Improving Robustness in Algebraic Multigrid

Scott MacLachlan

Delft Institute of Applied Mathematics,
Delft University of Technology
s.p.maclachlan@ewi.tudelft.edu

December 12, 2006

Roadmap

- Why AMG?
  - Background & motivation
- Adaptive multigrid
  - Fewer assumptions
  - Improved robustness
- Algebraic coarsening
  - Heuristics and theory
  - Compatible relaxation
- Current & future challenges
  - New problems
  - Theoretical questions

Target Applications

- Heterogeneous
  - Variations in material properties (porous media)
  - Multi-physics/multi-model (fluid-structure)
- Stochastic & uncertain
  - Kriging/geostatistics (porous media)
  - Monte-Carlo (lattice QCD)
- Unstructured meshes
  - Irregular geometry
  - Local refinement
- New challenges
  - Complex-valued systems (lattice QCD)
  - Indefinite systems (Helmholtz)

Starting Point

Discretization must represent important features of model

- preserve symmetry and/or definiteness
- small elements (to capture heterogeneity)
- irregular meshes

Look for efficient solvers for heterogeneous discrete models

- Large matrix sizes
- Large condition numbers
- Multiscale structure of operator

Stationary Iterative Methods

- Want to improve approximation, $x^{(0)}$, to $x = A^{-1}b$
- Residual, $r^{(0)}$, is a measure of the error

$$r^{(0)} = b - Ax^{(0)} = Ax - Ax^{(0)} = A(x - x^{(0)})$$

- Choose $B^{-1} \approx A^{-1}$
- Take $x^{(1)} = x^{(0)} + B^{-1}r^{(0)}$

Error propagation form: $e^{(1)} = (I - B^{-1}A)e^{(0)}$

Stationary Iterative Methods

- Want to improve approximation, $x^{(0)}$, to $x = A^{-1}b$
- Residual, $r^{(0)}$, is a measure of the error

$$r^{(0)} = b - Ax^{(0)} = Ax - Ax^{(0)} = A(x - x^{(0)})$$

- Choose $B^{-1} \approx A^{-1}$
- Take $x^{(1)} = x^{(0)} + B^{-1}r^{(0)}$

Error propagation form: $e^{(1)} = (I - B^{-1}A)e^{(0)}$

$$e^{(2)} = (I - B^{-1}A)e^{(1)}$$

Stationary Iterative Methods

- Want to improve approximation, $x^{(0)}$, to $x = A^{-1}b$
- Residual, $r^{(0)}$, is a measure of the error

$$r^{(0)} = b - Ax^{(0)} = Ax - Ax^{(0)} = A(x - x^{(0)})$$

- Choose $B^{-1} \approx A^{-1}$
- Take $x^{(1)} = x^{(0)} + B^{-1}r^{(0)}$

Error propagation form: $e^{(1)} = (I - B^{-1}A)e^{(0)}$

$$e^{(2)} = (I - B^{-1}A)^2 e^{(0)}$$

Stationary Iterative Methods

- Want to improve approximation, $x^{(0)}$, to $x = A^{-1}b$
- Residual, $r^{(0)}$, is a measure of the error

$$r^{(0)} = b - Ax^{(0)} = Ax - Ax^{(0)} = A(x - x^{(0)})$$

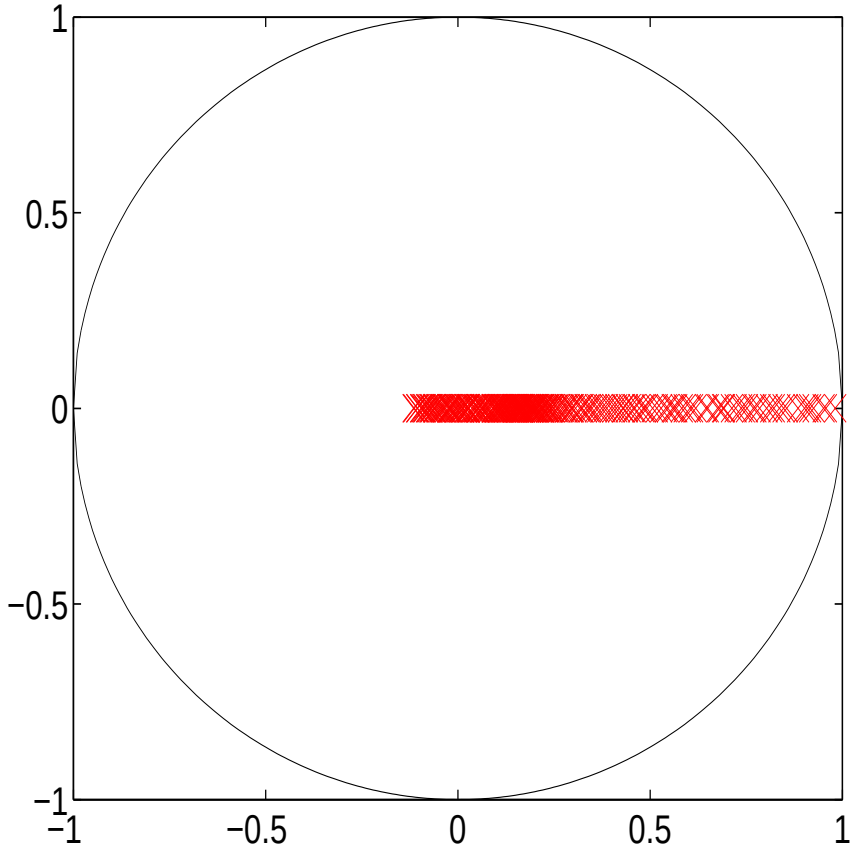
- Choose $B^{-1} \approx A^{-1}$
- Take $x^{(1)} = x^{(0)} + B^{-1}r^{(0)}$

Error propagation form:

$$\begin{aligned} e^{(1)} &= (I - B^{-1}A)e^{(0)} \\ e^{(2)} &= (I - B^{-1}A)^2 e^{(0)} \\ &\vdots \\ e^{(n)} &= (I - B^{-1}A)^n e^{(0)} \end{aligned}$$

Convergence of Stationary Iterations

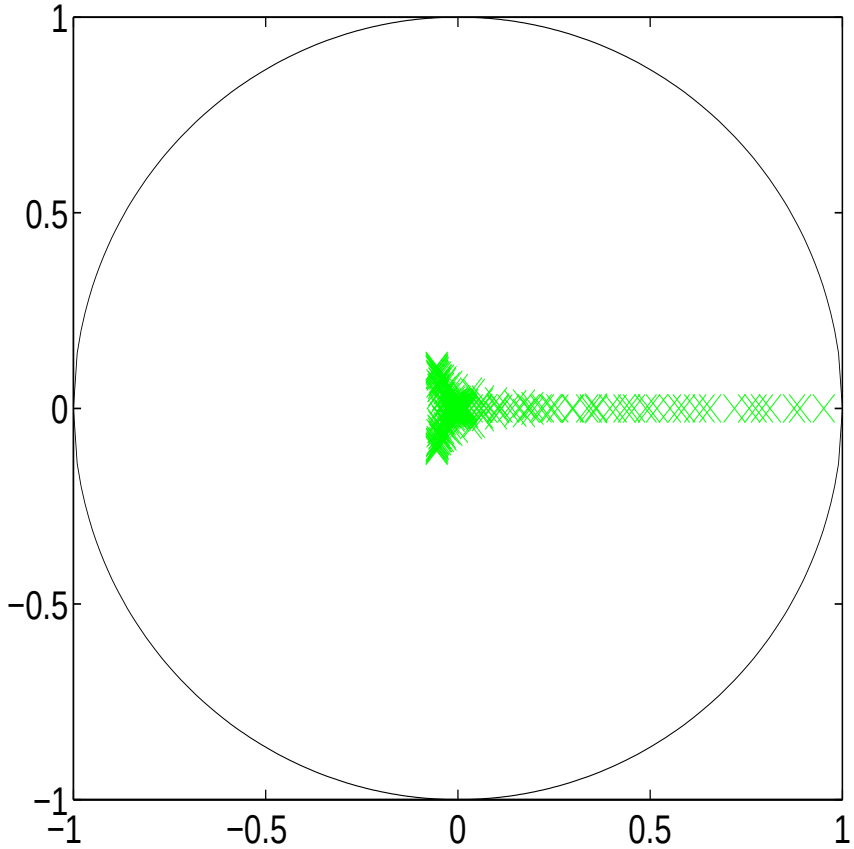
Convergence depends on spectrum of $I - B^{-1}A$



