Aspects of Multigrid

Kees Oosterlee [1,2]

[1]Delft University of Technology, Delft.
[2]CWI, Center for Mathematics and Computer Science, Amsterdam,

SIAM Chapter Workshop Day, May 30th 2018

CWI Centrum Wiskunde & Informatica

Agenda

- Towards PDE problems with random coefficients
- Introduction to Multigrid
  - Poisson Equation
  - Smoothing and Coarse Grid Correction
  - Problems with Jumping Coefficients
- Cell-Centered Multigrid

  Joint work with Francisco Gaspar, Prashant Kumar

CWI Centrum Wiskunde & Informatica

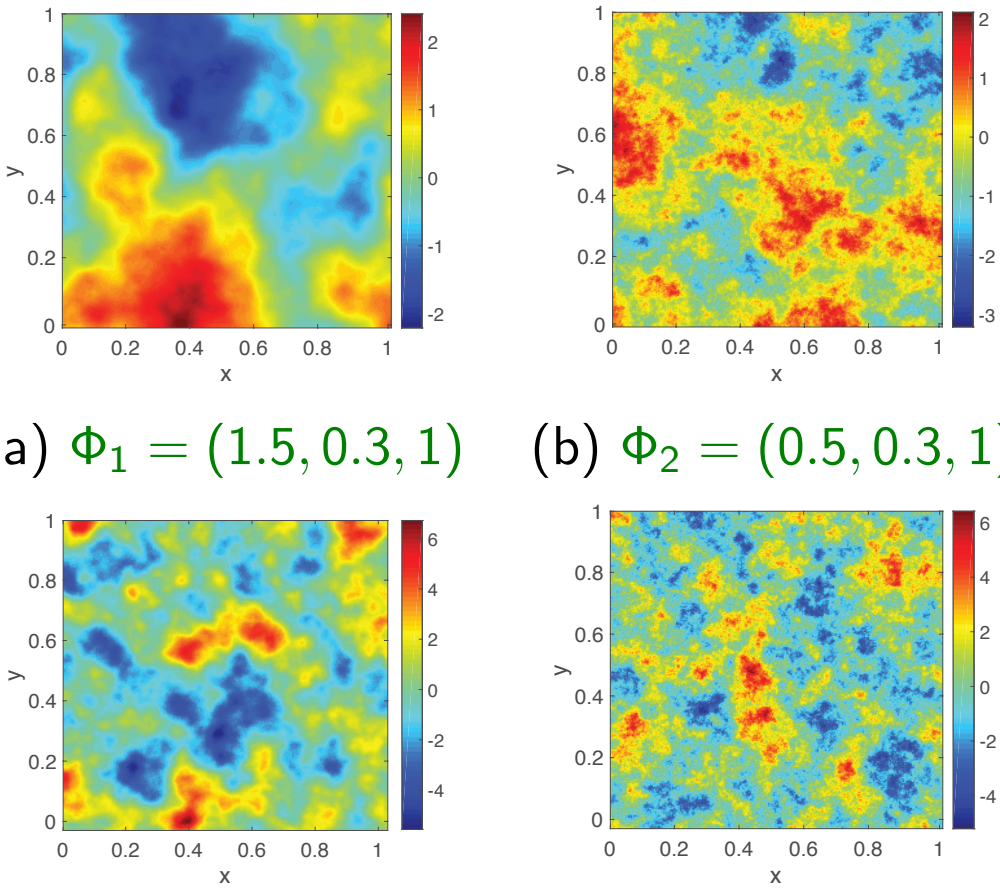
Problem Setting

- The elliptic PDE of our interest describes a steady-state single-phase flow in a random porous medium.
- Let ω be an event in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with sample space Ω, σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$, the permeability in the porous medium is described by $k(\mathbf{x}, \omega) : \overline{\mathcal{D}} \times \Omega \to \mathbb{R}$. The PDE is given by

$$-\nabla \cdot (k(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) = f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{D}, \omega \in \Omega, \tag{1}$$

  with f a source term.
- The engineering interest is typically found in expected values of linear functionals of solution u, denoted by $Q := Q(u)$.

Random Coefficient fields



(a) $\Phi_1 = (1.5, 0.3, 1)$ (b) $\Phi_2 = (0.5, 0.3, 1)$

(c) $\Phi_3 = (1.5, 0.1, 3)$ (d) $\Phi_4 = (0.5, 0.1, 3)$

Figure: Logarithm of the permeability field, $\log K$, generated using four reference parameter sets $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$.

Varying Coefficients

- When using Monte Carlo methods, in the case of elliptic PDEs with random coefficient fields, many samples of the random field are generated and for each field the numerical solution should be computed.
- For a fixed sample of the random field, we deal with an elliptic PDE with varying coefficients, due to the randomness.
- Robust and efficient iterative solution methods are very relevant for PDEs with variable coefficients.

$\Rightarrow$ Multigrid comes in naturally as a highly efficient solution method for the resulting PDEs with jumping coefficients.

