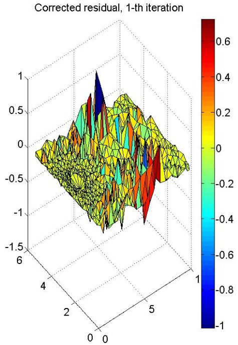


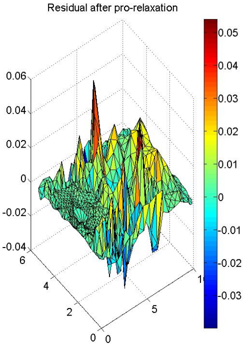


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

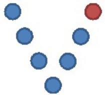


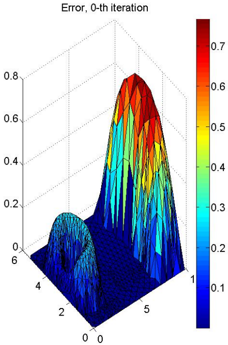


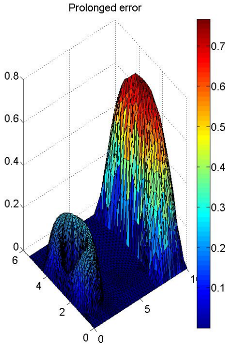


V-Cycle AMG - Example

- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

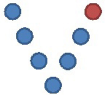


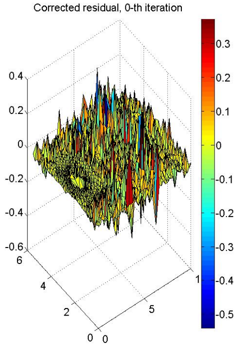


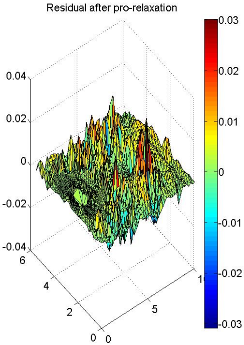


V-Cycle AMG - Example

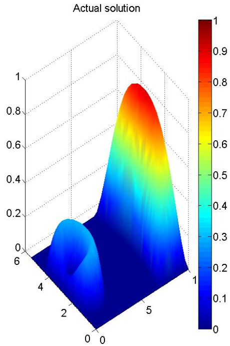
- Pre-relaxation & Interpolation
- Prolongation
- Correction & Pro-relaxation

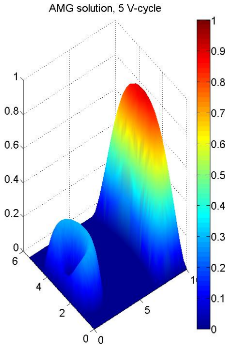




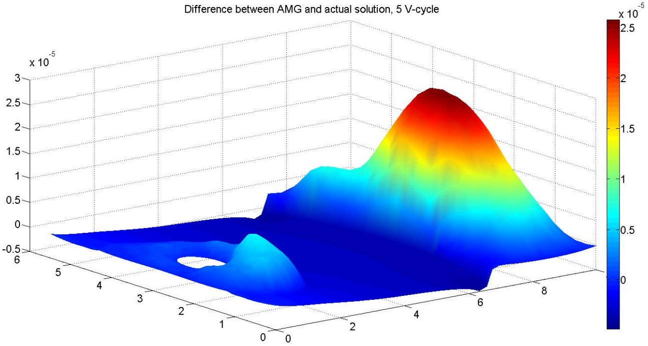


V-Cycle AMG - Example



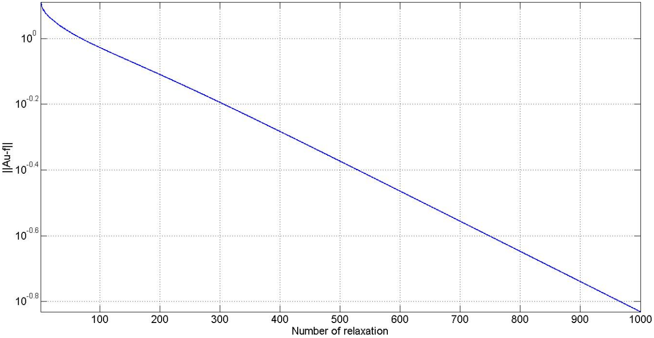


V-Cycle AMG - Example



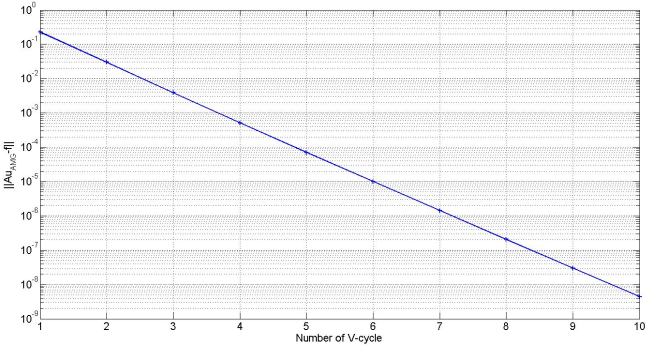
V-Cycle AMG - Example

Convergence of the residual for pure relaxation method (Weighted Jacobi Iteration with $\omega = 2/3$.)



V-Cycle AMG - Example

Convergence of the AMG method (Weighted Jacobi Iteration with $\omega = 2/3, \nu_1 = \nu_2 = 2$, depth of V-Cycle=4.)



References

William L. Briggs, Van Emden Henson, Stefe F. McCormick (2000)
A Multigrid Ttutorial, second edition.

K.Stüben (1999)
Algebraic Multigrid (AMG): An Introduction with Applications.

U. Trottenberg, C. Oosterlee, A. Schüller(2001)
Multigrid.

Thank you