Weighted Jacobi Iteration: $e^{(n)} = (I - \frac{3}{4}D^{-1}A)^n e^{(0)}$

Convergence of Stationary Iterations

Convergence depends on spectrum of $I - B^{-1}A$



Gauss-Seidel Iteration: $e^{(n)} = (I - L^{-1}A)^n e^{(0)}$

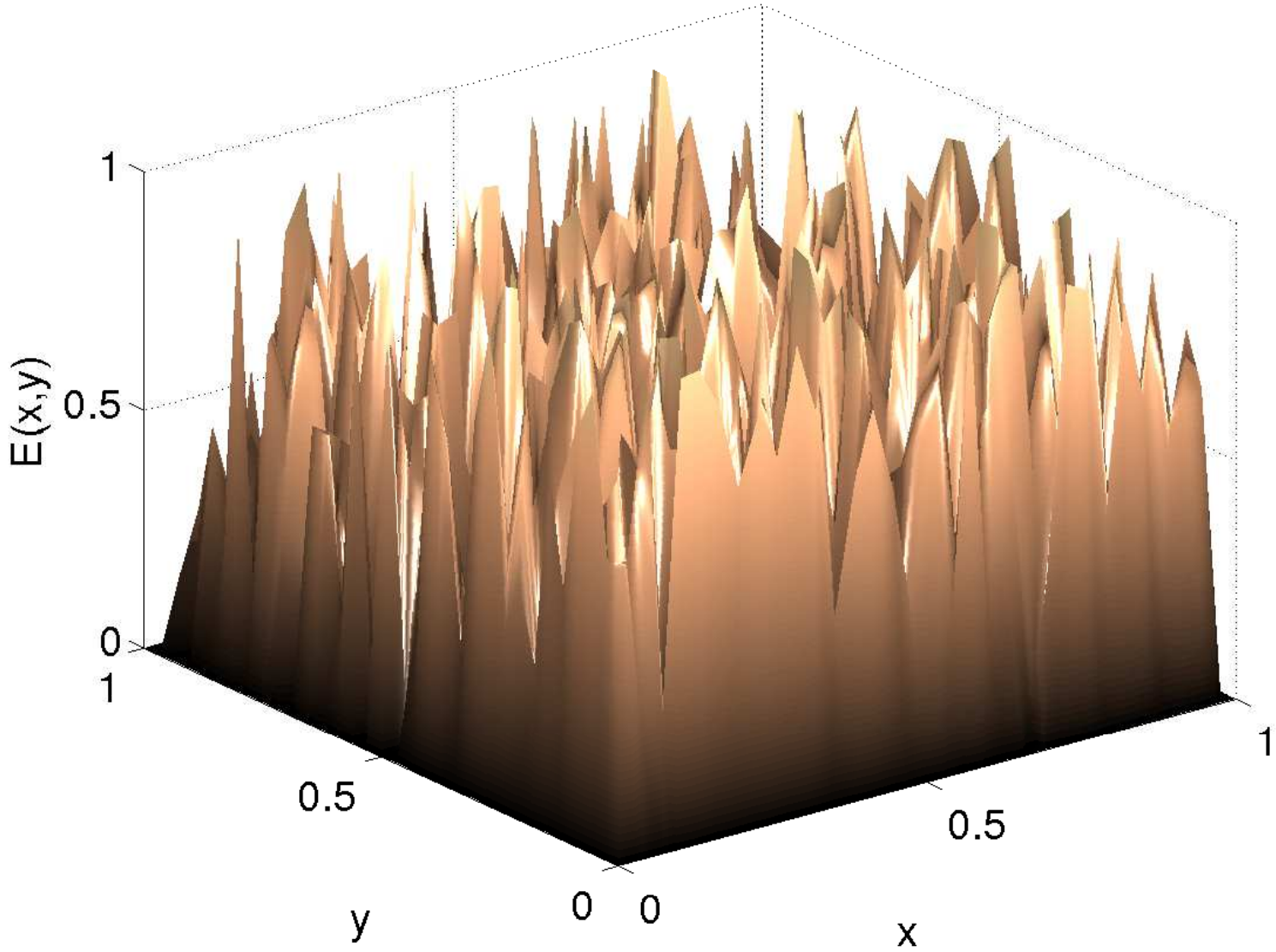
Failing in a Structured Way

Small $B^{-1}A$-Rayleigh quotients cause trouble

$$\lambda_{\max}(I - B^{-1}A) = 1 - \min_{y} \frac{y^T A y}{y^T B y}$$

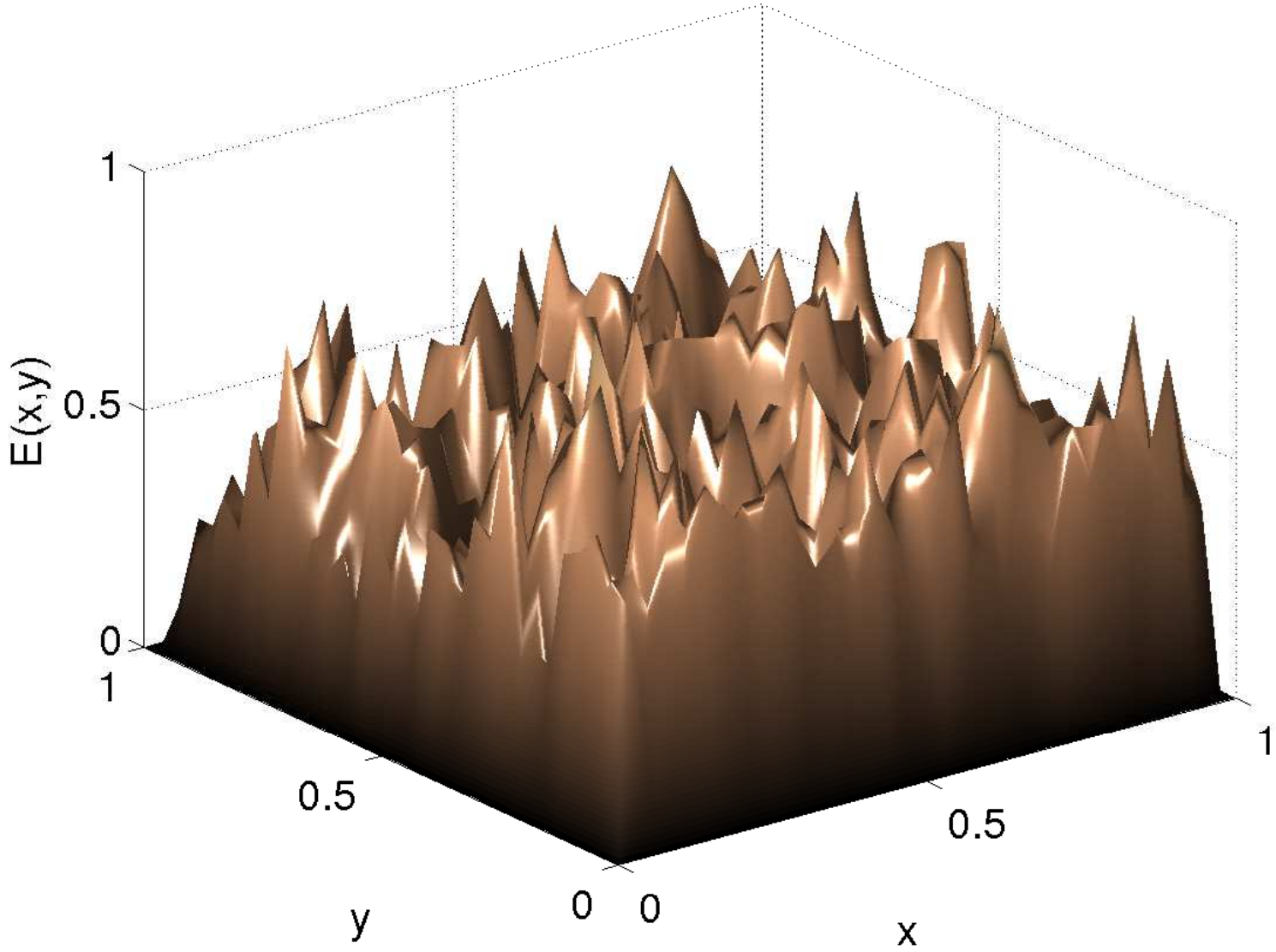
Can we use this to our advantage?