CWI Centrum Wiskunde & Informatica

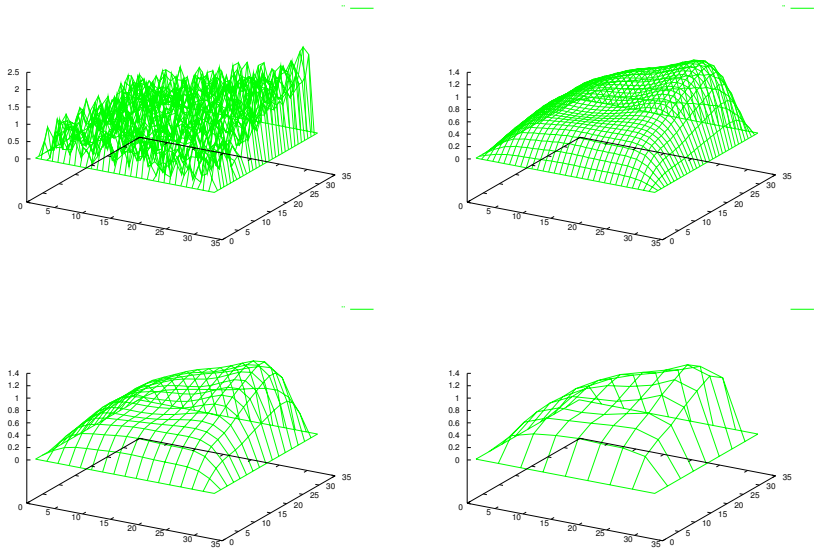
Efficient solution methods for discrete equations

- Multigrid is the classical hierarchical approach for nicely elliptic equations
- Poisson Equation, in 2D or 3D, $-\Delta u = f$.
- $O(N)$ complexity requires hierarchical approaches.
  - Approximation on different scales
  - Compute only those "parts" of a solution on fine scales that really require the fine resolution.
  - Compute/correct the remaining parts on coarser scales
  - Gives $O(N)$ complexity
- Advantages increase with increasing problem size
- The idea behind multigrid is very general

CWI Centrum Wiskunde & Informatica

Multigrid since 1973 (Brandt, Hackbusch, . . .)

- Iterative methods like Jacobi and Gauss-Seidel converge slowly on fine grids, however, they smooth the error $u_h - \overline{u}$
- Smooth errors can be approximated well on coarser grids (with much less grid points)
- Multigrid is a $\mathcal{O}(N)$- method !



CWI Centrum Wiskunde & Informatica

Smoothing effect

- Look at the 2D Fourier expansion of the error:

$$e_h(x,y) = \sum_{k,l=1}^{p-1} \alpha_{k,l} \sin k\pi x \sin l\pi y = \sum_{k,l=1}^{p-1} \alpha_{k,l} \varphi^{k,l}$$

- Gauss-Seidel iterative methods smooth the error in a numerical solution!
- The fact that this error becomes smooth means that the high frequency components, i.e.,

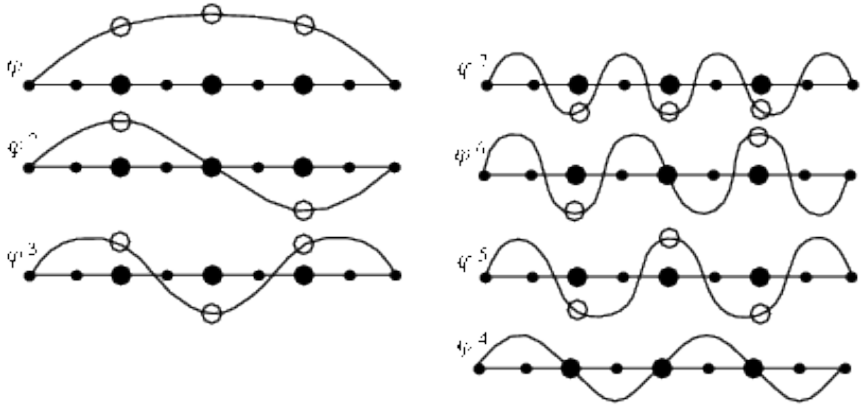
$$\alpha_{k,l} \sin k\pi x \sin l\pi y \quad \text{with } k \text{ or } l \text{ large}$$

become small after a few iterations, whereas the low frequency components

$$\alpha_{k,l} \sin k\pi x \sin l\pi y \quad \text{with } k \text{ and } l \text{ small}$$

hardly change.

CWI Centrum Wiskunde & Informatica



low frequency components, visible also on Ω_{2h}

high frequency components, not visible on Ω_{2h}

$\Rightarrow$ High frequency components cannot be corrected on a coarse grid !

- Coarse grid correction makes sense, if low frequencies dominate the error.

$$\sum_{k,l=1}^{p-1} \alpha_{k,l}\varphi^{k,l} = \sum^{high} \alpha_{k,l}\varphi^{k,l} + \sum^{low} \alpha_{k,l}\varphi^{k,l}$$

where

$$\sum^{low} \alpha_{k,l}\varphi^{k,l} = \sum_{k,l=1}^{p/2-1} \alpha_{k,l}\varphi^{k,l}$$

and

$$\sum^{high} \alpha_{k,l}\varphi^{k,l} = \sum_{\substack{k,l \\ p/2 \leq \max(k,l)}}^{p-1} \alpha_{k,l}\varphi^{k,l} \quad .$$

CWI Centrum Wiskunde & Informatica

Defect equation

- For any approximation u_h^m of the solution u_h, we denote the error by $e_h^m := u_h - u_h^m$, and the defect (or residual) by

$$\boxed{r_h^m := b_h - A_h u_h^m}$$

- The defect equation

$$A_h e_h^m = r_h^m$$

is equivalent to the original equation, since

$$u_h = u_h^m + e_h^m.$$

- This leads to the procedure

$$u_h^m \;\rightarrow\; r_h^m = b_h - A_h u_h^m \rightarrow A_h e_h^m = r_h^m \rightarrow u_h = u_h^m + e_h^m \;\; .$$

- This procedure is not a meaningful numerical process.