Smoothing Property



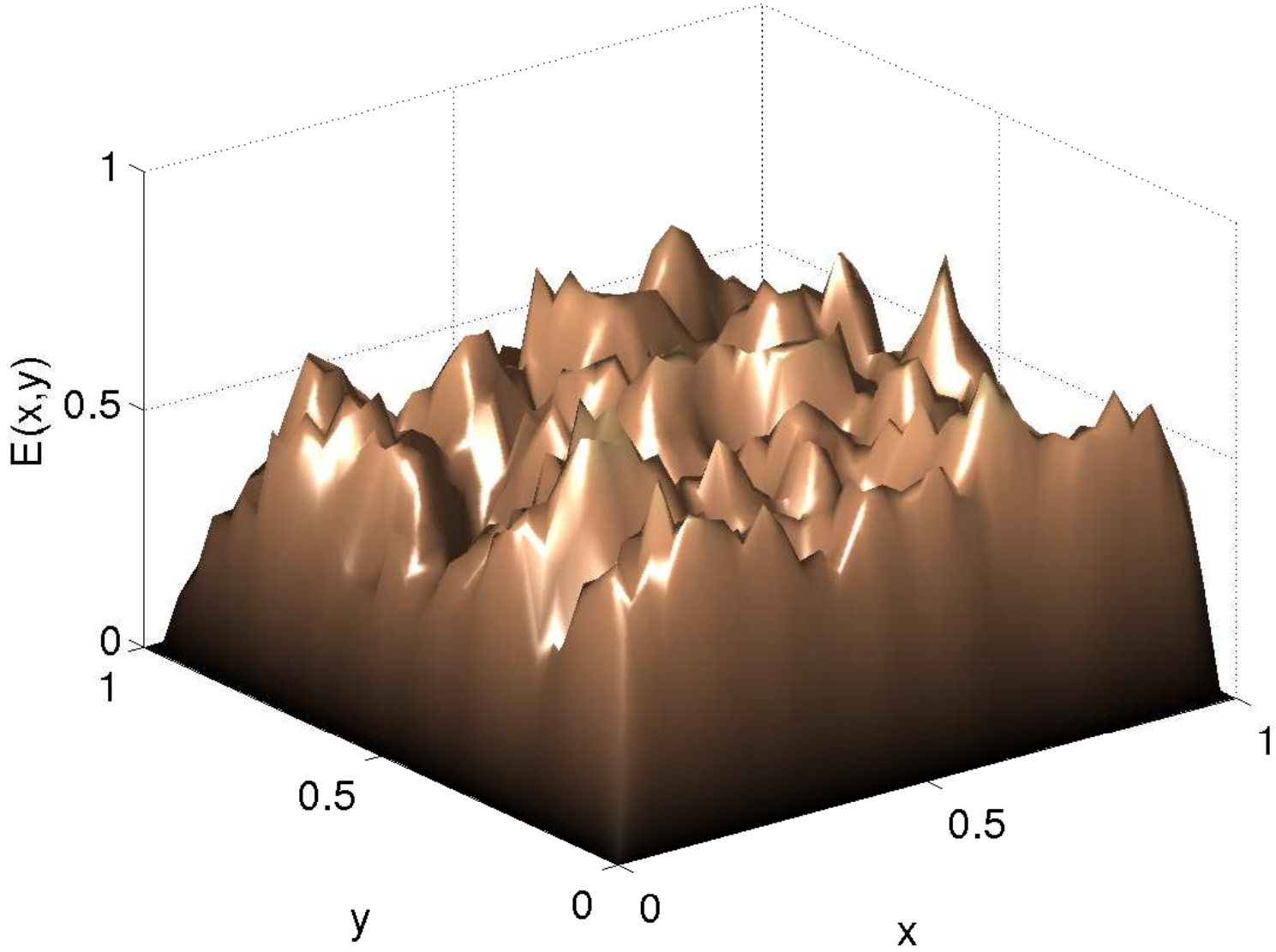
Random initial error

Smoothing Property



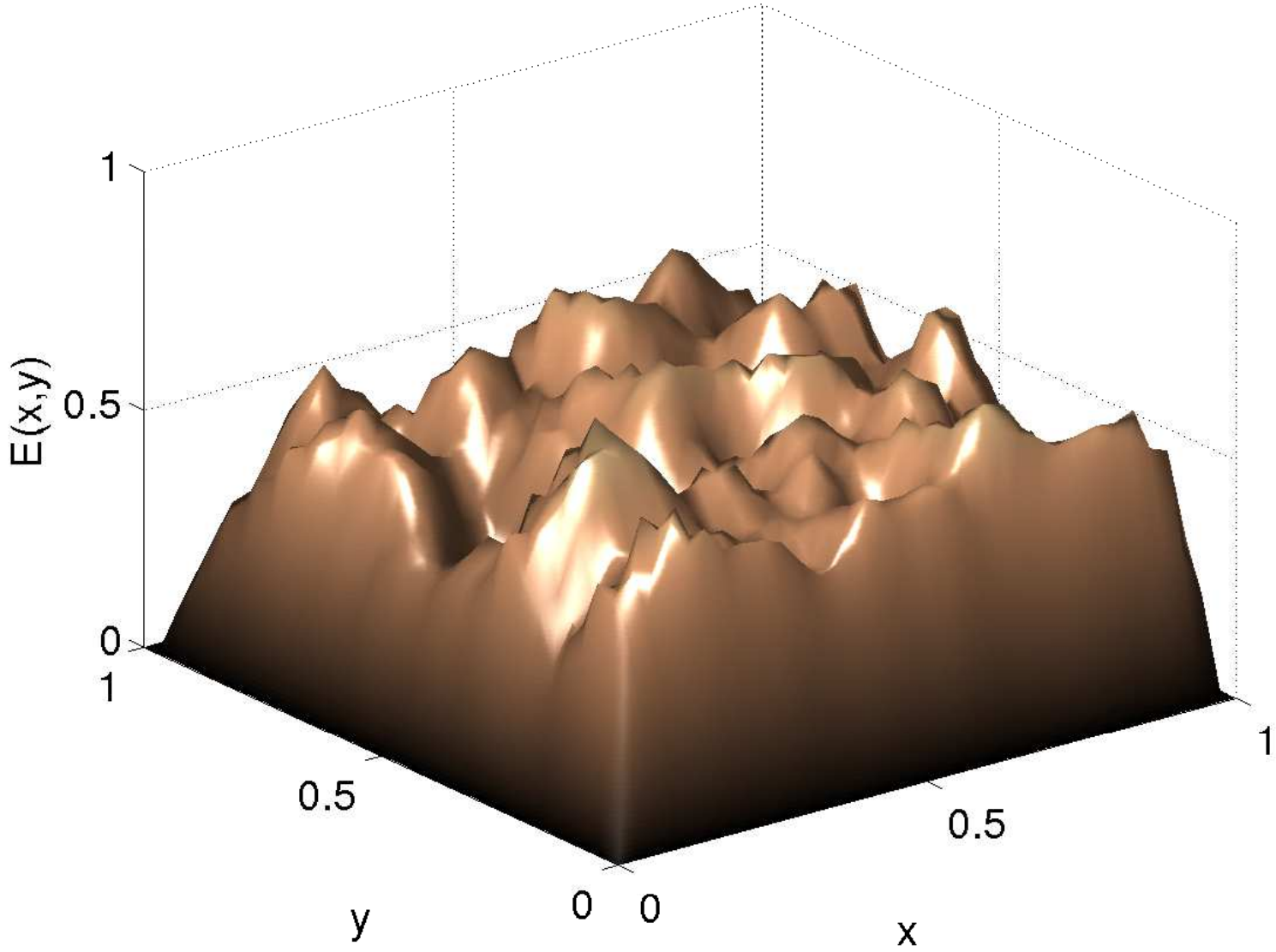
Error after 1 weighted Jacobi iteration

Smoothing Property



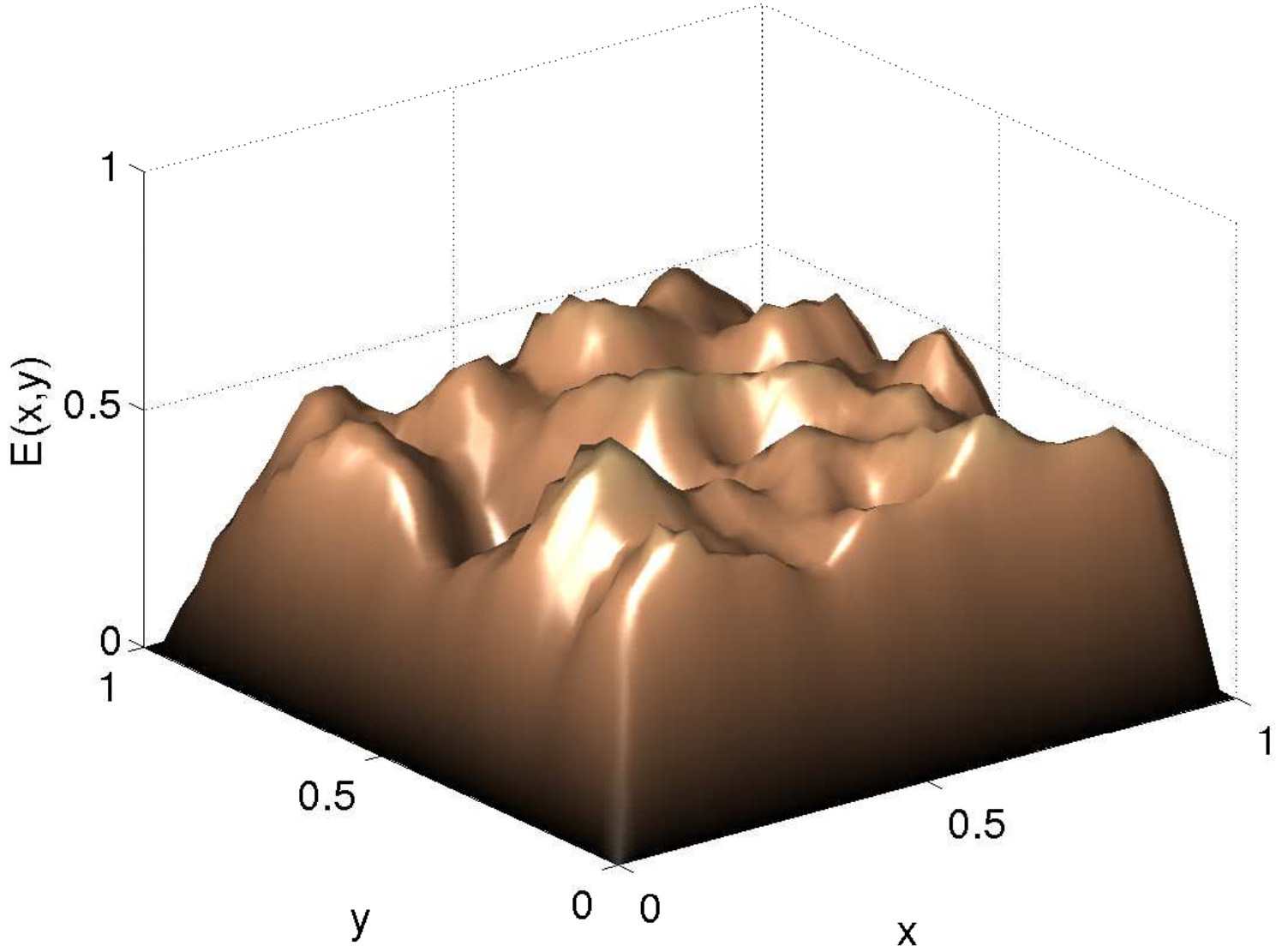
Error after 2 weighted Jacobi iterations

Smoothing Property



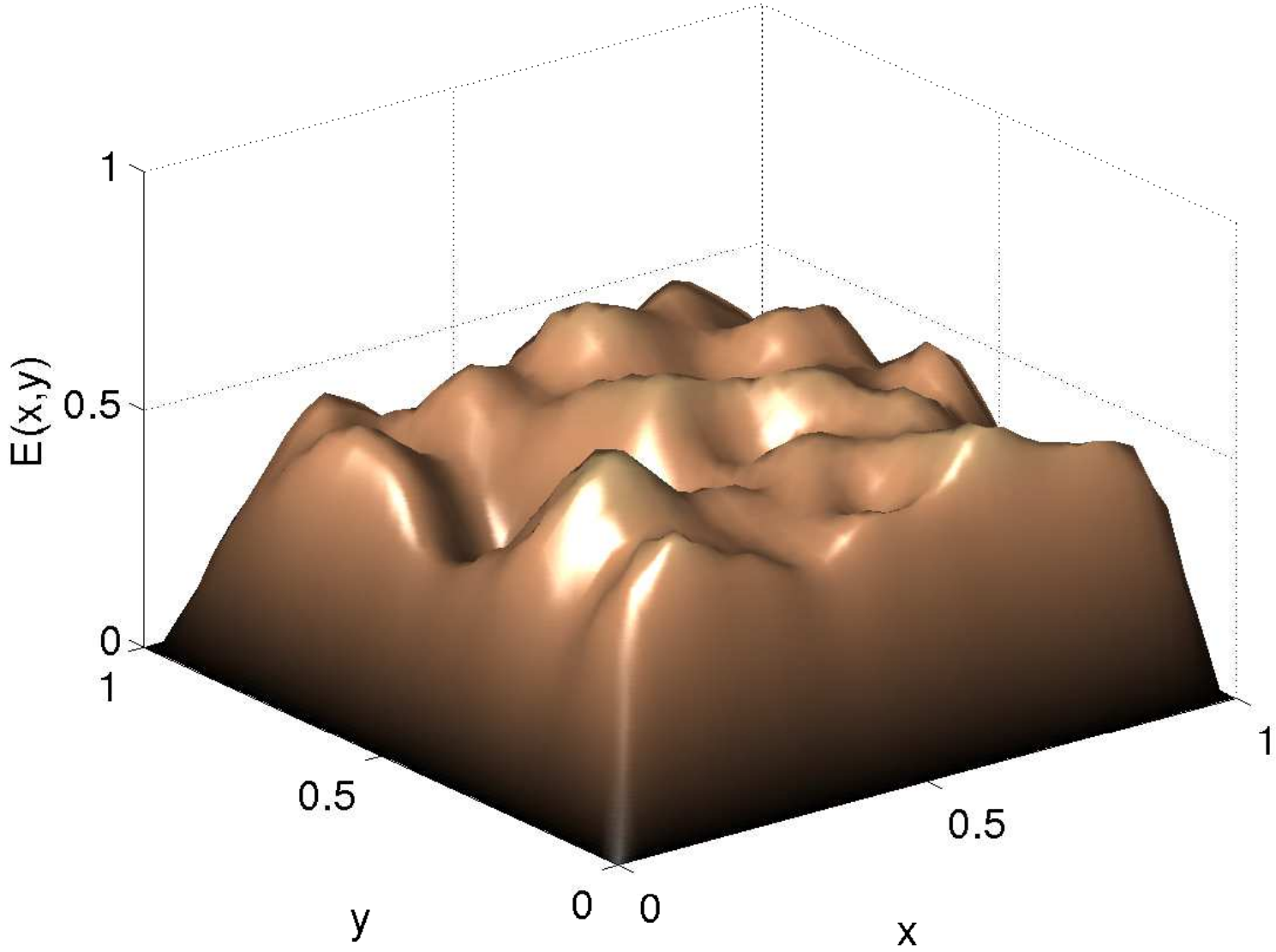
Error after 3 weighted Jacobi iterations