CWI Centrum Wiskunde & Informatica

Iteration matrix

- However, if A_h is approximated by a "simpler" operator $\widehat{A}_h$ such that $\widehat{A}_h^{-1}$ exists, $\widehat{e}_h^i$ in $\widehat{A}_h \widehat{e}_h^m = r_h^m$ gives a new approximation

$$u_h^{m+1} := u_h^m + \widehat{e}_h^m.$$

- The procedural formulation then looks like

$$\boxed{u_h^m \Rightarrow r_h^m = b_h - A_h u_h^m \rightarrow \widehat{A}_h \widehat{e}_h^m = r_h^m \Rightarrow u_h^{m+1} = u_h^m + \widehat{e}_h^m} \quad .$$

Coarse grid correction

- Use an appropriate approximation A_H of A_h on a coarser grid Ω_H, for instance the grid with mesh size $H = 2h$. The defect equation is replaced by

$$A_H \widehat{e}_H^m = r_H^m.$$

  $A_H : \mathcal{G}(\Omega_H) \rightarrow \mathcal{G}(\Omega_H), \quad \dim \mathcal{G}(\Omega_H) < \dim \mathcal{G}(\Omega_h)$ and ${A_H}^{-1}$ exists.

- As r_H^m and $\widehat{e}_H^m$ are grid functions on Ω_H, we need two transfer operators

$$I_h^H : \mathcal{G}(\Omega_h) \rightarrow \mathcal{G}(\Omega_H), \quad I_H^h : \mathcal{G}(\Omega_H) \rightarrow \mathcal{G}(\Omega_h)$$

- I_h^H is used to restrict r_h^m to Ω_H:

$$\boxed{r_H^m := I_h^H r_h^m} \ ,$$

  and I_H^h is used to interpolate (or prolongate) the correction $\widehat{e}_H^m$ back to Ω_h:

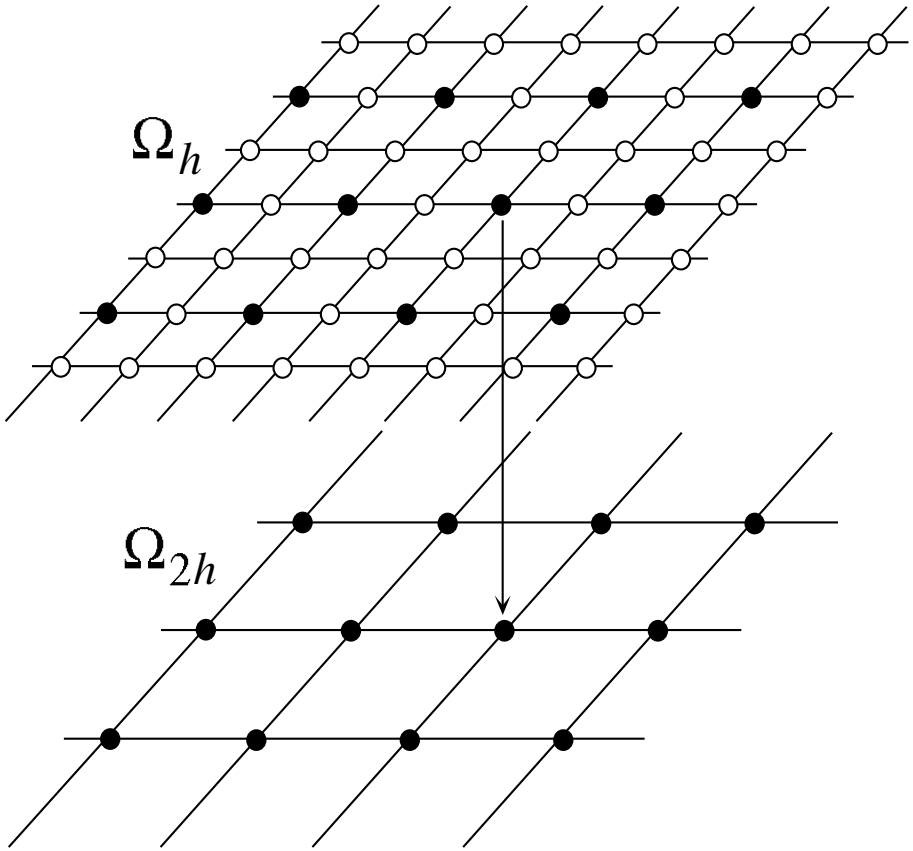
$$\boxed{\widehat{e}_h^m := I_H^h \widehat{e}_H^m} \ .$$

CWI Centrum Wiskunde & Informatica

- The simplest example for a restriction operator is the "injection" operator

$$r_H(P) = I_h^H r_h(P) := r_h(P) \qquad \text{for} \quad P \in \Omega_H \subset \Omega_h \ ,$$

- A fine and a coarse grid with the injection operator:



The two grid iteration: Correction scheme

$\Rightarrow$ Combine the two processes of smoothing and of coarse grid correction.

- Consider a linear problem $A_h u_h = b_h$ on grid G_h

(1)	ν_1 smoothing steps on the fine grid:	$\widehat{u}_h \Longleftarrow S^{\nu_1}(u_h^0, b_h)$
(2)	computation of residuals on the fine grid:	$r_h := b_h - A_h \widehat{u}_h$
(3)	restriction of residuals from fine to coarse:	$r_H := I_h^H r_h$
(4)	solution of the coarse grid problem:	$A_H e_H = r_H$
(5)	prolongation of corrections from coarse to fine:	$e_h := I_H^h e_H$
(6)	adding the corrections to the current fine grid approximation:	$\widehat{u}_h \Longleftarrow \widehat{u}_h + e_h$
(7)	on the fine grid:	$u_h^1 \Longleftarrow S^{\nu_2}(\widehat{u}_h, b_h)$

- Steps (1) and (7) are pre– and post–smoothing, steps (2)...(6) form the "coarse grid correction cycle".

CWI Centrum Wiskunde & Informatica

- One obtains the iteration operator Q_h^H of the (h, H) two–grid cycle:

$$\boxed{Q_h^H = {S_h}^{\nu_2} K_h^H {S_h}^{\nu_1} \qquad \text{with} \qquad K_h^H := I_h - I_H^h {A_H}^{-1} I_h^H A_h} \ .$$

- The following components of the (h, H)–cycle have to be specified:
  - the smoothing procedure SMOOTH (u_h^m, A_h, f_h)
  - the numbers ν_1, ν_2 of smoothing steps,
  - the coarse grid Ω_H,
  - the fine–to–coarse restriction operator I_h^H,
  - the coarse grid operator A_H,
  - the coarse–to–fine interpolation operator I_H^h.

  $\Rightarrow$ Generalisation to multiple grids direct

Multigrid components for Poisson's equation

- Smoothing iteration: lexicographic Gauss-Seidel method
- Coarse grid discretization:

$$L_H u_H = \frac{1}{H^2} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}_H .$$

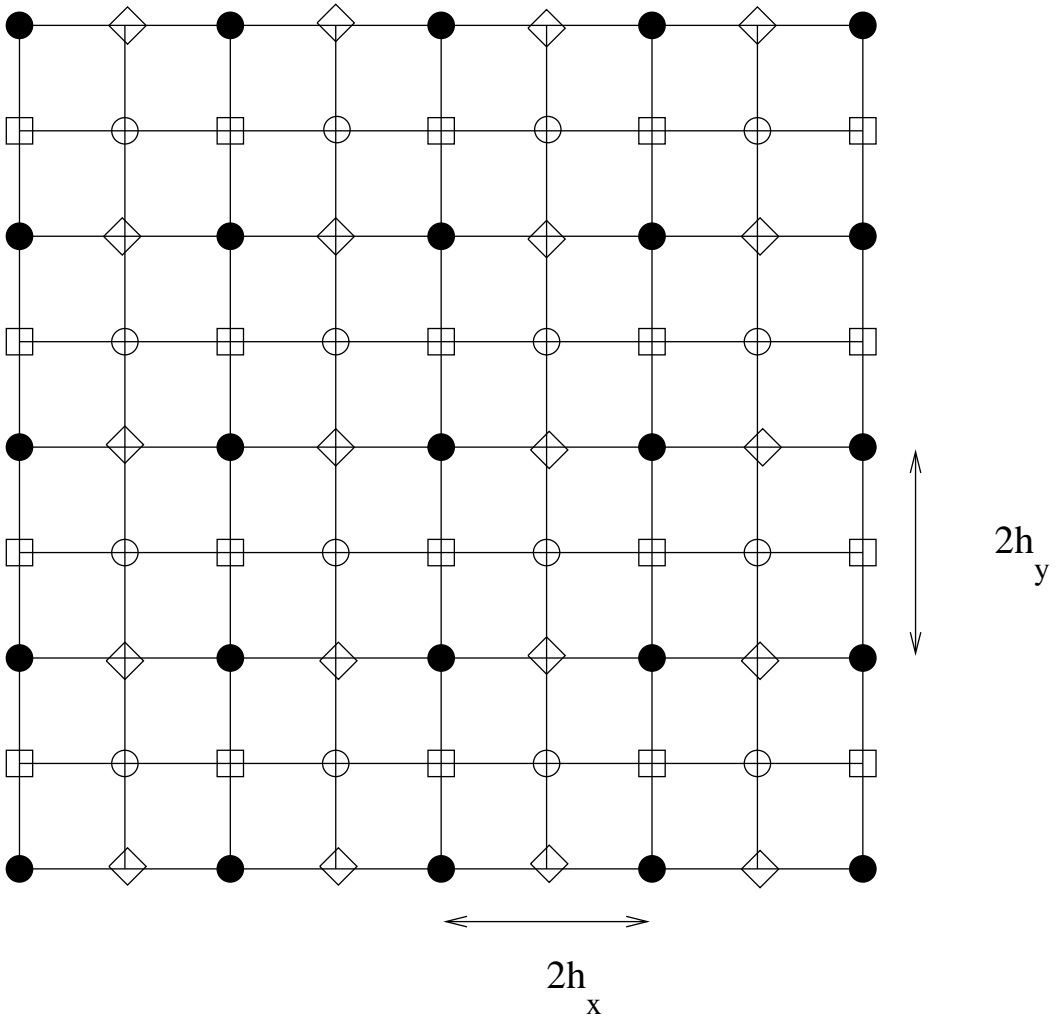
- Restriction: Next to injection, the full-weighting operator:
  Applying this operator to a grid function $r_h(x,y)$ at a coarse grid point $(x,y) \in \Omega_{2h}$ means

$$\begin{aligned} r_{2h}(x,y) &= I_h^{2h} r_h(x,y) \\ &= \frac{1}{16}\Big[4r_h(x,y) &+ 2r_h(x+h,y) + 2r_h(x-h,y) \\ &&+ 2r_h(x,y+h) + 2r_h(x,y-h) \\ &&+ r_h(x+h,y+h) + r_h(x+h,y-h) \\ &&+ r_h(x-h,y+h) + r_h(x-h,y-h)\Big] . \end{aligned}$$