Smoothing Property



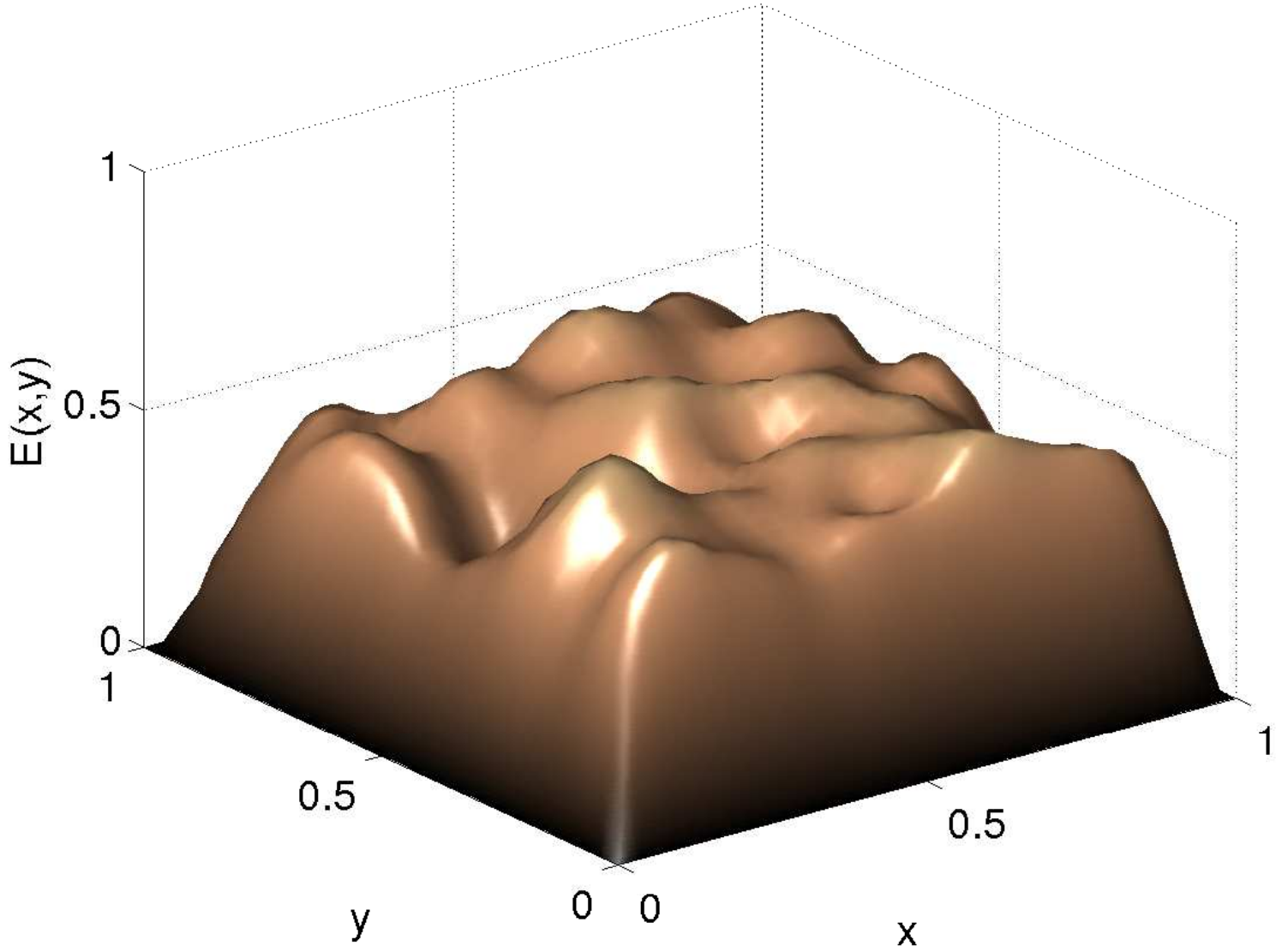
Error after 4 weighted Jacobi iterations

Smoothing Property



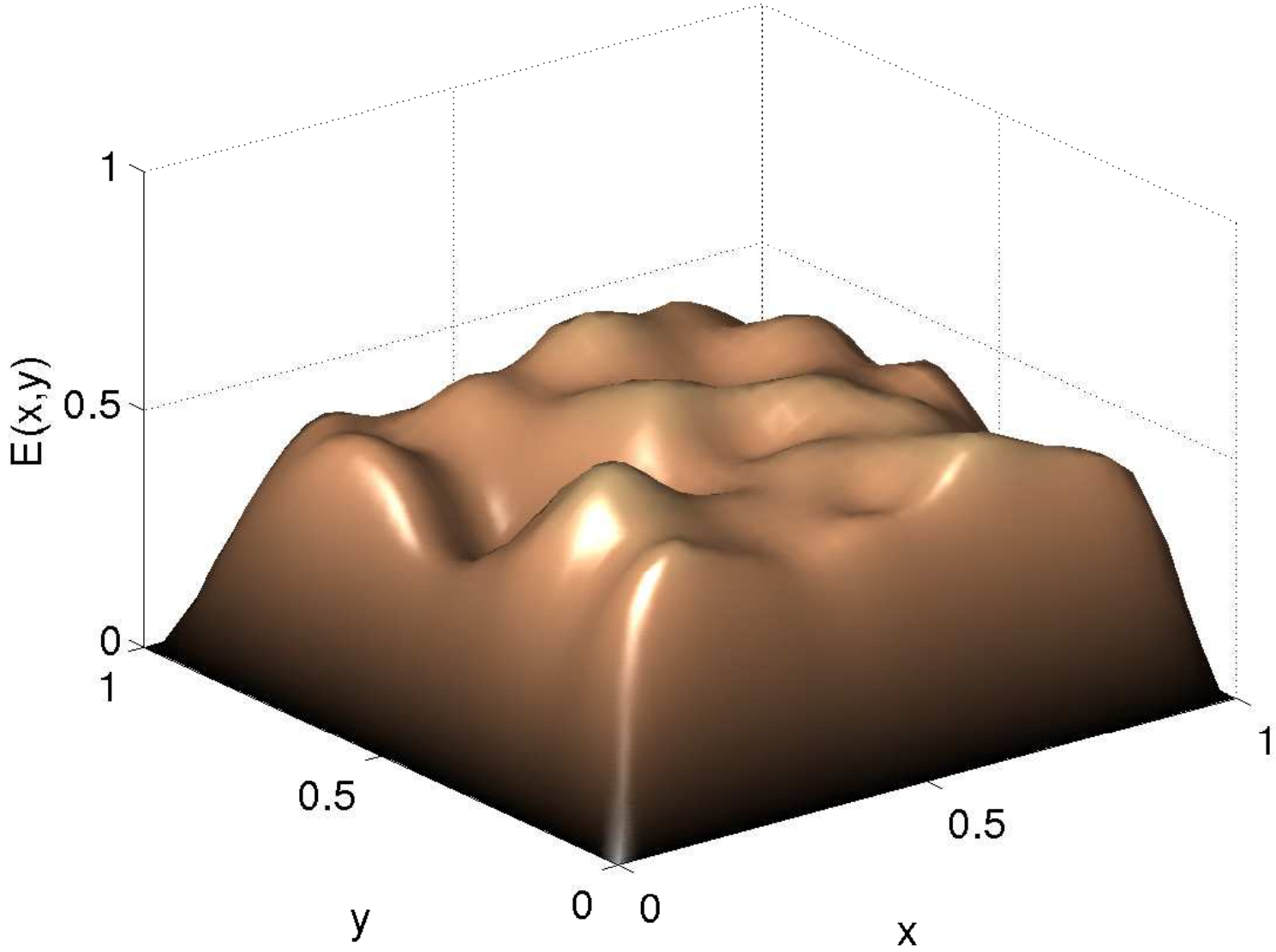
Error after 5 weighted Jacobi iterations