CWI Centrum Wiskunde & Informatica

Multigrid components continued

- Prolongation operator



Computational work

- With these MG components, h-independent multigrid convergence for the Poisson equation: Convergence does not depend on the finest grid size.
- $(||r_h^m||/||r_h^0||)^{1/m}, \quad r_h^m = f_h - A_h u_h^m$.

Cycle	$h = 1/512$	$1/256$	$1/128$	$1/64$	$1/32$	$1/16$
V(1,1):	0.19	0.19	0.19	0.19	0.19	0.19
V(2,1):	0.12	0.12	0.12	0.12	0.12	0.12

- The total computational work W_l of one complete 2D multigrid cycle (V-cycle: $\gamma = 1$; W-cycle $\gamma = 2$):

$$W_\ell \dot{\leq} \begin{cases} \frac{4}{3} C \cdot N_\ell & (\text{for } \gamma = 1) \\ 2C \cdot N_\ell & (\text{for } \gamma = 2) \\ 4C \cdot N_\ell & (\text{for } \gamma = 3) \\ O(N_\ell \log N_\ell) & (\text{for } \gamma = 4) \end{cases} .$$

CWI Centrum Wiskunde & Informatica

An example: Anisotropic equation

- For the anisotropic equation (plus bound. conditions):

$$-\varepsilon u_{xx} - u_{yy} = b(x, y)$$

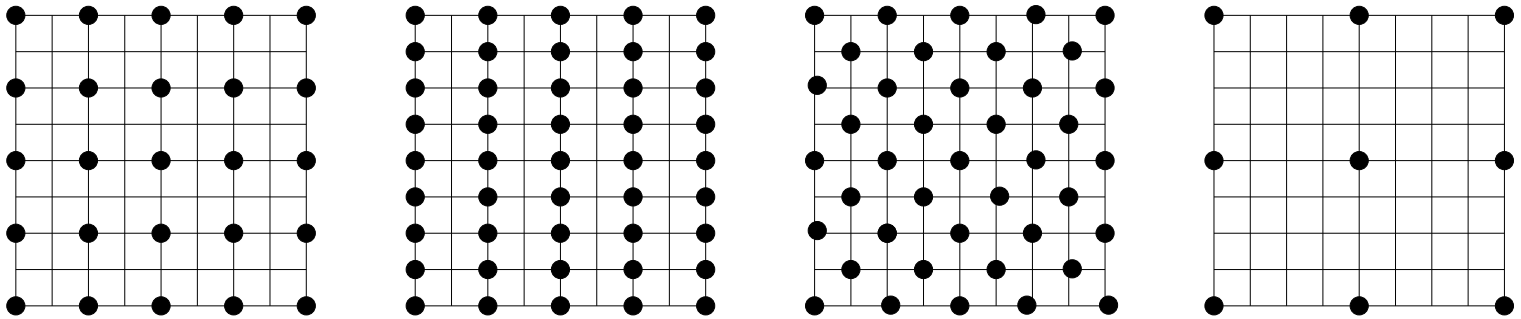
  with $0 < \varepsilon << 1$.

- Discretization stencil:

$$L_h(\varepsilon) = \frac{1}{h^2} \begin{bmatrix} & -1 & \\ -\varepsilon & 2(1+\varepsilon) & -\varepsilon \\ & -1 & \end{bmatrix}_h .$$

- Smoothness of errors after Gauss-Seidel smoothing in only one direction: Semi-coarsening in y-direction is appropriate. OR: Change smoothing so that the error will become smooth in both directions.
- Change MG components, either smoothing or coarsening!

Multigrid components: Choice of coarse grid



- The choice of grid depends on the smoothness of the error.
- For irregular finite volume/ finite element grids coarse "grids" are chosen based on the connections in the matrix. In this case, it is better to say that "coarse matrices" are constructed.
- It is possible to determine, based on matrix properties (M-matrix, for example), where the error will be smooth and accordingly how to coarsen algebraically (algebraic multigrid, AMG).

Discussion about Algebraic Multigrid

- In the context of *algebraic multigrid*, two prevailing methods have proved their use and robustness;

$\Rightarrow$ Classical AMG and aggregation-based multigrid.

- The origin of these algebraic methods is found in the early days of multigrid
- AMG $\Leftarrow$ Black-box multigrid for structured vertex-centered Cartesian grids, by Joel Dendy and co-workers.
- In classical AMG, black-box multigrid was essentially enhanced by a flexible coarsening strategy.

CWI Centrum Wiskunde & Informatica

Discussion about Algebraic Multigrid

- Aggregation-based multigrid, developed in the work by Jan Mandel (smoothed aggregation), is based on constant transfer operators and simple coarse grid discretization (aggloreation/aggregation).

$\Rightarrow$ Related to cell-centered multigrid, proposed by Piet Wesseling et al.

- Cell-centered multigrid components were augmented with robust smoothing, like Incomplete Lower-Upper decomposition (ILU) relaxation.
- We can now analyze and understand the convergence of these methods for problems with jumping, or random coefficients!

CWI Centrum Wiskunde & Informatica

Jumping Coefficients Problem

- Classical PDE problem with jumping coefficients,

$$-\nabla \cdot (k(\mathbf{x})\nabla u(\mathbf{x})) = f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{D}, \tag{2}$$

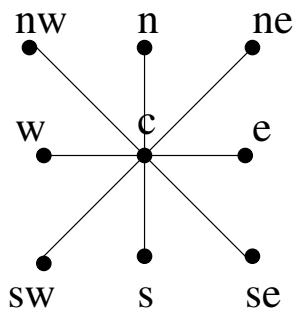
$$u(\mathbf{x}) = g(\mathbf{x}), \quad \mathbf{x} \in \partial\mathcal{D}, \tag{3}$$

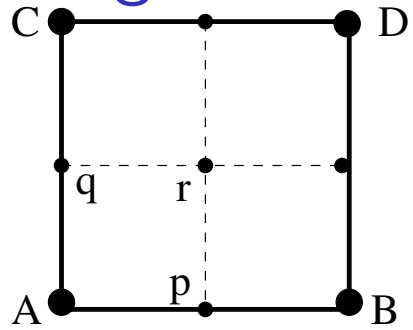
where $k(\mathbf{x})$ is a function which may be discontinuous across internal boundaries. Model for the random coefficient PDE

Black-box MG for vertex-centered discretization

- Prolongation: matrix-dependent interpolation, based on the matrix elements
- Restriction: Adjoint of prolongation
- Galerkin coarse grid matrix: $A_{2h} := I_h^{2h} A_h I_{2h}^h, \quad A_{4h} := I_{2h}^{4h} A_{2h} I_{4h}^{2h}, \quad$ etc.

Matrix-Dependent Prolongation–Jumping Coefficients





- For fine grid points p: $e_{h,p} = w_A e_{H,A} + w_B e_{H,B}$.

$$
\begin{aligned}
w_A &= \min(1, \max(0, w_w)); \; w_B = \min(1, \max(0, w_e)), \\
w_w &= \frac{d_w}{d_w + d_e}, \; w_e = \frac{d_e}{d_w + d_e} \\
d_w &= \max(|a_p^{sw} + a_p^{w} + a_p^{nw}|, |a_p^{sw}|, |a_p^{nw}|) \\
d_e &= \max(|a_p^{se} + a_p^{s} + a_p^{ne}|, |a_p^{se}|, |a_p^{ne}|),
\end{aligned}
$$

On points r: $e_h(r)$ is determined so that $A_h I_{2h}^h e_{2h} = 0$ at r,

Cell-Centered Multigrid (Delft, 1980/1990ties)

- Work in the group of prof. Wesseling in the 1980ties
- Discretize by *cell-centered finite volume method* based on the harmonic average of $k(\mathbf{x})$. Gives, for each interior cell, a five-point scheme

$$c^h_{i_1,i_2} u_{i_1,i_2} + w^h_{i_1,i_2} u_{i_1-1,i_2} + e^h_{i_1,i_2} u_{i_1+1,i_2} + s^h_{i_1,i_2} u_{i_1,i_2-1} + n^h_{i_1,i_2} u_{i_1,i_2+1} = f^h_{i_1,i_2}, \quad (4)$$

where

$$\begin{aligned}
w^h_{i_1,i_2} &= -\frac{2}{h^2}\frac{k_{i_1,i_2}k_{i_1-1,i_2}}{k_{i_1,i_2}+k_{i_1-1,i_2}}, \quad e^h_{i_1,i_2} = -\frac{2}{h^2}\frac{k_{i_1,i_2}k_{i_1+1,i_2}}{k_{i_1,i_2}+k_{i_1+1,i_2}}, \\
s^h_{i_1,i_2} &= -\frac{2}{h^2}\frac{k_{i_1,i_2}k_{i_1,i_2-1}}{k_{i_1,i_2}+k_{i_1,i_2-1}}, \quad n^h_{i_1,i_2} = -\frac{2}{h^2}\frac{k_{i_1,i_2}k_{i_1,i_2+1}}{k_{i_1,i_2}+k_{i_1,i_2+1}}, \\
c^h_{i_1,i_2} &= -(w^h_{i_1,i_2} + e^h_{i_1,i_2} + n^h_{i_1,i_2} + s^h_{i_1,i_2}),
\end{aligned}$$