Smoothing Property



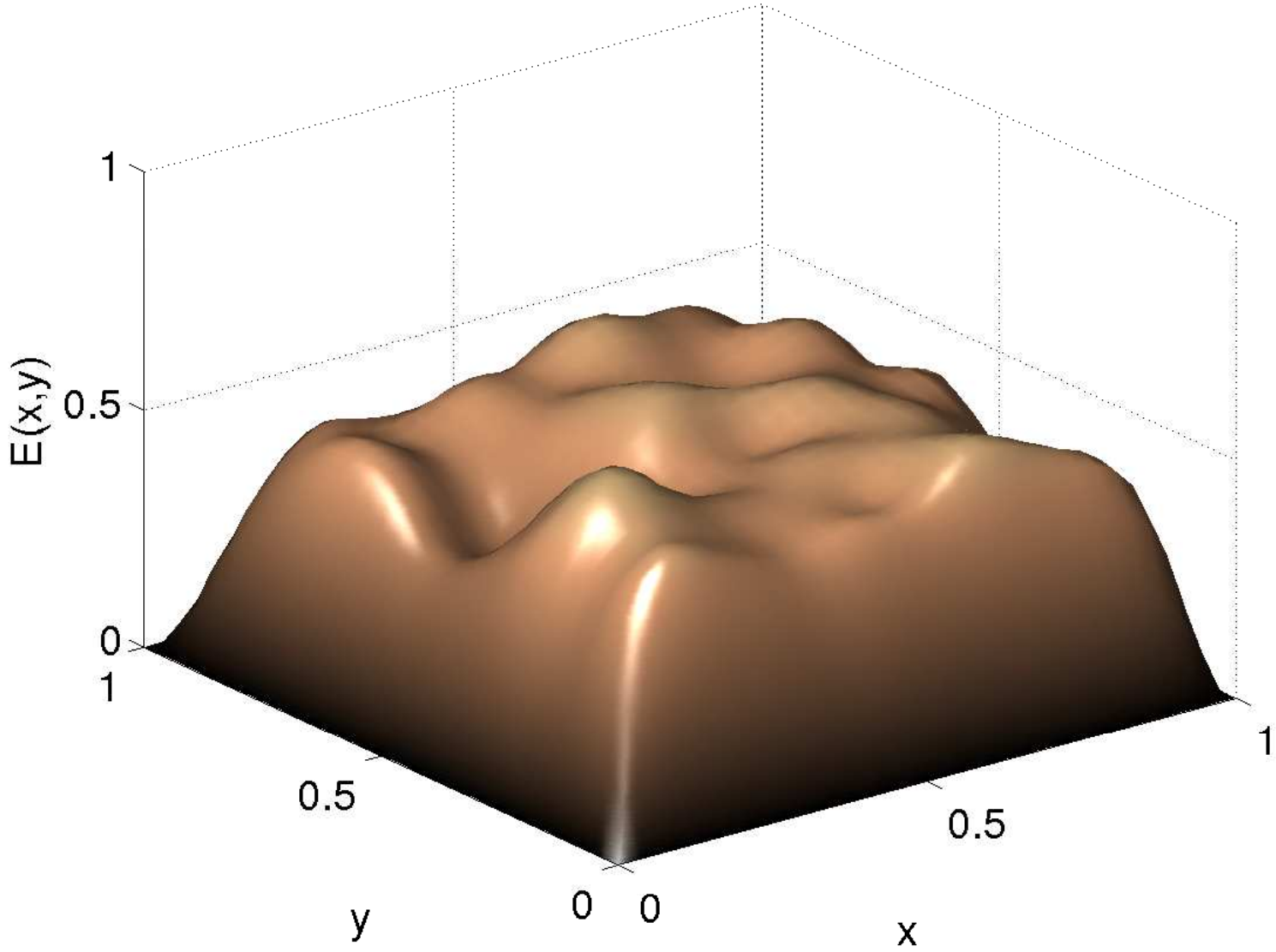
Error after 6 weighted Jacobi iterations

Smoothing Property



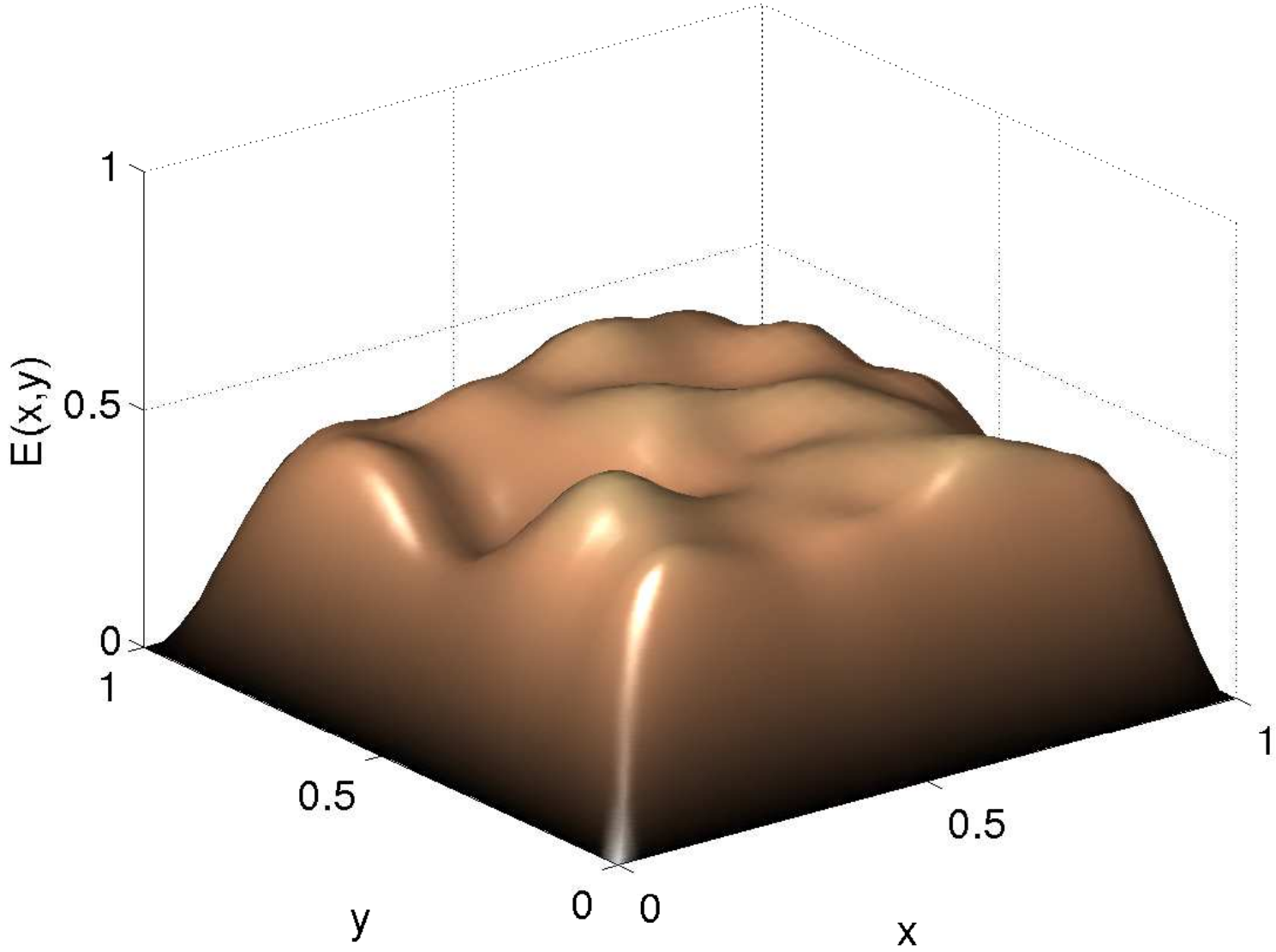
Error after 7 weighted Jacobi iterations

Smoothing Property



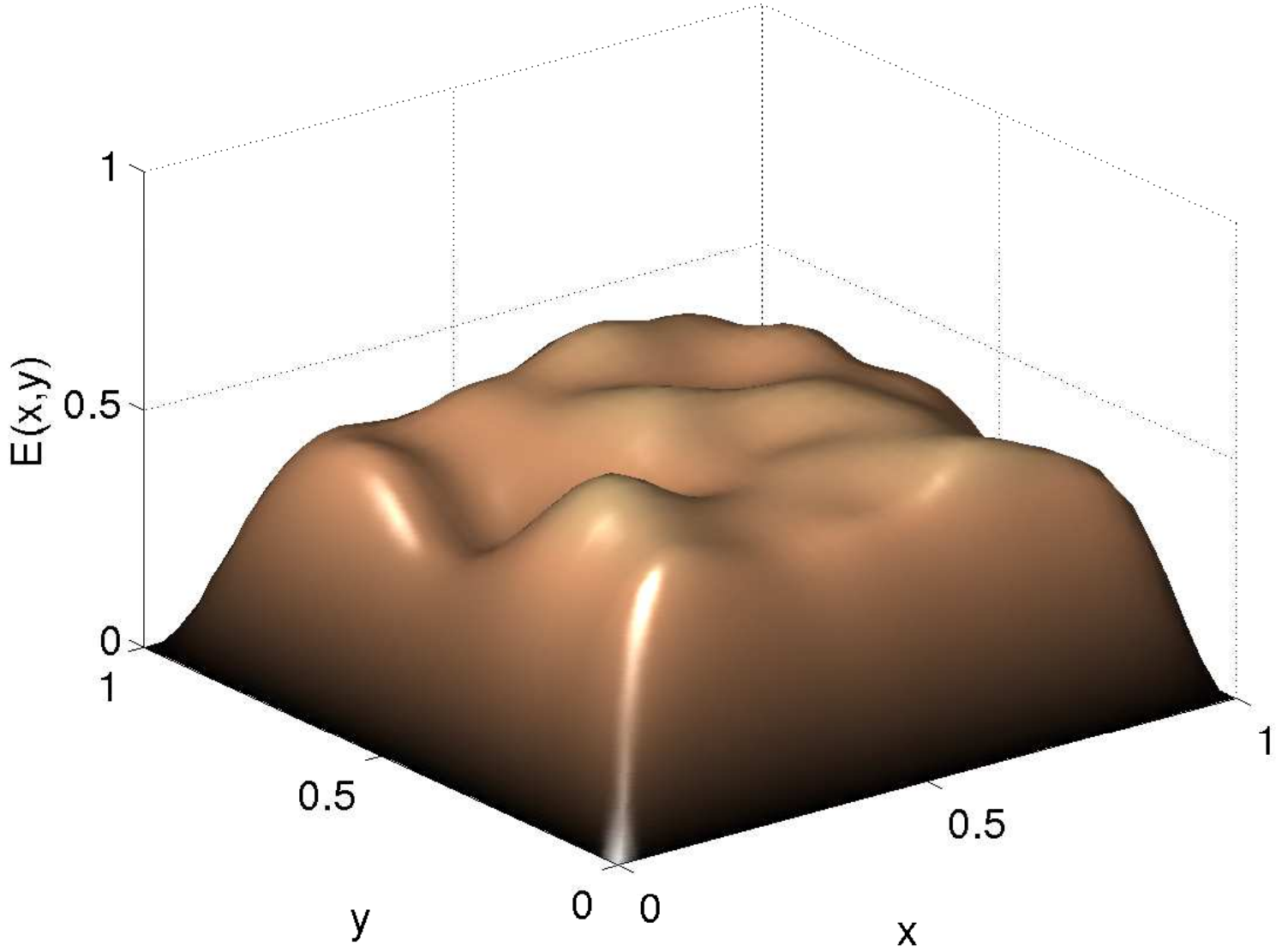
Error after 8 weighted Jacobi iterations

Smoothing Property



Error after 9 weighted Jacobi iterations

Smoothing Property



Error after 10 weighted Jacobi iterations

Complementarity

- Error after a few weighted Jacobi iterations has structure
- Instead of throwing out the method, look to complement its failings

How can we best correct error modes that are slow to be reduced by relaxation?

Complementarity

- Error after a few weighted Jacobi iterations has structure
- Instead of throwing out the method, look to complement its failings

How can we best correct error modes that are slow to be reduced by relaxation?