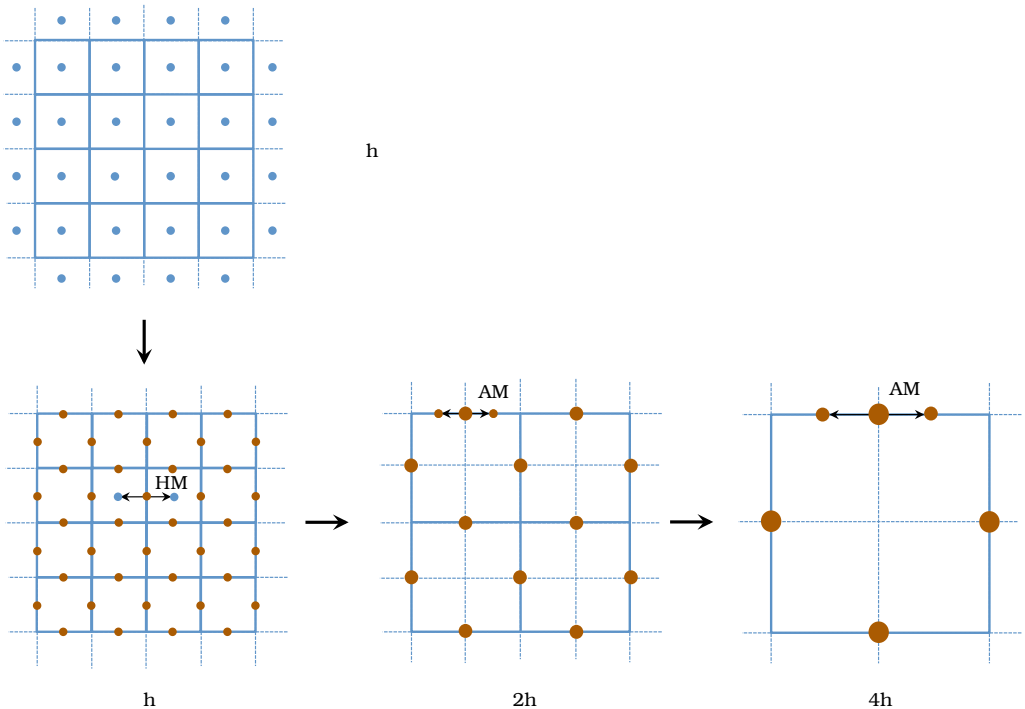
CWI Centrum Wiskunde & Informatica

Jumping Coefficients - Cell-Centered Multigrid

- Standard coarsening
- Since the unknowns are located in the cell-centers, this results in a non-nested hierarchy of grids.
- Simple prolongation I_{2h}^{h}: the piecewise constant interpolation.
- Restriction operator I_{h}^{2h} is the adjoint of the prolongation
- The coarse grid operators are constructed by direct discretization, defining the diffusion coefficient at edges of coarse cells appropriately.

Jumping Coefficients - Cell-Centered Multigrid

- Assume coefficient $k(\mathbf{x})$ is piecewise constant on the fine grid.
- The flux over an edge, dependent on the solution in the two adjacent cells, is based on the harmonic average.
- The value of the diffusion coefficient at a coarse edge between two coarse cells, is the arithmetic average of corresponding fine grid coefficients.



Jumping Coefficients - Cell-Centered Multigrid

- With L_h the fine-grid operator based on the cell-centered finite volume discretization on a uniform grid.
- And I_{2h}^h the piecewise constant prolongation operator and I_h^{2h} its adjoint.

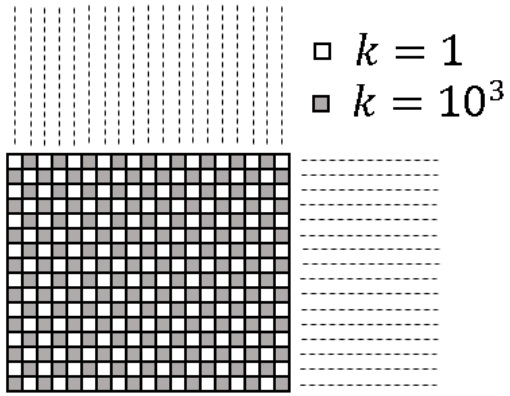
$\Rightarrow$ Then, the Galerkin coarse grid operator $L_{2h} = \frac{1}{2} I_h^{2h} L_h I_{2h}^h$ is equivalent to a direct discretization on the coarse grid, based on the arithmetic average of the corresponding fine grid coefficients.

- So, a direct coarse grid discretization is equivalent to the Galerkin approach, but computationally more efficient (Molenaar 1996, our paper).

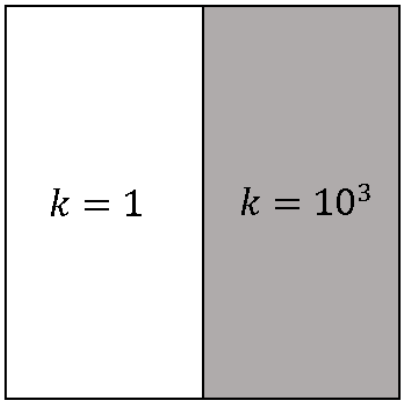
CWI Centrum Wiskunde & Informatica

Jumping Coefficients - Test Cases

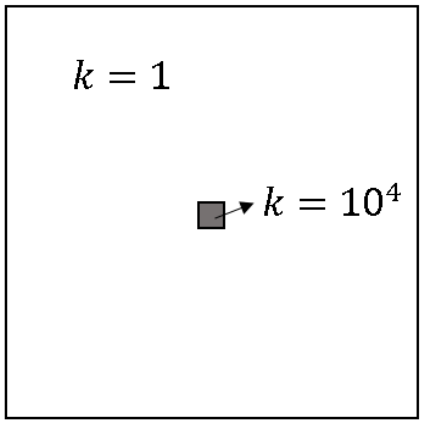
- The following jumping coefficient benchmark problems, characterized by the distribution of the diffusion coefficient, are considered here:

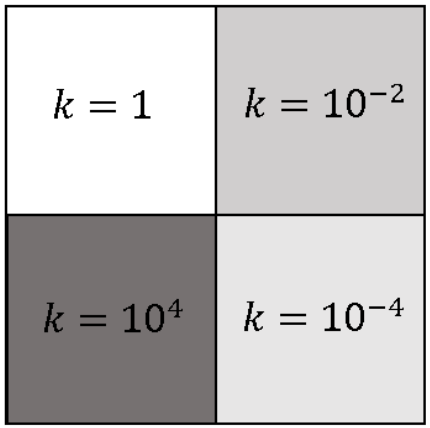


(a) Chessboard



(b) Vertical jump





Jumping Coefficients - Test Cases

$\Rightarrow$ Lexicographic Gauss-Seidel smoother

			$W(1,0)$	$W(1,1)$	$W(2,2)$
Chessboard	GS	(CP,CR)	0.44 (0.45)	0.20 (0.20)	0.04 (0.04)
Vertical jump	GS	(CP,CR)	0.44 (0.45)	0.19 (0.20)	0.05 (0.04)
Inhomogeneous square	GS	(CP,CR)	0.47 (0.54)	0.26 (0.30)	0.15 (0.16)
Four corner problem	GS	(CP,CR)	0.42 (0.44)	0.17 (0.17)	0.04 (0.04)

Table: Asymptotic two-grid convergence factors predicted by LFA and the corresponding computed average multigrid convergence factors using W-cycles, in parenthesis, for the four examples.

CWI Centrum Wiskunde & Informatica

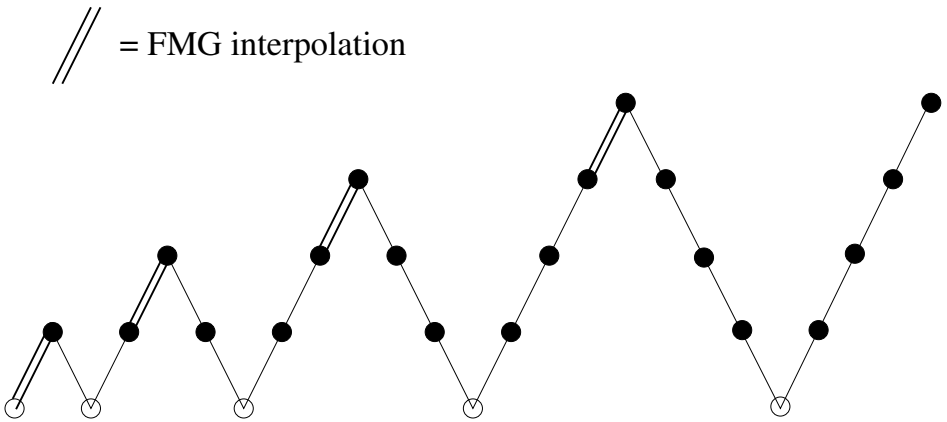
Conclusions

- Multigrid for problems with random coefficients
- Basic introduction (reminder) of multigrid
- Smoothing and coarse grid correction
- MG components for jumping coefficients
- Cell-centered multigrid (from Delft)

$\Rightarrow$ Prashant Kumar will show the excellent MG convergence for problems with random coefficients.

CWI Centrum Wiskunde & Informatica

Structure of Full Multigrid



- It can be proved that at most one V(1,1) iteration is needed to obtain an approximation with $O(h^2)$ accuracy for Poisson's equation.
- Example with known analytic solution.
- Solution is obtained within 120 milliseconds on a 256^2 grid.

Grid	$\|\|u - u_h\|\|_\infty$	V(0,1)	V(1,1)	F(0,1)	F(1,1)
32^2	.31(-5)	.26(-4)	.47(-5)	.86(-5)	.32(-5)
64^2	.77(-6)	.83(-5)	.12(-5)	.13(-5)	.77(-6)
128^2	.19(-6)	.27(-5)	.31(-6)	.20(-6)	.19(-6)
256^2	.48(-7)	.87(-6)	.78(-7)	.48(-7)	.48(-7)

CWI Centrum Wiskunde & Informatica

Jumping Coefficients - Cell-Centered Multigrid

Let L_h be the fine-grid operator based on the cell-centered finite volume discretization of problem (1) on a uniform grid of mesh size $h = \ell/M$ with M even. Let P_{2h}^h be the piecewise constant prolongation operator and R_h^{2h} its adjoint. Then, the Galerkin coarse grid operator $L_{2h} = \frac{1}{2} R_h^{2h} L_h P_{2h}^h$ is equivalent to a direct discretization on the coarse grid based on the arithmetic average of the corresponding fine grid coefficients.

Proof.

We prove the equivalence for a coarse grid cell $D_{2h}^{i_1,i_2}$ such that none of its edges lies on the boundary of the domain. The equivalence for coarse cells close to boundaries with Dirichlet or Neumann boundary conditions can be proven similarly. By applying the restriction operator R_h^{2h} in (**??**), the equation associated with the cell $D_{2h}^{i_1,i_2}$ by using the Galerkin approach is given by

$$\begin{aligned} \frac{1}{2}(R_h^{2h} L_h P_{2h}^h u)_{i_1,i_2} &= \frac{1}{8}((L_h P_{2h}^h u)_{2i_1,2i_2} + (L_h P_{2h}^h u)_{2i_1-1,2i_2} + (L_h P_{2h}^h u)_{2i_1,2i_2-1} + \\ &\quad (L_h P_{2h}^h u)_{2i_1-1,2i_2-1}). \end{aligned} \tag{5}$$

Taking into account that the prolongation operator is piecewise constant, we obtain