- Slow-to-converge errors for Poisson are smooth
- Smooth vectors can be easily represented using fewer degrees of freedom

Coarse-Grid Correction

- Smooth vectors can be accurately represented using fewer degrees of freedom
- Idea: transfer job of resolving smooth components to a coarser grid version of the problem

Need:

- Complementary process for resolving smooth components of the error on the coarse grid
- Way to combine the results of the two processes

Variational Coarsening

- Correct the approximation after relaxation, $x^{(1)}$, from an auxiliary (coarse-grid) problem
- Need interpolation map, P, from coarse grid to fine grid
- Corrected approximation will be $x^{(2)} = x^{(1)} + Px_c$

What is the *best* x_c for correction?

A-norm and A-inner product

- Asking for the *best* solution implies a metric
- Symmetric and positive-definite matrix, A, defines an inner product and a norm:

$$\langle x, y \rangle_A = y^T A x \quad \text{and} \quad \|x\|_A^2 = x^T A x$$

- *Best* then means closest to the exact solution in norm

$$y^\star = \underset{y}{\text{argmin}} \, \|x - y\|_A$$

Variational Coarsening

- Want to correct the approximation after relaxation, $x^{(1)}$, from a coarse-grid version of of the problem
- Need interpolation map, P, from coarse grid to fine grid
- Corrected approximation will be $x^{(2)} = x^{(1)} + Px_c$

What is the *best* x_c for correction?

- *Best* means closest to the exact solution in norm

$$x_c = \underset{y_c}{\operatorname{argmin}} \|x - (x^{(1)} + Py_c)\|_A$$

- *Best* x_c satisfies $(P^T AP)x_c = P^T A(x - x^{(1)}) = P^T r^{(1)}$

Multigrid

Multigrid Components

Relax: $x^{(1)} = x^{(0)} + D^{-1} r^{(0)}$

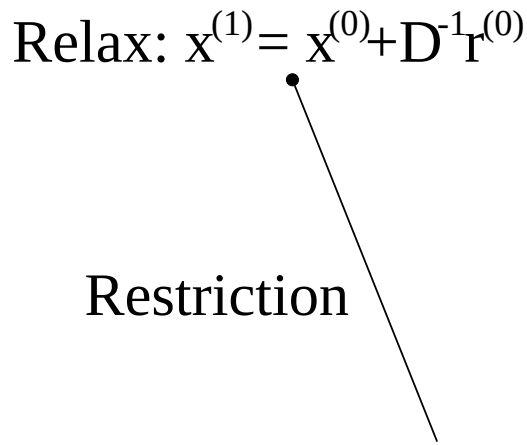
- Relaxation

- Use a smoothing process (such as Jacobi or Gauss-Seidel) to eliminate oscillatory errors
- Remaining error satisfies $Ae^{(1)} = r^{(1)} = b - Ax^{(1)}$

Multigrid

Multigrid Components

- Relaxation
- Restriction

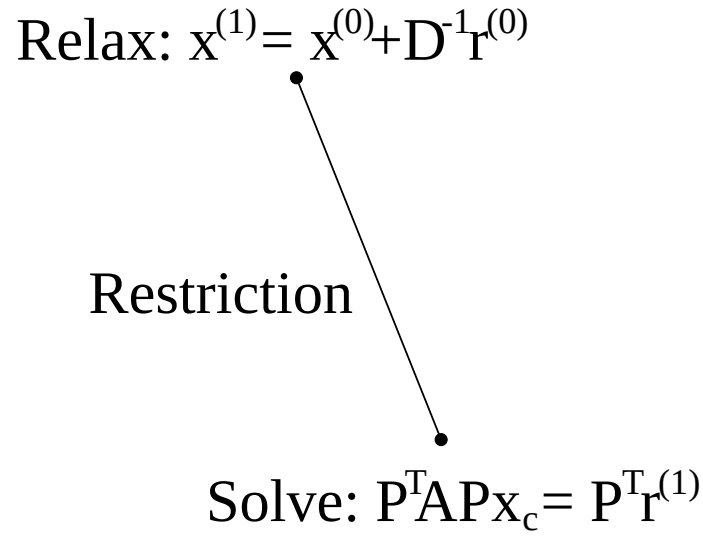


- Transfer residual to coarse grid
- Compute $P^T r^{(1)}$

Multigrid

Multigrid Components

- Relaxation
- Restriction
- Coarse-Grid Correction



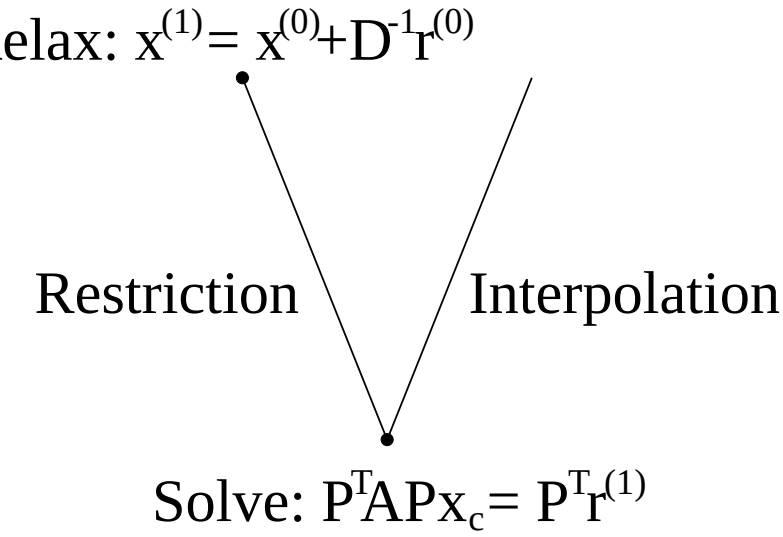
- Use coarse-grid correction to eliminate smooth errors
- Best correction, x_c, in terms of A-norm satisfies

$$P^T A P x_c = P^T r^{(1)}$$

Multigrid

Multigrid Components

- Relaxation
- Restriction
- Coarse-Grid Correction
- Interpolation

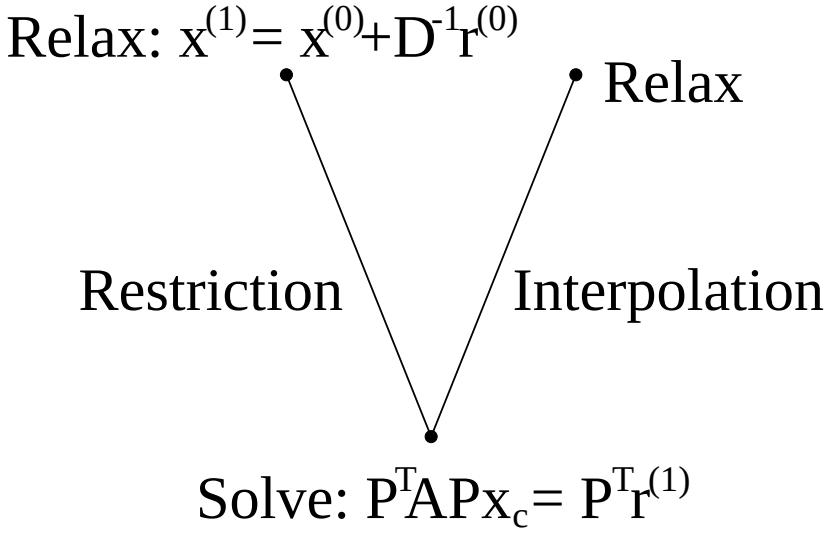


- Transfer correction to fine grid
- Compute $x^{(2)} = x^{(1)} + Px_c$

Multigrid

Multigrid Components

- Relaxation
- Restriction
- Coarse-Grid Correction
- Interpolation
- Relaxation

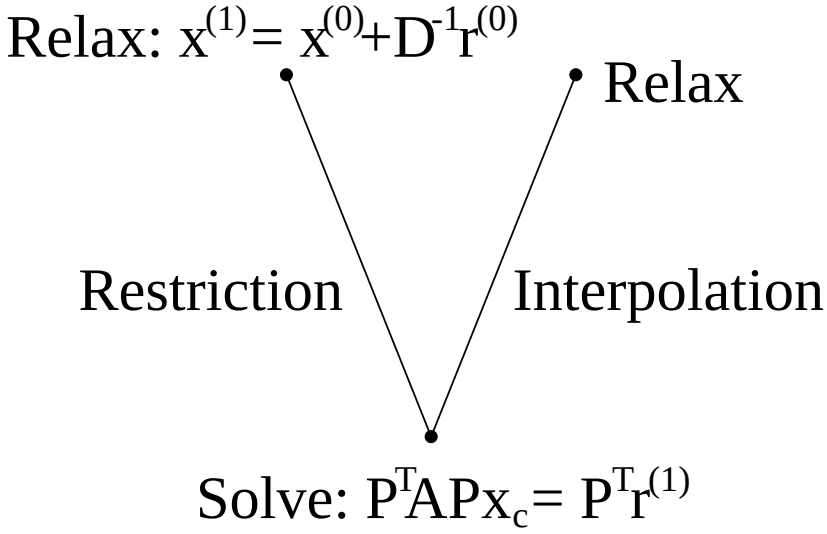


- Relax once again to remove oscillatory error introduced in coarse-grid correction

Multigrid

Multigrid Components

- Relaxation
- Restriction
- Coarse-Grid Correction
- Interpolation
- Relaxation



Direct solution of coarse-grid problem isn't practical

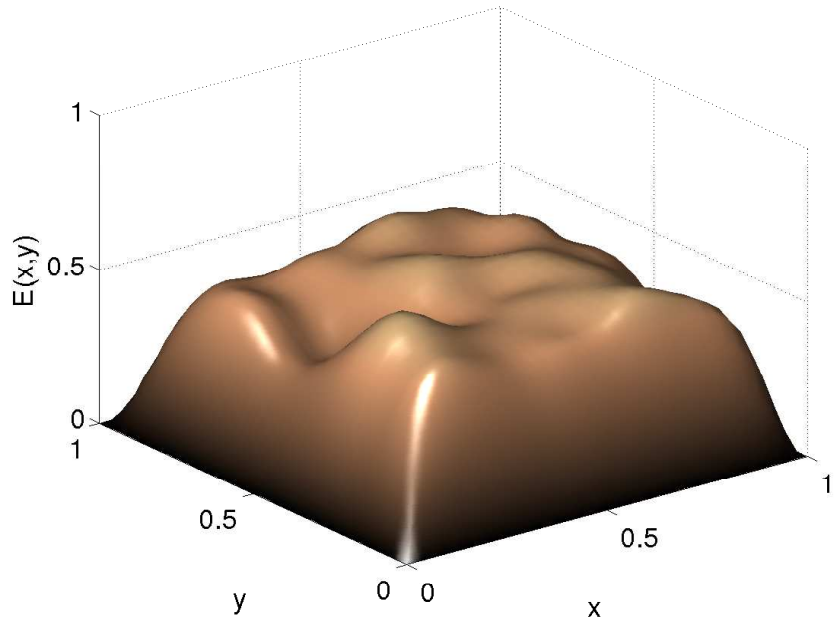
Recursion!

Apply same methodology to solve coarse-grid problem

Geometric Multigrid

For homogeneous operators, relaxation is predictable

- Jacobi/Gauss-Seidel relaxation
- Regular coarsening
- Linear interpolation



Fully explained by local mode (Fourier) analysis

Algebraic Picture

On any level, for any A, error reduced by

1. Relaxation
2. Coarse-grid correction

Coarse-grid correction treats errors in Range(P)

- Range(P) must include errors for which relaxation is slow
- Relaxation must be effective on Range(P)$^{\perp}$

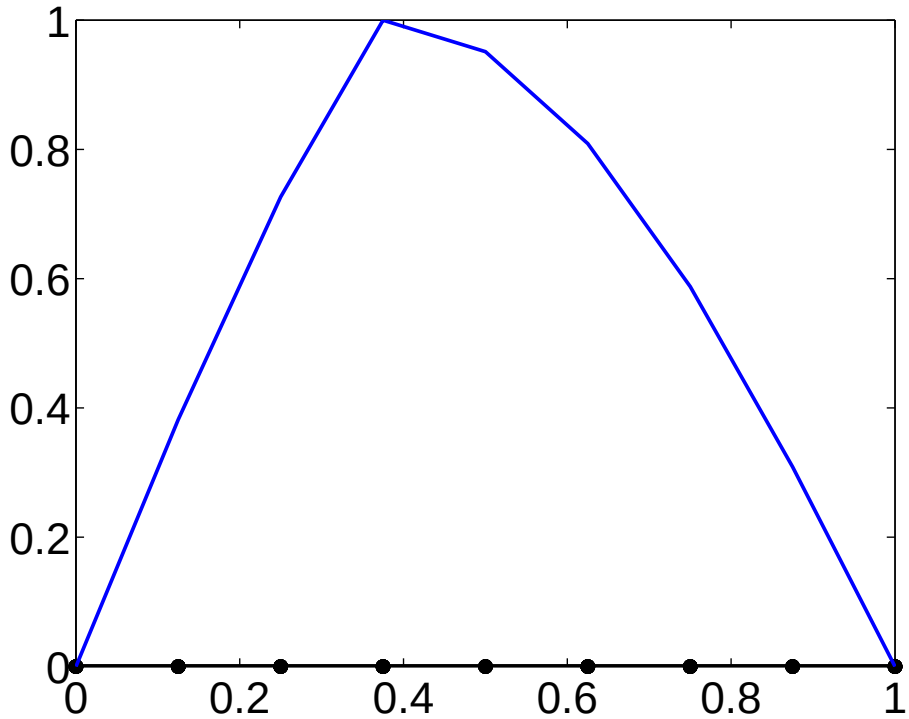
$$\text{Domain}(A) = \text{Range}(P) \oplus \text{Range}(P)^{\perp}$$

"Smooth" Errors

Heterogeneity strongly influences performance of relaxation

Slowest to converge error for $\frac{d}{dx}\left(\sigma\frac{du}{dx}\right)$, for

$$\sigma = \begin{cases} 10^{-8} & x \leq \frac{3}{8} \\ 1 & x > \frac{3}{8} \end{cases}$$



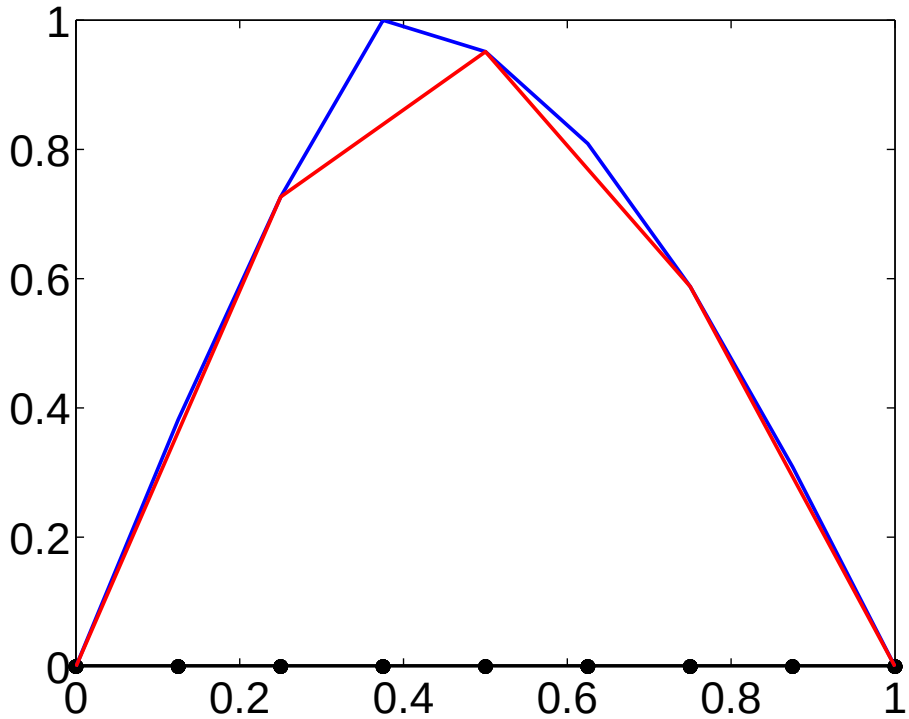
"Smooth" Errors

Heterogeneity strongly influences performance of relaxation

Slowest to converge error for $\frac{d}{dx}\left(\sigma\frac{du}{dx}\right)$, for

$$\sigma = \begin{cases} 10^{-8} & x \leq \frac{3}{8} \\ 1 & x > \frac{3}{8} \end{cases}$$

and linear interpolant from coarse grid



"Smooth" Errors

Heterogeneity strongly influences performance of relaxation

- The abrupt change in character of slow-to-converge errors is reflected in matrix entries

$$A = \frac{1}{h^2}\begin{bmatrix} 2\times 10^{-8} & -10^{-8} & & & & & & \\ -10^{-8} & 2\times 10^{-8} & -10^{-8} & & & & & \\ & -10^{-8} & 1+10^{-8} & -1 & & & & \\ & & -1 & 2 & -1 & & & \\ & & & -1 & 2 & -1 & & \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & \end{bmatrix}$$

"Smooth" Errors

Heterogeneity strongly influences performance of relaxation

- The abrupt change in character of slow-to-converge errors is reflected in matrix entries
- Idea: Use the entries in the matrix operator to help define interpolation

Algebraic Multigrid Interpolation

- Assume a partition into fine (F) and coarse (C) grid sets
- Define interpolation based only on entries in A
- Start with assumption that errors left after relaxation have small residuals: for $i \in F$,

$$(Ae)_i \approx 0$$

$$a_{ii} e_i = -\sum_{j \in F} a_{ij} e_j - \sum_{k \in C} a_{ik} e_k$$

- Use assumptions about slow-to-converge error to collapse connections to $j \in F$ onto $k \in C \cap \{k : a_{ik} \neq 0\}$

A. Brandt, S. McCormick, J. Ruge, in *Sparsity and Its Applications*, 1984
J. Ruge and K. Stüben, in *Multigrid Methods*, 1987

Calibrating Interpolation

What if we don't know what to assume about slow-to-converge errors?

A. Brandt and D. Ron, in *Multilevel Optimization in VLSICAD*, 2003

M. Brezina et al., SISC 2004, **25**:1896-1920

Calibrating Interpolation

What if we don't know what to assume about slow-to-converge errors?

Run relaxation to find out!

- Run relaxation on $Ax = 0$ with a random initial guess
- This exposes the local character of slow-to-converge errors
- Use resulting vector as a prototype of errors to be corrected by interpolation within algebraic multigrid

---

A. Brandt and D. Ron, in *Multilevel Optimization in VLSICAD*, 2003

M. Brezina et al., SISC 2004, **25**:1896-1920

Adaptive Multigrid

Automatic probing of relaxation and algebraic coarsening

- Given matrix A, Relaxation operation $B^{-1}r$
- Iterate on homogeneous problem, $Ax = 0$, with a random initial guess
- Create AMG-style interpolation such that prototype of slow-to-converge error is in its range
- Create coarse-grid problem and recurse

Adaptive Multigrid

Automatic probing of relaxation and algebraic coarsening

- Given matrix A, Relaxation operation $B^{-1}r$
- Iterate on homogeneous problem, $Ax = 0$, with a random initial guess
- Create AMG-style interpolation such that prototype of slow-to-converge error is in its range
- Create coarse-grid problem and recurse

Relaxation can be anything

Adaptive Multigrid

Automatic probing of relaxation and algebraic coarsening

- Given matrix A, Relaxation operation $B^{-1}r$
- Iterate on homogeneous problem, $Ax = 0$, with a random initial guess
- Create AMG-style interpolation such that prototype of slow-to-converge error is in its range
- Create coarse-grid problem and recurse

Relaxation can be anything,
even the multigrid method itself!

- Allows for iterative improvement of a poorly performing multigrid cycle

Linear Elasticity

- Model displacement, u, of an elastic body under external forces

$$-\mu\Delta u - (\lambda + \mu)\nabla\nabla \cdot u = f$$

- μ, λ are Lamé coefficients, defined as

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \text{and} \quad \mu = \frac{E}{2(1+\nu)}$$

- Fix Poisson ratio, $\nu = 0.32$ (steel)
- Let Young modulus, E, vary between 1 (nylon/polypro) and 10^{σ} (100 = titanium, 1000 = diamond)
- Know properties of slow-to-converge errors for small σ

Numerical Results: Linear Elasticity

3D cube, 201,720 DOFs, exponential distribution of E

	Standard SA			Adaptive SA		
σ	ρ_{MG}	Itns	CPU (s)	ρ_{MG}	Itns	CPU (s)
2	0.115	9	26.0	0.214	12	267.7
3	0.247	14	35.7	0.310	16	275.6
4	0.395	20	50.0	0.404	21	289.4
5	0.556	32	73.6	0.497	27	381.2

M. Brezina et al., SISC 2004, **25**:1896-1920

Lattice Quantum Chromodynamics

- Modelling interactions between fermions on a lattice
- Goal: Solve $H(u, \rho)f = b$, for multiple source vectors, b, at each step of a Monte Carlo simulation
- Difficulty: u is a complex unitary field defined on lattice edges, phases chosen randomly based on parameter, β
- H is Hermitian, but indefinite, so solve normal equations
- As ρ approaches a critical value, H^*H becomes singular (for any β)
- Structure of low-energy modes strongly depends on u
  - When $\beta \to \infty$, $u \to 1$, H^*H looks like a second-order discrete differential operator
  - For each state, new characterization of low-energy modes

Numerical Results: Lattice QCD

128×128 periodic lattice
average residual reduction per iteration

	Diagonal-PCG				AdaptiveMG-PCG			
$\rho - \rho_{\text{cr}}$	0.3	0.1	0.05	0.01	0.3	0.1	0.05	0.01
$\beta = 2$	0.85	0.94	0.96	0.99	0.31	0.31	0.31	0.33
$\beta = 3$	0.86	0.93	0.97	0.98	0.31	0.40	0.42	0.42
$\beta = 5$	0.83	0.92	0.96	0.99	0.28	0.29	0.31	0.31

Adaptive MG setup time: 13.7 seconds
Adaptive MG-PCG solve time: 0.8 seconds
Diagonal-PCG solve time: 4.7 seconds

J. Brannick et al., to appear in *Proc. DD16*, 2007

Choosing Coarse Grids

- Difficult to say what best coarse grid is
  - Enough coarse-grid points so that interpolation is accurate for all slow-to-converge errors
  - Significant reduction in number of grid points
- Interpolation chosen to complement failings of relaxation
- Galerkin coarse-grid operator must stay manageable

Coarse-grid selection must make this possible

Maximal Independent Sets

- Want local interpolation operators
- Idea: Coarsen so that every fine-grid node has at least one coarse-grid neighbor

Problem: not all connections are equal

$$-\epsilon u_{xx} - u_{yy} \rightarrow \begin{bmatrix} & -\frac{1}{h_y^2} & \\ -\frac{\epsilon}{h_x^2} & \left(\frac{2\epsilon}{h_x^2} + \frac{2}{h_y^2}\right) & -\frac{\epsilon}{h_x^2} \\ & -\frac{1}{h_y^2} & \end{bmatrix}$$

- Really want every fine-grid node to be somehow strongly coupled to at least one coarse-grid node

AMG coarsening

- Classical AMG defines strong connections for each node based directly on the matrix entries:

$$S_i = \left\{ j : -a_{ij} \geq \theta \max_{k \neq i} \{-a_{ik}\} \right\}$$

- Coarse grid chosen as maximal independent set over strong connections

AMG coarsening

- Classical AMG defines strong connections for each node based directly on the matrix entries:

$$S_i = \left\{ j : -a_{ij} \geq \theta \max_{k \neq i} \{-a_{ik}\} \right\}$$

- Coarse grid chosen as maximal independent set over strong connections
- Strong connections based on nice M-matrix properties
- Break down if near null space of A is oscillatory
  - Diagonal rescaling, $A \rightarrow DAD$
  - Finite element anisotropy

Compatible Relaxation

Alternative: choose coarse grids so that we know that interpolation can be chosen to complement relaxation

Principle of Compatible Relaxation:

If relaxation on fine-grid submatrix fast to converge

Then a good interpolation operator exists

Good means that resulting multigrid method has small convergence factor

---

A. Brandt, ETNA 2000, **10**:1-20

Convergence Theory

- Fine-grid relaxation: $I - \omega B_{ff}^{-1} A_{ff}$ is efficient if

$$c_1 \langle B_{ff} x_f, x_f \rangle \leq \langle A_{ff} x_f, x_f \rangle \leq c_2 \langle B_{ff} x_f, x_f \rangle$$

  for reasonable ω, c_1, c_2

- Under right assumptions, can show multigrid convergence is bounded less than 1, with bound dependent on $\frac{c_2}{c_1}$

R. Falgout and P. Vassilevski, SISC 2004, **42**:1669-1693

S. MacLachlan, S. McCormick, & T. Manteuffel, NLAA 2006, **13**:599-620

Satisfying the Theory

Want to turn these results into a practical algorithm

Idea: Choose partition so that we know fine-grid relaxation converges quickly

Know weighted Jacobi relaxation on A_{ff} converges quickly when A_{ff} is diagonally dominant

We can guarantee good 2-level convergence factors by choosing A_{ff} to be diagonally-dominant

NP-completeness

- Define θ-dominance of A_{ff} as

$$a_{ii} \geq \theta \sum_{j \in F} |a_{ij}|$$

- Theory can be satisfied as long as A_{ff} is θ-dominant

Want A_{ff} to be the largest submatrix of A that is θ-dominant

NP-completeness

- Define θ-dominance of A_{ff} as

$$a_{ii} \geq \theta \sum_{j \in F} |a_{ij}|$$

- Theory can be satisfied as long as A_{ff} is θ-dominant

Want A_{ff} to be the largest submatrix of A that is θ-dominant
This is an NP-complete problem

Greedy Algorithm Approach

Want an $O(n)$ (not NP) coarse-grid selection algorithm

- Initialize all points to be in U; F, C to be empty
- For each point i, compute diagonal dominance measure

$$\hat{\theta}_i = \frac{|a_{ii}|}{\sum_{j \in F \cup U} |a_{ij}|}$$

  If $\hat{\theta}_i \geq \theta$, put i into F, remove it from U
- While U is non-empty
  - Find $j = \underset{i \in U}{\operatorname{argmin}} \hat{\theta}_i$
  - Remove j from U, put it in C
  - For each neighboring point i of j, update $\hat{\theta}_i$
    If $\hat{\theta}_i \geq \theta$, put i into F, remove it from U

Test Problems

Test problems based on finite element discretizations of
$-\nabla \cdot K(x,y)\nabla p(x,y)$

- Laplace equation, $K(x,y) = 1$
- Smooth coefficient, $K(x,y) = 10^{-8} + 10(x^2 + y^2)$
- Randomly chosen coefficient, $K(x,y) = 10^{-8}$ on 20% of the cells, chosen randomly, $K(x,y) = 1$ otherwise
- Anisotropic coefficient, $K(x,y) = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0.01 \end{smallmatrix}\right]$

Algorithm:

- Greedy algorithm to select coarse grids
- Classical AMG to define interpolation
- Usual AMG-V(1,1) cycles with Gauss-Seidel relaxation

Multilevel AMG results

Coefficient	Grid	c_A	t_{setup}	t_{solve}	# iters.	ρ
$K(x,y)=1$	512^2	1.33	1.3	0.7	5	0.13
	1024^2	1.33	5.1	2.5	5	0.14
	2048^2	1.33	21.9	10.5	5	0.14
smooth	512^2	1.33	1.3	0.6	5	0.13
	1024^2	1.33	5.1	2.5	5	0.14
	2048^2	1.33	21.7	10.4	5	0.14
random	512^2	2.06	2.3	1.2	6	0.35
	1024^2	2.08	9.6	4.8	6	0.40
	2048^2	2.10	41.0	19.8	6	0.46
anisotropic	512^2	2.39	1.5	1.0	5	0.13
	1024^2	2.41	6.2	4.1	5	0.20
	2048^2	2.43	25.8	17.7	5	0.20

QCD Revisited

Brannick et al. was a brute-force approach

- Complex system converted to equivalent real form
- Expensive adaptive smoothed aggregation technique

Want an AMG algorithm that naturally handles

- Complex-valued operators
- Hermitian, complex-symmetric, non-symmetric operators
- Strong heterogeneity

Complex Helmholtz

Time-harmonic simplification of Maxwell's equations yields

$$-\Delta u + \imath\omega u = f$$

Differentially dominant terms occur in real part

- Use systems approach
  - Base coarse grids, interpolation on dominant term
- Apply AMG coarsening to real part ($-\Delta u$)
- Galerkin product based on complex matrix

D. Lahaye et al., IEEE Trans. Magn. 2000, **36**:1535-1538
S. Reitzinger et al., J. Comp. App. Math., **155**:405-421

Gauge Laplacian

Simplified model from lattice QCD gives

$$u_{i,j} - \kappa \left(e^{\imath\beta\theta_{i,j}} u_{i-1,j} + e^{\imath\beta\phi_{i,j}} u_{i,j-1} + e^{\imath\beta\theta_{i+1,j}} u_{i+1,j} + e^{\imath\beta\phi_{i,j+1}} u_{i,j+1} \right) = f_{ij},$$

where

- $\kappa \approx 0.25$
- $\beta \geq 0$
- $\{\theta_{ij}\}$, $\{\phi_{ij}\}$ chosen from given probability distribution

Real part no longer dominates

Need AMG based on full complex operator

Acoustic Wave Equation

Fourier transform in time yields

$$-\nabla \cdot \frac{1}{\rho(x)} \nabla \hat{p}(x) - \frac{\omega^2}{K} \hat{p}(x) = f(x)$$

For constant density, precondition with geometric multigrid for

$$-\nabla \cdot \nabla \hat{p}(x) - \frac{\omega^2}{c^2}(1 + \imath\alpha)\hat{p}(x) = f(x)$$

New challenges:

- Multigrid for variable-density problems
- AMG for indefinite matrices

Erlangga, Oosterlee, Vuik, SISC 2006, **27**:1471-1492

Practical Theory for AMG

Most AMG theory is neither computable nor sharp

Typically,

- Many conditions or idealized algorithm
- Only an upper bound is given
- Upper bound depends on solution of eigenvalue problem

Practical theory would give

- Bound dependent on easily computed properties
- Bound achieved for some data & initial guess
- Insight into choices to be made within AMG

Summary

- Heterogeneity & uncertainty add new complications to linear solvers
- Algebraic picture of multigrid gives insight
- Adaptive framework replaces assumptions on relaxation
- Added expense can be recovered for some applications

Summary

- Heterogeneity & uncertainty add new complications to linear solvers
- Algebraic picture of multigrid gives insight
- Adaptive framework replaces assumptions on relaxation
- Added expense can be recovered for some applications
- New coarsening approaches show promise
- Principles of compatible relaxation offer alternative to heuristics

Summary

- Heterogeneity & uncertainty add new complications to linear solvers
- Algebraic picture of multigrid gives insight
- Adaptive framework replaces assumptions on relaxation
- Added expense can be recovered for some applications
- New coarsening approaches show promise
- Principles of compatible relaxation offer alternative to heuristics
- New applications leading to new challenges
- Better theoretical tools leading to better computational tools